MIS Quarterly

RESEARCH NOTE

DOES PLS HAVE ADVANTAGES FOR SMALL SAMPLE SIZE OR NON-NORMAL DATA?

Dale L. Goodhue
Terry College of Business, MIS Department, University of Georgia,
Athens, GA 30606 U.S.A. {dgoodhue@terry.uga.edu}

William Lewis
{william.w.lewis@gmail.com}

Ron Thompson
Schools of Business, Wake Forest University,
Winston-Salem, NC 27109 U.S.A. {thompsrl@wfu.edu}

Table of Contents

Appendix A. Results for Studies

Table A1. Study 1: Path Estimates for Simple Model (3 Indicators)
Table A2. Study 1: Power Results for Simple Model (3 Indicators)
Table A3. Study 1: R^2 In the Simple Model (3 Indicators)
Table A4. Study 2: Path Estimates for Simple Model, Non-Normal Data
Table A5. Study 2: Power Results for Simple Model, Non-Normal Data
Table A6. Study 3(a): Path Estimates for Simple Model (6 Indicators)
Table A7. Study 3(a): Power Results for Simple Model (6 Indicators)
Table A8. Study 3(b): Path Estimates for Simple Model (3 Indicators, Lower Loadings)
Table A9. Study 3(b): Power Results for Simple Model (3 Indicators, Lower Loadings)
Table A10. Study 4: Path Estimates for More Complex Model
Table A11. Study 4: Power Results for More Complex Model
Table A12. Study 4: Power Results for Complex Model and Non-Normal Data

Appendix A

Results for Studies

Table A1. Study 1: Path Estimates for Simple Model (3 Indicators)*					
Gamma1: Large Effect Size					
n =	**20**	**40**	**90**	**150**	**200**
Actual	***.480***	***.480***	***.480***	***.480***	***.480***
Regression	.383	.398	.391	.393	.392
PLS	.388	.403	.399	.400	.399
LISREL	.456†	.486†	.485	.484	.484
Gamma2: Medium Effect Size					
n =	**20**	**40**	**90**	**150**	**200**
Actual	***.314***	***.314***	***.314***	***.314***	***.314***
Regression	.258	.255	.258	.254	.256
PLS	.262	.273	.270	.263	.262
LISREL	.290†	.317†	.320	.315	.314
Gamma3: Effect Size					
n =	**20**	**40**	**90**	**150**	**200**
Actual	***.114***	***.114***	***.114***	***.114***	***.114***
Regression	.081	.099	.099	.096	.089
PLS	.090	.106	.115	.107	.099
LISREL	.107†	.116†	.120	.119	.110
Gamma4: No Effect					
n =	**20**	**40**	**90**	**150**	**200**
Actual	***.000***	***.000***	***.000***	***.000***	***.000***
Regression	-.010	-.009	-.001	-.002	-.002
PLS	-.011	-.017	.003	-.002	-.003
LISREL	-.012	-.010	.000	-.004	-.003

*Each construct is measured with three indicators, with loadings of .70, .80, and .90.
†N = 20 or N = 40 is far below any recommended minimum sample size for LISREL.

Table A2. Study 1: Power Results for Simple Model (Three Indicators)

Gamma1: Large Effect Size					
n =	**20**	**40**	**90**	**150**	**200**
Predicted*	***<.46‡***	***.78***	***.99***	***.99***	***.99***
95% C.I.	N/A	(.74, .82)	(.98, 1.0)	(.98, 1.0)	(.98, 1.0)
Regression	.39	.75	.99	1.00	1.00
PLS	.31	.76	.98	1.00	1.00
LISREL	.18†	.69†	.98	1.00	1.00
Gamma2: Medium Effect Size					
n =	**20**	**40**	**90**	**150**	**200**
Predicted*	***<.25‡***	***.44***	***.81***	***.96***	***.99***
95% C.I.	N/A	(.40, .48)	(.78, .84)	(.94, .98)	(.98, 1.0)
Regression	.22	.40	.76	.92	.97
PLS	.15	.41	.75	.92	.97
LISREL	.10†	.37†	.79	.95	.98
Gamma3: Small Effect Size					
n =	**20**	**40**	**90**	**150**	**200**
Predicted*	***< .25‡***	***< .25‡***	***< .25‡***	***< .25‡***	***.33***
95% C.I.	N/A	N/A	N/A	N/A	(.29, .37)
Regression	.07	.09	.16	.27	.29
PLS	.04	.10	.18	.28	.29
LISREL	.06†	.11†	.19	.28	.31
Gamma4: No Effect					
n =	**20**	**40**	**90**	**150**	**200**
Allowable	***.05***	***.05***	***.05***	***.05***	***.05***
95% C.I.	(.03, .07)	(.03, .07)	(.03, .07)	(.03, .07)	(.03, .07)
Regression	.05	.06	.05	.04	.03
PLS	.04	.06	.05	.05	.04
LISREL	.05†	.08†	.05	.04	.04

*The predicted values for the power results are for normally distributed data, and are derived from Cohen (1988).
†N = 20 or N = 40 is far below any recommended minimum sample size for LISREL.
‡Cohen's tables do not go to this low of a sample size.

