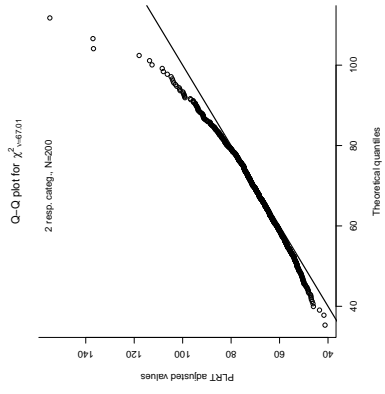


Pairwise likelihood ratio tests and model
selection criteria for structural equation
models with ordinal variables

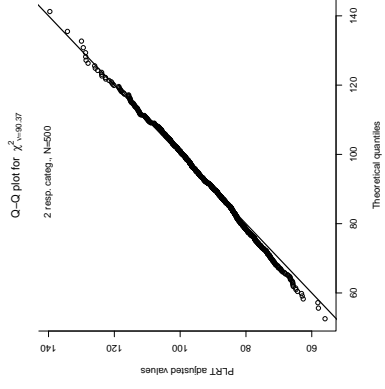
Supplementary Material

Table 1: Empirical type I error rates and average estimated adjusted degrees of freedom for the overall-fit test statistics: $PLRT_{SEM-MV}$, $T_{DWLS-MV}$ and T_{ULS-MV} , for variables with 2, 4, and 7 response categories, sample sizes 200, 500, 1000, and nominal significance levels 5% and 1%; in parenthesis 95% confidence intervals are provided only when the nominal value of type I error is not included

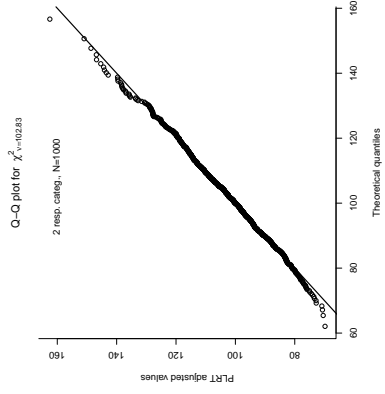
2 resp. categ.	$\alpha = 5\%$			$\alpha = 1\%$			average est. adj. df		
	200	500	1000	200	500	1000	200	500	1000
N									
$PLRT_{SEM-MV}$	0.034 (0.020, 0.048)	0.032 (0.018, 0.046)	0.041	0.005	0.004	0.011	67.01	90.37	102.83
$T_{DWLS-MV}$	0.051	0.042	0.045	0.007	0.006	0.013	74.45	99.78	112.25
T_{ULS-MV}	0.027 (0.013, 0.041)	0.036	0.045	0.004	0.004	0.013	73.52	97.94	109.85
4 resp. categ.									
$PLRT_{SEM-MV}$	0.018 (0.004, 0.032)	0.036	0.027 (0.013, 0.041)	0.002 (0, 0.008)	0.005	0.003 (0, 0.009)	47.49	68.45	80.73
$T_{DWLS-MV}$	0.047	0.053	0.040	0.005	0.010	0.007	64.50	85.29	95.59
T_{ULS-MV}	0.022 (0.008, 0.036)	0.041	0.036	0.004	0.010	0.006	62.92	82.81	92.56
7 resp. categ.									
$PLRT_{SEM-MV}$	0.023 (0.009, 0.037)	0.028 (0.014, 0.042)	0.044	0.001 (0, 0.007)	0.005	0.009	43.35	62.34	73.70
$T_{DWLS-MV}$	0.117 (0.103, 0.131)	0.073 (0.059, 0.087)	0.069 (0.055, 0.083)	0.031 (0.025, 0.037)	0.015	0.022 (0.016, 0.028)	62.93	80.50	89.26
T_{ULS-MV}	0.074 (0.060, 0.088)	0.060	0.060	0.011	0.011	0.020 (0.014, 0.026)	61.09	77.81	86.07