Table A3. Study 1: R^2 in the Simple Model (Three Indicators)

	Regression		PLS		LISREL	
Sample Size	R^{2*}	Adjusted R^{2**}	R^{2*}	Adjusted R^{2**}	R^{2*}	Adjusted R^{2**}
20	.37	.20	.50	.36	.52	.39
40	.31	.23	.38	.31	.45	.39
90	.26	.22	.30	.26	.39	.36
150	.25	.23	.27	.25	.37	.35
200	.24	.23	.26	.25	.36	.35

*While regression users typically report adjusted R^2 values, users of PLS or LISREL typically report the only R^2 value produced by the software package, the unadjusted R^2.
**Adjusted R^2 for PLS and LISREL was calculated using the formula:

$$1 - (1 - R^2)\frac{n - i}{n - p}$$

where i is one if there is an intercept or zero if there is not an intercept; *p* is the total number of regressors in the model (excluding the constant term) and *n* is the sample size (SAS User's Guide: Statistics, Version 5, 1985, p. 715).

Table A4. Study 2: Path Estimates for Non-Normal Data (Normal, Average Skew, Moderate Skew, Extreme Skew, High Negative Kurtosis)

Gamma1: Large Effect Size										
n =	40					90				
Actual*	.480					.480				
	Normal	Ave. Skew	Mod. Skew	Extrm Skew	Negative Kurt.	Normal	Ave. Skew	Mod. Skew	Extrm Skew	Negative Kurt.
Reg.	.398	.380	.379	.330	.360	.391	.390	.393	.334	.368
PLS	.403	.391	.390	.351	.375	.399	.401	.399	.350	.376
LISREL	.486*	.478*	.500*	.407*	.455*	.485	.487	.484	.428	.461
Gamma2: Medium Effect Size										
n =	40					90				
Actual*	.314					.314				
	Normal	Ave. Skew	Mod. Skew	Extrm Skew	Negative Kurt.	Normal	Ave. Skew	Mod. Skew	Extrm Skew	Negative Kurt.
Reg.	.255	.259	.249	.208	.246	.258	.253	.255	.198	.246
PLS	.273	.278	.275	.233	.270	.270	.269	.267	.220	.262
LISREL	.317*	.311*	.413*	.269*	.306*	.320	.321	.312	.254	.312
Gamma3: Small Effect Size										
n =	40					90				
Actual*	.114					.114				
	Normal	Ave. Skew	Mod. Skew	Extrm Skew	Negative Kurt.	Normal	Ave. Skew	Mod. Skew	Extrm Skew	Negative Kurt.
Reg.	.099	.089	.096	.075	.084	.099	.098	.089	.099	.068
PLS	.106	.107	.107	.088	.090	.115	.102	.112	.080	.099
LISREL	.116*	.111*	.072*	.107*	.106*	.120	.110	.125	.120	.085
Gamma4: No Effect										
n =	40					90				
Actual*	.000					.000				
	Normal	Ave. Skew	Mod.	Extrm Skew	Negative Kurt.	Normal	Ave. Skew	Mod. Skew	Extrm Skew	Negative Kurt.
Reg.	-.009	.002	-.008	.004	-.002	-.001	.002	.001	.000	-.011
PLS	-.017	.006	-.005	-.004	.002	.003	.002	.005	-.002	-.011
LISREL	-.010*	-.005*	.058*	.005*	-.004*	.000	.001	.005	-.004	-.009

*N = 40 is far below any recommended minimum sample size for LISREL.

Table A5. Study 2: Power Results for Non-Normal Data (Normal, Average Skew, Moderate Skew, Extreme Skew, High Negative Kurtosis)

Gamma1: Large Effect Size										
n =	40					90				
Predicted†	.78 (for normal distributions)					.99 (for normal distributions)				
95% C.I.	(.74, .82)					(.98, 1.00)				
	Normal	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**	**Normal**	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**
Reg.	.75	.70	.67	.53	.65	.99	.97	.97	.87	.97
PLS	.76	.71	.68	.51	.65	.98	.97	.98	.88	.97
LISREL	.69*	.64*	.62*	.39*	.54*	.98	.97	.98	.87	.97
Gamma2: Medium Effect Size										
n =	**40**					**90**				
Predicted†	.44 (for normal distributions)					.80 (for normal distributions)				
95% C.I.	(.40, .48)					(.77, .85)				
	Normal	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**	**Normal**	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**
Reg.	.40	.39	.37	.28	.36	.76	.74	.76	.53	.71
PLS	.41	.42	.39	.24	.39	.75	.76	.74	.48	.71
LISREL	.37*	.40*	.36*	.22*	.33*	.79	.78	.75	.50	.73
Gamma3: Small Effect Size										
n =	**40**					**90**				
Predicted†	< .25					< .25				
95% C.I.	N/A					N/A				
	Normal	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**	**Normal**	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**
Reg.	.09	.09	.11	.07	.07	.16	.17	.15	.12	.14
PLS	.10	.13	.10	.07	.08	.18	.17	.19	.12	.15
LISREL	.11*	.10*	.09*	.08*	.08*	.19	.16	.20	.11	.15
Gamma4: No Effect										
n =	**40**					**90**				
Allowable	.05					.05				
95% C.I.	(.03, .07)					(.03, .07				
	Normal	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**	**Normal**	**Ave. Skew**	**Mod. Skew**	**Extrm Skew**	**Negative Kurt.**
Reg.	.06	.05	.04	.05	.06	.05	.04	.06	.05	.06
PLS	.06	.07	.06	.06	.07	.05	.07	.06	.06	.06
LISREL	.08*	.07*	.05*	.06*	.06*	.05	.07	.05	.06	.06

*N = 40 is far below any recommended minimum sample size for LISREL.

†All of the predicted values for the power are for normally distributed data, and are derived from Cohen (1988). Obviously, the predicted value may not apply for non-normal data.

Table A6. Study 3(a): Path Estimates for Simple Model (Six Indicators)*

Gamma1: Large Effect Size				
n =	**40**	**90**	**150**	**200**
Actual	***.480***	***.480***	***.480***	***.480***
Regression	.432	.431	.435	.427
PLS	.431	.432	.438	.430
LISREL	.482†	.482	.486	.477
Gamma2: Medium Effect Size				
n =	**40**	**90**	**150**	**200**
Actual	***.314***	***.314***	***.314***	***.314***
Regression	.283	.280	.280	.281
PLS	.297	.291	.287	.287
LISREL	.317†	.314	.312	.314
Gamma3: Small Effect Size				
n =	**40**	**90**	**150**	**200**
Actual	***.114***	***.114***	***.114***	***.114***
Regression	.099	.104	.100	.106
PLS	.113	.121	.115	.121
LISREL	.112†	.118	.111	.118
Gamma4: No Effect				
n =	**40**	**90**	**150**	**200**
Actual	***.000***	***.000***	***.000***	***.000***
Regression	.006	-.002	.001	-.001
PLS	.004	-.001	-.004	-.005
LISREL	.009†	.002	.001	.009

*Each construct is measured with six indicators: two with loadings set at .70, two at .80, and two at .90.
†N = 40 is far below any recommended minimum sample size for LISREL.

Table A7. Study 3(a): Power Results for Simple Model (Six Indicators)*				
Gamma1: Large Effect Size				
n =	**40**	**90**	**150**	**200**
Regression	.84	.99	1.00	1.00
PLS	.85	.99	1.00	1.00
LISREL	.81†	.99	1.00	1.00
Gamma2: Medium Effect Size				
n =	**40**	**90**	**150**	**200**
Regression	.47	.85	.98	1.00
PLS	.48	.86	.98	.99
LISREL	.45†	.87	.98	1.00
Gamma3: Small Effect Size				
n =	**40**	**90**	**150**	**200**
Regression	.10	.20	.31	.40
PLS	.12	.20	.30	.40
LISREL	.09†	.20	.30	.42
Gamma4: No Effect				
n =	**40**	**90**	**150**	**200**
Allowable	***.05***	***.05***	***.05***	***.05***
95% C.I.	**(.03, .07)**	**(.03, .07)**	**(.03, .07)**	**(.03, .07)**
Regression	.05	.06	.06	.04
PLS	.06	.06	.05	.04
LISREL	.05†	.05	.06	.04

*Each construct is measured with six indicators: two with loadings set at .70, two at .80, and two at .90.
†N = 40 is far below any recommended minimum sample size for LISREL.