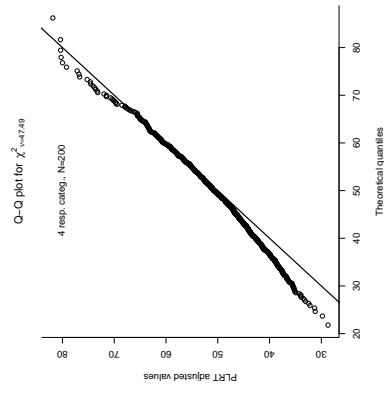
(a) 2 response categories, $N = 200$



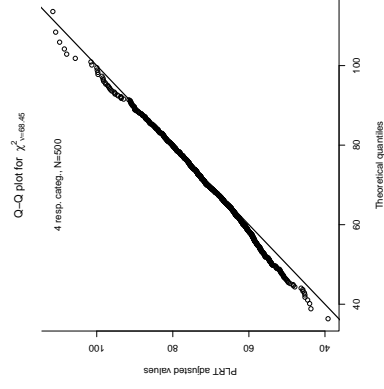
(b) 2 response categories, $N = 500$



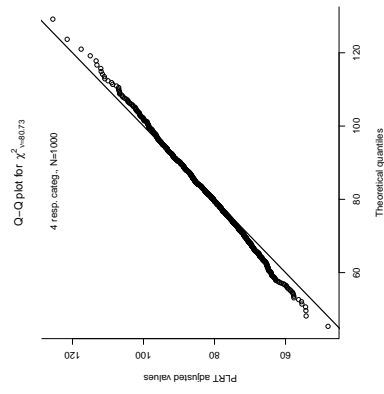
(c) 2 response categories, $N = 1000$



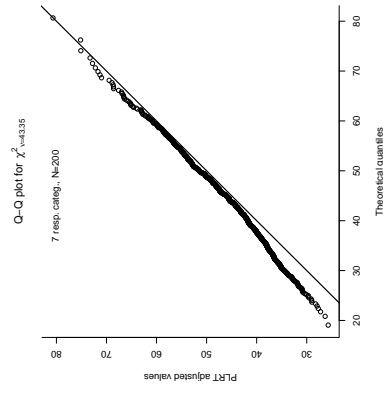
(d) 4 response categories, $N = 200$



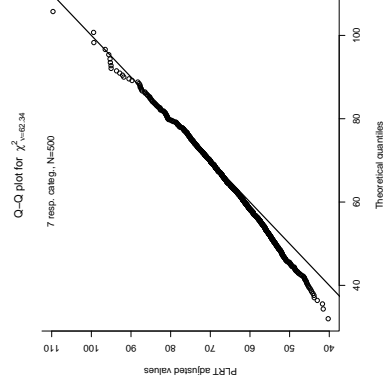
(e) 4 response categories, $N = 500$



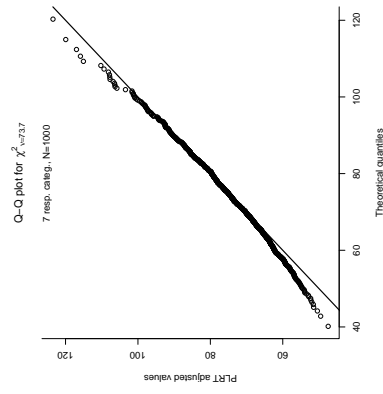
(f) 4 response categories, $N = 1000$



(g) 7 response categories, $N = 200$



(h) 7 response categories, $N = 500$



(i) 7 response categories, $N = 1000$

Figure 1: Q-Q plots for $PLRT_{SEM-MV}$ for variables with 2, 4, and 7 response categories, and sample sizes 200, 500, 1000; $PLRT_{SEM-MV}$ is compared against a chi-squared distribution with degrees of freedom equal to the average estimated adjusted degrees of freedom reported in Table 1; the solid line is the line $y = x$.

Table 2: Empirical power rates and average estimated adjusted degrees of freedom for the overall-fit test statistics: $PLRT_{SEM-MV}$, $T_{DWLS-MV}$ and T_{ULS-MV} , for variables with 2, 4, and 7 response categories, sample sizes 200, 500, 1000, and nominal significance levels 5% and 1%; the fitted model (Model 1a) misspecifies the factor correlation by fixing it equal to 0.3 while the true value is 0.4

2 resp. categ.	$a = 5\%$			$a = 1\%$			average est. adj. df		
	N	200	500	1000	200	500	1000	200	500
$PLRT_{SEM-MV}$	0.105	0.278	0.527	0.040	0.130	0.361	51.37	67.26	75.06
$T_{DWLS-MV}$	0.200	0.361	0.598	0.096	0.203	0.435	65.20	82.81	90.81
T_{ULS-MV}	0.146	0.315	0.568	0.058	0.169	0.407	64.44	81.50	89.20
4 resp. categ.									
$PLRT_{SEM-MV}$	0.130	0.438	0.734	0.051	0.253	0.588	33.03	42.89	47.52
$T_{DWLS-MV}$	0.247	0.544	0.791	0.130	0.377	0.685	49.51	59.78	63.98
T_{ULS-MV}	0.191	0.502	0.779	0.100	0.336	0.651	48.87	58.83	62.87
7 resp. categ.									
$PLRT_{SEM-MV}$	0.186	0.483	0.798	0.064	0.320	0.656	28.35	35.71	39.17
$T_{DWLS-MV}$	0.405	0.594	0.858	0.258	0.466	0.764	46.54	52.38	54.81
T_{ULS-MV}	0.338	0.562	0.838	0.185	0.423	0.736	45.83	51.68	54.06

Table 3: Empirical power rates and average estimated adjusted degrees of freedom for the overall-fit test statistics: $PLRT_{SEM-MV}$, $T_{DWLS-MV}$ and T_{ULS-MV} , for variables with 2, 4, and 7 response categories, sample sizes 200, 500, and nominal significance levels 5% and 1%; the fitted model (Model 1b) misspecifies the factor correlation by fixing it equal to 0 while the true value is 0.4

2 resp. categ.	$a = 5\%$		$a = 1\%$		average est. adj. df		
	N	200	500	200	500	200	500
$PLRT_{SEM-MV}$		0.956	1	0.901	1	48.43	61.96
$T_{DWLS-MV}$		0.969	1	0.936	1	60.29	74.18
T_{ULS-MV}		0.963	1	0.918	1	58.81	71.93
4 resp. categ.							
$PLRT_{SEM-MV}$		0.995	1	0.988	1	30.55	38.28
$T_{DWLS-MV}$		0.997	1	0.994	1	43.68	51.04
T_{ULS-MV}		0.996	1	0.994	1	42.59	49.38
7 resp. categ.							
$PLRT_{SEM-MV}$		0.992	1	0.984	1	26.25	31.96
$T_{DWLS-MV}$		0.998	1	0.994	1	40.58	44.10
T_{ULS-MV}		0.997	1	0.991	1	39.44	42.71

Table 4: Empirical power rates and average estimated adjusted degrees of freedom for the overall-fit test statistics: $PLRT_{SEM-MV}$, $T_{DWLS-MV}$ and T_{ULS-MV} , for variables with 4 response categories, sample sizes 200, 500, 1000, and significance levels 5% and 1%; the fitted model (Model 0) misspecifies three loadings by fixing them equal to 0 while their true value is 0.2