Table A8. Study 3(b): Path Estimates for Simple Model (3 Indicators, Lower Loadings)*				
Gamma1: Large Effect Size				
n =	**40**	**90**	**150**	**200**
Actual	***.480***	***.480***	***.480***	***.480***
Regression	.352	.341	.347	.346
PLS	.386	.355	.357	.354
LISREL	.601†	.487	.490	.486
Gamma2: Medium Effect Size				
n =	**40**	**90**	**150**	**200**
Actual	***.314***	***.314***	***.314***	***.314***
Regression	.222	.226	.224	.225
PLS	.245	.248	.238	.236
LISREL	.202†	.322	.318	.315
Gamma3: Small Effect Size				
n =	**40**	**90**	**150**	**200**
Actual	***.114***	***.114***	***.114***	***.114***
Regression	.085	.088	.085	.077
PLS	.100	.108	.101	.089
LISREL	.176*	.121	.120	.109
Gamma4: No Effect				
n =	**40**	**90**	**150**	**200**
Actual	***.000***	***.000***	***.000***	***.000***
Regression	-.007	-.001	.000	-.002
PLS	.000	.001	.002	.000
LISREL	.066†	-.001	.003	-.003

*Each construct is measured with three indicators, with loadings of .60, .70 and .80.
†N = 40 is far below any recommended minimum sample size for LISREL.

Table A9. Study 3(b): Power Results for Simple Model (3 Indicators, Lower Loadings)*				
Gamma1: Large Effect Size				
n =	**40**	**90**	**150**	**200**
Regression	.63	.93	1.00	1.00
PLS	.63	.93	1.00	1.00
LISREL	.41[†]	.92	.99	1.00
Gamma2: Medium Effect Size				
n =	**40**	**90**	**150**	**200**
Regression	.29	.64	.83	.93
PLS	.30	.64	.83	.92
LISREL	.18[†]	.62	.83	.93
Gamma3: Small Effect Size				
n =	**40**	**90**	**150**	**200**
Regression	.07	.13	.20	.20
PLS	.09	.14	.21	.21
LISREL	.04[*]	.12	.20	.21
Gamma4: No Effect				
n =	**40**	**90**	**150**	**200**
Allowable	***.05***	***.05***	***.05***	***.05***
95% C.I.	**(.03, .07)**	**(.03, .07)**	**(.03, .07)**	**(.03, .07)**
Regression	.06	.04	.04	.03
PLS	.05	.07	.05	.04
LISREL	.05[†]	.04	.04	.04

*Each construct is measured with three indicators, with loadings of .60, .70 and .80.
[†]N = 40 is far below any recommended minimum sample size for LISREL.

Table A10. Study 4: Path Estimates for More Complex Model												
n =	20			40			90			150		
	Reg.	PLS	LISREL	Reg.	PLS	LISREL	Reg.	PLS	LISREL	Reg.	PLS	LISREL
E1E4	.005	-.006	-.004†	.021	.020	.009†	.010	.010	-.003	.003	.002	-.012
E2E4	.254	.248	.247†	.262	.267	.287†	.254	.258	.289	.254	.258	.286
E3E4	.280	.282	.288†	.254	.263	.311†	.254	.273	.322	.263	.269	.319
K1E1	.278	.325	.307†	.270	.301	.310†	.277	.299	.326	.275	.288	.323
K2E2	.318	.371	.360†	.303	.338	.356†	.303	.319	.359	.301	.312	.356
K3E3	.327	.377	.357†	.328	.367	.383†	.334	.354	.395	.332	.345	.395
K1E4	.215	.221	.212†	.219	.225	.259†	.218	.222	.261	.224	.230	.268
K2E4	.283	.289	.283†	.265	.270	.299†	.269	.274	.305	.275	.278	.312
K3E4	.006	.002	.001†	.018	.016	-.004†	.020	.018	.002	.021	.020	.003

†N = 20 or n = 40 is far below any recommended minimum sample size for LISREL.