N	$a = 5\%$			$a = 1\%$			average estimated df		
	200	500	1000	200	500	1000	200	500	1000
$PLRT_{SEM-MV}$	0.127	0.597	0.974	0.023	0.336	0.899	44.70	65.27	77.19
$T_{DWLS-MV}$	0.251	0.710	0.983	0.081	0.474	0.933	65.97	88.16	99.50
T_{ULS-MV}	0.153	0.609	0.968	0.041	0.364	0.882	64.27	85.48	96.26

Table 5: Empirical type I error rates and average estimated adjusted degrees of freedom for the test statistics: $PLRT_{MV}$, $T_{DWLS-MV}$ and T_{ULS-MV} , testing nested models (Model 2 vs Model 0) for variables with 4 response categories, sample sizes 200, 500, and nominal significance levels 5% and 1%; Model 2 allows three loadings to be estimated which are correctly fixed to 0 in Model 0; in parenthesis 95% confidence intervals are provided only when the nominal value of type I error is not included

Model 2 vs Model 0	$a = 5\%$		$a = 1\%$		average est. df	
N	200	500	200	500	200	500
$PLRT_{MV}$	0.031 (0.017, 0.045)	0.049	0.005	0.010	2.85	2.94
$T_{DWLS-MV}$	0.059	0.057	0.012	0.013	2.90	2.96
T_{ULS-MV}	0.055	0.054	0.010	0.013	2.89	2.95

Table 6: Empirical type I error rates for the test statistics: $PLRT_{MV}$, $T_{DWLS-MV}$ and T_{ULS-MV} , testing two-group nested models (Models A, B, and C) for variables with 4 response categories, sample sizes 200, 500, 1000, and nominal significance levels 5% and 1%; Model A is the unconstrained model, Model B is the loading-invariant one, and Model C is the threshold- and loading-invariant model; in parenthesis 95% confidence intervals are provided only when the nominal value of type I error is not included

Model B versus A	$a = 5\%$			$a = 1\%$		
group size	200	500	1000	200	500	1000
$PLRT_{MV}$	0.017 (0.004, 0.031)	0.028 (0.014, 0.042)	0.040	0.001 (0, 0.007)	0.008	0.006
$T_{DWLS-MV}$	0.054	0.049	0.047	0.013	0.011	0.007
T_{ULS-MV}	0.058	0.050	0.055	0.012	0.012	0.009
<hr/>						
Model C versus B						
$PLRT_{MV}$	0.018 (0.005, 0.032)	0.047	0.051	0.002 (0, 0.008)	0.004	0.011
$T_{DWLS-MV}$	0.051	0.044	0.040	0.010	0.010	0.007
T_{ULS-MV}	0.053	0.055	0.042	0.010	0.006	0.008

R commands in Lavaan

The commands to estimate a SEM for ordinal variables with PL, DWLS, ULS and to get the mean-and-variance adjusted test statistics for overall fit under DWLS and ULS are

```
fitPL <- sem(model=Model, data=ord_data, estimator="PML")
```

```
fitDWLS <- sem(model=Model, data=ord_data, estimator="DWLS",
se="robust", test="Satterthwaite")
```

```
fitULS <- sem(model=Model, data=ord_data, estimator="ULS",
se="robust", test="Satterthwaite")
```

where `Model` is the model fitted to the data saved as `ord_data`. The objects `fitPL`, `fitDWLS`, `fitULS` include the estimates, standard errors, and mean-and-variance adjusted test statistics for overall fit. In `fitPL`, the test statistic $PLRT_{SEM-MV}$ where the thresholds are treated as nuisance parameter is printed by default. When parametric structure is assumed on thresholds as well, `lavaan:::ctr_pml_plrt2(fitPL)` should be used to obtain the right overall fit test statistic.

To compare nested models, the command `lavTestLRT()` with the models to be compared as input should be used. The output of this command includes the AIC_{PL} and BIC_{PL} values of the compared models when `estimator="PML"`.

To obtain the AIC_{PL} and BIC_{PL} use the command `lavaan:::ctr_pml_aic_bic()` with the fitted model as input.