Table A11. Study 4: Power Results for More Complex Model												
n =	20			40			90			150		
	Reg.	PLS	LISREL	Reg.	PLS	LISREL	Reg.	PLS	LISREL	Reg.	PLS	LISREL
E1E4	0.050	0.016	0.110†	0.062	0.052	0.084†	0.066	0.080	0.078	0.054	0.056	0.056
E2E4	0.196	0.056	0.166†	0.376	0.372	0.400†	0.734	0.756	0.730	0.928	0.928	0.904
E3E4	0.216	0.074	0.160†	0.374	0.350	0.426†	0.788	0.778	0.808	0.928	0.932	0.936
K1E1	0.238	0.290	0.108†	0.392	0.458	0.392†	0.764	0.780	0.764	0.940	0.946	0.946
K2E2	0.294	0.396	0.128†	0.468	0.554	0.468†	0.846	0.880	0.866	0.966	0.970	0.974
K3E3	0.318	0.370	0.136†	0.546	0.588	0.526†	0.930	0.938	0.928	0.992	0.990	0.994
K1E4	0.146	0.056	0.154†	0.312	0.306	0.374†	0.610	0.606	0.656	0.886	0.890	0.900
K2E4	0.214	0.092	0.180†	0.414	0.380	0.422†	0.788	0.790	0.750	0.950	0.950	0.938
K3E4	0.066	0.014	0.106†	0.072	0.062	0.096†	0.052	0.070	0.064	0.074	0.066	0.074
Avg. 7 True	0.232	0.191	0.147†	0.412	0.430	0.430†	0.780	0.790	0.786	0.941	0.944	0.942
Avg. 2 False	0.058	0.015	0.108†	0.067	0.057	0.090†	0.059	0.075	0.071	0.064	0.061	0.065

† N = 20 or n = 40 is far below any recommended minimum sample size for LISREL.

Table A12. Study 4: Complex Model With Non-Normal Data: Power

Comparing the Drop in Power Due to Extremely Skewed Data (Averaged Across 7 True Paths)
With the Drop in Power Due to Extremely Skewed Data in the Simple Model

Complex Model	Average Power Across Seven True Paths					
	N = 40			N = 90		
	LISREL	PLS	Regression	LISREL	PLS	Regression
Normal Distribution	0.43†	0.43	0.41	0.79	0.79	0.78
Extreme Skew (S = 1.8; K = 3.8)	0.24†	0.26	0.28	0.52	0.51	0.59
Reduction in Power Due to Non-Normal Data	0.19†	0.17	0.13	0.27	0.28	0.19

Simple Model	Power for Medium Effect Size					
	N = 40			N = 90		
	LISREL	PLS	Regression	LISREL	PLS	Regression
Normal Distribution	0.37†	0.41	0.40	0.79	0.75	0.76
Extreme Skew (S = 1.8; K = 3.8)	0.22†	0.24	0.28	0.50	0.48	0.53
Reduction in Power Due to Non-Normal Data	0.15†	0.17	0.12	0.29	0.27	0.24

†N = 40 is far below any recommended minimum sample size for LISREL.

Appendix B

Use of 100 Bootstrapping Resamples

A question could be asked concerning our decision to use 100 bootstrapping resamples for each PLS analysis, rather than a larger number such as 500 or 1,000. It is true that Chin (1998) recommends using 500 bootstrapping resamples to determine statistical significance, and it is also true that the bootstrapping literature recommends that amount or more. The reason is that when using a small original sample, or a sample from a difficult or irregular distribution, 500 resamples gives confidence that the resulting bootstrapping distribution is close to the underlying distribution. We decided to use only 100 resamples in our analyses of each PLS run for both substantive and practical reasons.

We can explain the substantive reason by looking at the difference between two different research situations. Consider a typical researcher who was interested in reporting the significance of the Gamma2 link in Figure 1, from his or her sample of 40 questionnaires. Analogously, we are interested in reporting an estimate of the likely significance that would be found by any researcher studying Gamma2 with 40 questionnaires. Both the typical researcher and we are interested in reporting a single value: the researcher wants to report the t-statistic of a single data set of n = 40; we want to report the power achieved by the PLS analyses of 500 data sets of n = 40 each.

If that researcher reported a value based on only 100 bootstrapping resamples for their PLS analysis, there is a danger that random chance in the data analysis could distort the findings. We agree it would be better for that researcher to use 500 or 1,000 resamples to bolster confidence that the single reported value is very close to the true value.

We also want to report a single value. But we develop that single value from PLS analyses of 500 data sets, and 100 resamples for each of those. In other words, if the typical researcher used only 100 resamples for his or her data set, the reported value would be based on only 100 resamples. When we use only 100 resamples for each data set, we are also using 500 data sets. Therefore, every reported PLS power value in our study is based on 50,000 resamples.

To be very clear, in all of our tables every reported value generated by bootstrapping is based on at least 500 samples and 50,000 resamples. With this much replication, we doubt there is much likelihood the picture would change by doing 5 or 10 times as much replication (i.e., with 500 or 1,000 resamples per data set giving 250,000 or 500,000 bootstrapping resamples for each value reported in our tables). To test this assumption, we reran the 500 data sets with N = 90 from the more complex model in Figure 4, now using PLS with 1,000 bootstrapping resamples. The comparison between 100 resamples and 1,000 resamples is shown in Table B1. Each reported value in the 1,000 resamples columns of Table B1 is based on 500,000 bootstrapping resamples.

The differences between 100 and 1,000 resamples are quite small: the average difference in power is .002; the largest difference is .018. None of these changes are statistically significant either individually or as a group. Changes of this magnitude do not impact the overall results of the study. Given the above logical analysis and the spot checking displayed in Table B1, we believe an increase to 500 or 1,000 bootstrapping resamples is not necessary.

Our practical reason for using 100 resamples is the surprising amount of work that would be involved in using more resamples. Given the large numbers of runs involved in the total research project, it should be clear to the reader that this type of research would not be practical if it were not somehow automated. The automation process we used for every PLS analysis in our tables involved creating a large "deck" file of 50,500 PLS runs; for example, the 3 indicator N = 200 run (including Gamma1 through Gamma4) required us to create a PLS command file or "deck" file which is .7 gigabytes in size.

The problem with processing files this large is magnified significantly by the fact that PLS-Graph has some peculiarities that cause it to crash occasionally (especially with larger files). In general for an individual analysis this is not a problem—the same PLS run can be rerun immediately, and virtually always completes the second time. But when we execute a deck file of 50,500 PLS runs, it virtually always crashes somewhere in the process. When a crash occurs, our solution is to chop the large deck file into smaller and smaller pieces until no crashes occur, and then reassemble the results at the end. This, however, moves us away from an automated process, and to a more manual one. Increasing the number of resamples to 500 would have required the processing of deck files with 252,500 PLS runs each. Although we would have liked to use 500 or 1,000 resamples instead of 100, it simply wasn't practical to do so; we would have been unable to complete our analyses.

Table B1. Testing PLS Results with 1,000 Bootstrapping Resamples (Complex Model)					
n = 90					
	Path Estimate		**Power Results**		
Path	**100 resamples**	**1,000 resamples**	**100 resamples**	**1,000 resamples**	**Power Difference**
Eta1-Eta4	.010	.010	.080	.066	.014
Eta2-eta4	.258	.258	.756	.746	.010
Eta3-Eta4	.273	.273	.778	.790	-.012
Ksi1-Eta1	.299	.299	.780	.776	.004
Ksi2-Eta2	.319	.319	.880	.898	-.018
Ksi3-Eta3	.354	.354	.938	.930	.008
Ksi1-Ksi4	.222	.222	.606	.606	.000
Ksi2-Ksi4	.274	.274	.790	.788	.002
Ksi3-Ksi4	.018	.018	.070	.060	.010
			Average		.002

Appendix C

Detailed Presentation, Studies 2 and 3

Study 2: Simple Model and Non-Normal Data

Since much of the data used in behavioral research is not normally distributed and PLS could have an advantage with non-normal data, we tested the impact of non-normal data on the performance of the three techniques with Study 2. To determine initial realistic values of skew and kurtosis, we first used a convenience sample of four IS-related data sets that we considered as representative of IS research, and then chose more extreme values of skew and kurtosis from the literature (Fleishman 1978; Vale and Maurelli 1983). For the four representative IS data sets, we examined the distributions of the data for all of the measurement indicators used in each. From this we obtained values for what we call the "average" degree of skewness (0.63) and kurtosis (0.45) across the responses for the indicators. We then added one standard deviation to both the average skew (0.63 + 0.50 = 1.13) and the average kurtosis (0.45 + 1.16 = 1.61) to determine values for "more non-normal" distributions. These two conditions we labeled "average skew" and "moderate skew."

To test the possible influence of more extreme values, we chose the most extreme value of skew and the most extreme value of negative kurtosis from Fleishman's (1978) table of possible values of skew and kurtosis, resulting in two new conditions to test. The first, which we termed "high skew," had values of 1.8 for skew and 3.8 for kurtosis. The second, which we termed "negative kurtosis," had values of .25 for skew and -1.00 for kurtosis.

For these four new conditions of non-normality (each with the values for skew and kurtosis stated above), we used Figure 1 as our model and Fleishman's tables and his approach for generating non-normal data. Because at n = 20 and n = 150 the results are quite predictable (at n = 20 all techniques have much less than 40% power for a medium effect size; at n = 150 all techniques have power much greater than 80% for a medium effect size) we focused on n = 40 and n = 90 in this study. Once again, for each distribution and sample size combination, we generated 500 data sets, and analyzed them as we had for the normally distributed data.

Table A4 (parameter estimates) and Table A5 (power) paralleling the earlier Tables A1 and A2, are included in Appendix A. The key results for n = 40 and n = 90 can be seen in Figure 4 of the paper proper. The picture is not very different for the two sample sizes. Skew up to about 1.1 and kurtosis up to 1.6 has only a slight impact on the statistical power of the analysis, regardless of which technique is used. When skew is increased to 1.8 and kurtosis to 3.8, there is a noticeable drop in power and accuracy from what would be seen with normally distributed data. However, going to the other extreme on kurtosis with a large negative (-1.00 in our study), power and accuracy are not much reduced from the moderate skew condition.

Interestingly, there doesn't seem to be any important difference between the abilities of these three statistical techniques to handle non-normal data. The one exception to this is that for n = 40, high negative kurtosis, and a medium effect size, LISREL's power (.33) has a statistically significant drop ($p < .05$) below PLS's (.39). Since neither of these power values is even close to .80, and since no one would suggest using LISREL for such small sample sizes, this is not a surprising or interesting finding.

Study 3: Number of Indicators, Reliability

The data in Studies 1 and 2 utilized three indicators per construct with loadings of 0.70, 0.80, and 0.90. However, researchers doing field studies often have a greater number of indicators available to measure each of the constructs in the research model, and empirical work has demonstrated that the number of indicators used in the analysis can impact the path estimates of PLS (McDonald 1996). As the number of indicators used to represent a construct increases, so too does the reliability. Similarly, researchers often work with construct items that have lower loadings. It seems important, therefore, to extend our previous studies by examining the impact, if any, that changes in the number of indicators and changes in reliability have on the results.

To explore the impact of increasing the number of indicators, we generated a group of data sets using six indicators per construct (with loadings of two at .7, two at .8, and two at .9). The results for the path estimates and power are shown in Appendix A, Tables A6 and A7. The results from these tests were consistent with our previous analyses. Since the number of indicators (and therefore the overall reliability, now .91) of the construct measures increased, the path estimates and the power were slightly higher than when three indicators were used. Nevertheless, there were no changes in the relative efficacy of the three analysis techniques. Differences between the three techniques were modest and not statistically significant. As with Studies 1 and 2, LISREL produced higher and more accurate path estimates for the large and medium effect sizes at sample sizes of 90 and above; PLS generally produced estimates that were slightly higher than regression.

With respect to power, the same observations held. Increasing the number of indicators did increase the power for all techniques—for example, at n = 40 and medium effect size, regression's power went from .40 to .47; the power for PLS went from .41 to .48. However, the larger number of indicators did not alter the relative efficacy of the three techniques. Furthermore, moving to six indicators did not change the fact that the rule of 10 is not a useful guide to appropriate sample size when power is a consideration.

The reliability of each of the constructs in the Study 1 model can be considered comfortably high (0.84). To test the impact of less reliable and more diverse indicators, we generated an additional group of data sets. The indicator loadings were set at 0.6, 0.7, and 0.8 for each of the constructs, which resulted in a Cronbach's alpha of .74. We note that with lower indicator loadings and a sample size of n = 40, we begin to run the risk of "empirical underidentification" in the LISREL runs. Although the Gamma4 submodel is theoretically just identified, if one of its indicators is not significantly different from zero in a particular run, LISREL may not converge.

The results of the analyses for these data sets are shown in Table A8 for the path estimates and A9 for power, and were also consistent with previous analyses. Although the path estimates were lower and the percentage of estimates that were statistically significant declined slightly for the lower loading data set, the overall pattern of results (when compared across the three techniques or against accepted standards) remained the same.

Appendix D

SAS Program Used to Generate the Data for this Study

This appendix presents the SAS program used to generate the data from the simple model for 500 samples of n = 40. Note that we use the SAS subroutine Rannor (which returns N(0,1) values) separately for every random error variable in the model, each with its own seed. In this way, values for each indicator are generated as part of its own independent string of random numbers, each with a different starting seed.

We added 5.2 to our indicator scores to translate all of the data points to values between 0 and 10. Note also that amount of random variance added to each indicator was chosen to bring the total variance of each indicator to 1.0. Since all construct scores and random error scores have a variance of 1, we get a variance of 1 for the indicators by adding the appropriate portion of the error variance. For example, for X1 have the following variance: $(.70)^2 \times (1)^2 + (.714)^2 \times (1)^2 = 1.0$.

```
DATA Temp.One;
File 'c:\MultiThreadRandomNumbers\MT-RNums500x40.txt';
Retain Seed1 5532806 Seed2 528793 Seed3 523857 seed4 5223263 Seed5 5238837 Seed6 523852187
Seed7 52385427 Seed8 52398857 Seed9 52385797 Seed10 52273857 Seed11 52385347 Seed12 52385794
Seed13 5238537 Seed14 5285957 Seed15 5238657 Seed16 53385794 Seed17 539754437
Seed18 586085957 Seed19 533826357 Seed20 532657;
Do I = 1 to 20000;
Call Rannor(Seed1,KSI1);
Call Rannor(Seed2,X1Err);
Call Rannor(Seed3,X2Err);
Call Rannor(Seed4,X3Err);
X1 = round(.70*KSI1 + .714*X1Err + 5.2);
X2 = round(.80*KSI1 + .60 *X2Err + 5.2);
X3 = round(.90*KSI1 + .436*X3Err + 5.2);
Call Rannor(Seed5,KSI2);
Call Rannor(Seed6, X4Err);
Call Rannor(Seed7, X5Err);
Call Rannor(Seed8, X6Err);
X4 = round(.70*KSI2 + .714*X4Err + 5.2);
X5 = round(.80*KSI2 + .60 *X5Err + 5.2);
X6 = round(.90*KSI2 + .436*X6Err + 5.2);
Call Rannor(Seed9,KSI3);
Call Rannor(Seed10, X7Err);
Call Rannor(Seed11, X8Err);
Call Rannor(Seed12, X9Err);
X7 = round(.70*KSI3 + .714*X7Err + 5.2);
X8 = round(.80*KSI3 + .60 *X8Err + 5.2);
X9 = round(.90*KSI3 + .436*X9Err + 5.2);
Call Rannor(Seed13,KSI4);
Call Rannor(Seed14, X10Err);
Call Rannor(Seed15, X11Err);
Call Rannor(Seed16, X12Err);
X10 = round(.70*KSI4 + .714*X10Err + 5.2);
X11 = round(.80*KSI4 + .60 *X11Err + 5.2);
X12 = round(.90*KSI4 + .436*X12Err + 5.2);
Call Rannor(Seed17,Eta1Err);
Call Rannor(Seed18,Y1Err);
Call Rannor(Seed19,Y2Err);
Call Rannor(Seed20,Y3Err);
```

```
ETA1 = .48*KSI1 + .314*KSI2 + .114*KSI3 + .000*KSI4  + .811*Eta1Err;
Y1 = round(.70*ETA1 + .714*Y1Err + 5.2);
Y2 = round(.80*ETA1 + .60 *Y2Err + 5.2);
Y3 = round(.90*ETA1 + .036*Y3Err + 5.2);
Put Y1 3.0 Y2 3.0 Y3 3.0 X1 3.0 X2 3.0 X3 3.0 X4 3.0 X5 3.0 X6 3.0 X7 3.0 X8 3.0 X9 3.0 X10 3.0 X11 3.0 X12 3.0;
end;
```

Generating Non-Normal Data

Replacing the three lines generating values for X1, X2 and X3 on the previous page with the following will generate highly negative kurtosis data for X1NN, X2NN, and X3NN. Similar code can be used to generate the other non-normal data. Changing the values of "a" through "d" in the code according to the table in Fleishman will give data of different specified characteristics.

```
X1N = .70*KSI1N + .714*X1ErrN ;
X2N = .80*KSI1N + .60 *X2ErrN ;
X3N = .90*KSI1N + .436*X3ErrN ;

***  Negative Kurtosis Random Set 2007-1:  Skew = .25, Kurtosis = -1.00
***  Ala Fleishman, 1978, Pyschometrika, Vol 43, No 4.;

  b = 1.263412800;
  c = .0774624390;
  d = -.1000360450;
  a = -c;

X1NN = round(a + b*X1N + c*X1N**2 + d*X1N**3) + 5.2;
X2NN = round(a + b*X2N + c*X2N**2 + d*X2N**3) + 5.2;
X3NN = round(a + b*X3N + c*X3N**2 + d*X3N**3) + 5.2;
```

References

Chin, W. W. 1998. "The Partial Least Squares Approach to Structural Equation Modeling," in *Modern Methods for Business Research*, G. A. Marcoulides (ed.), London: Psychology Press, pp. 295-336.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale, N: Lawrence Erlbaum Associates.

Fleishman, A. I. 1978. "A Method for Simulating Non-Normal Distributions," *Psychometrika* (43:4), pp. 521-532.

McDonald, R. P. 1996. "Path Analysis with Composite Variables," *Multivariate Behavioral Research* (31:2), pp. 239-270.

SAS User's Guide: Statistics, Version 5, 1985. Cary, NC: SAS Institute, Inc.