1. PPP Method and PP-KS method

Table 1 False Alarm Rate in Correct One-Factor CFA Model

	K =6			K=12		
	N=50	N=500	N=2000	N=50	N=500	N=2000
KS0.3	0.3813	0.4247	0.2793	0.2589	0.3608	0.2408
PPP0.1	0.2140	0.2013	0.1113	0.1027	0.1332	0.0647
KS0.5	0.1967	0.1720	0.0953	0.0908	0.1050	0.0491
PPP0.05	0.1667	0.1193	0.0660	0.0679	0.0636	0.0277

Table 2 Detection Rate for Local Misfit in Underspecified Two-factor CFA model

	K = 12		K=24	
	N=50	N=500	N=50	N=500
KS0.3	0.3577	0.5539	0.3764	0.5791
PPP0.1	0.1748	0.3782	0.1698	0.3917
KS0.5	0.1609	0.3459	0.1509	0.3603
PPP0.05	0.1250	0.2967	0.1042	0.3105

$$KS0.3 = \frac{\sum_{i=1}^{C}\sum_{r=1}^{R}(\rho_{ir} > 0.3)}{CR}$$

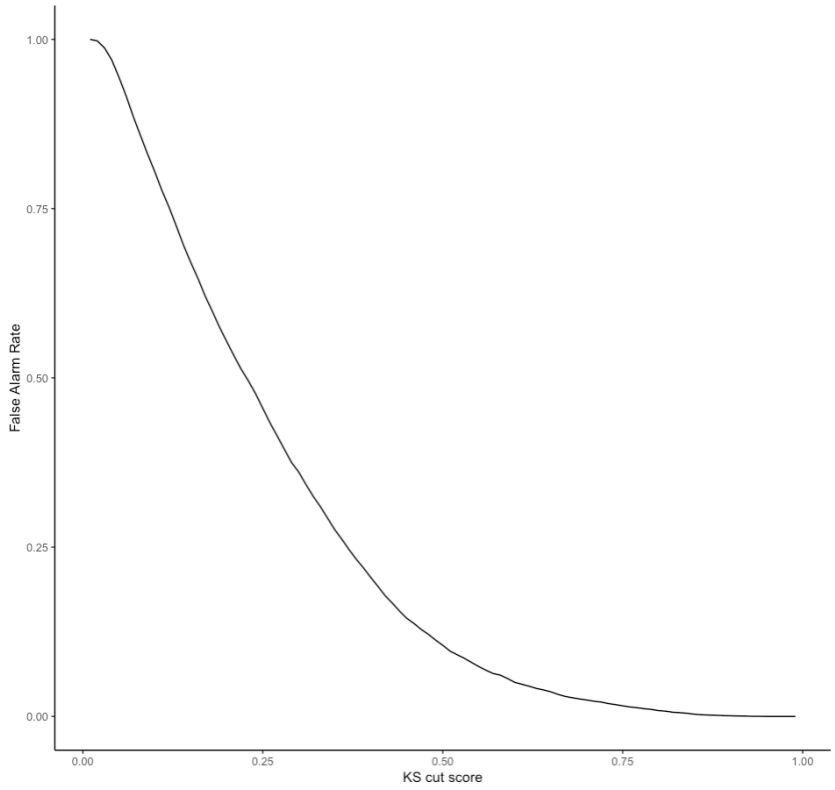
where i is the item pair and r is the repetition. ρ_{ir} is the predictive posterior KS test for correct model with saturated model for *I* item pair correlation in *r* iteration.

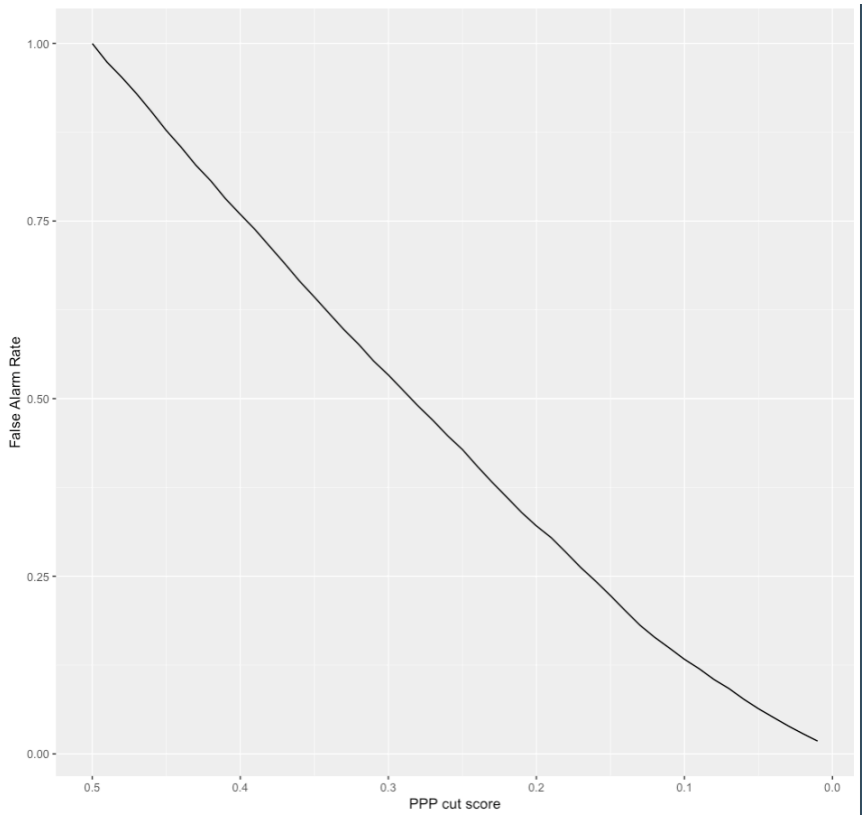
$$PPP0.05 = \frac{\sum_{i=1}^{C}\sum_{r=1}^{R}(P_{ir} > 0.95 \cup P_{ir} < 0.05)}{CR}$$

Where P_{ir} is the predictive posterior percentile for *i* item pair correlation in *r* iteration.

1.1 False Alarm Rate for Different cut scores

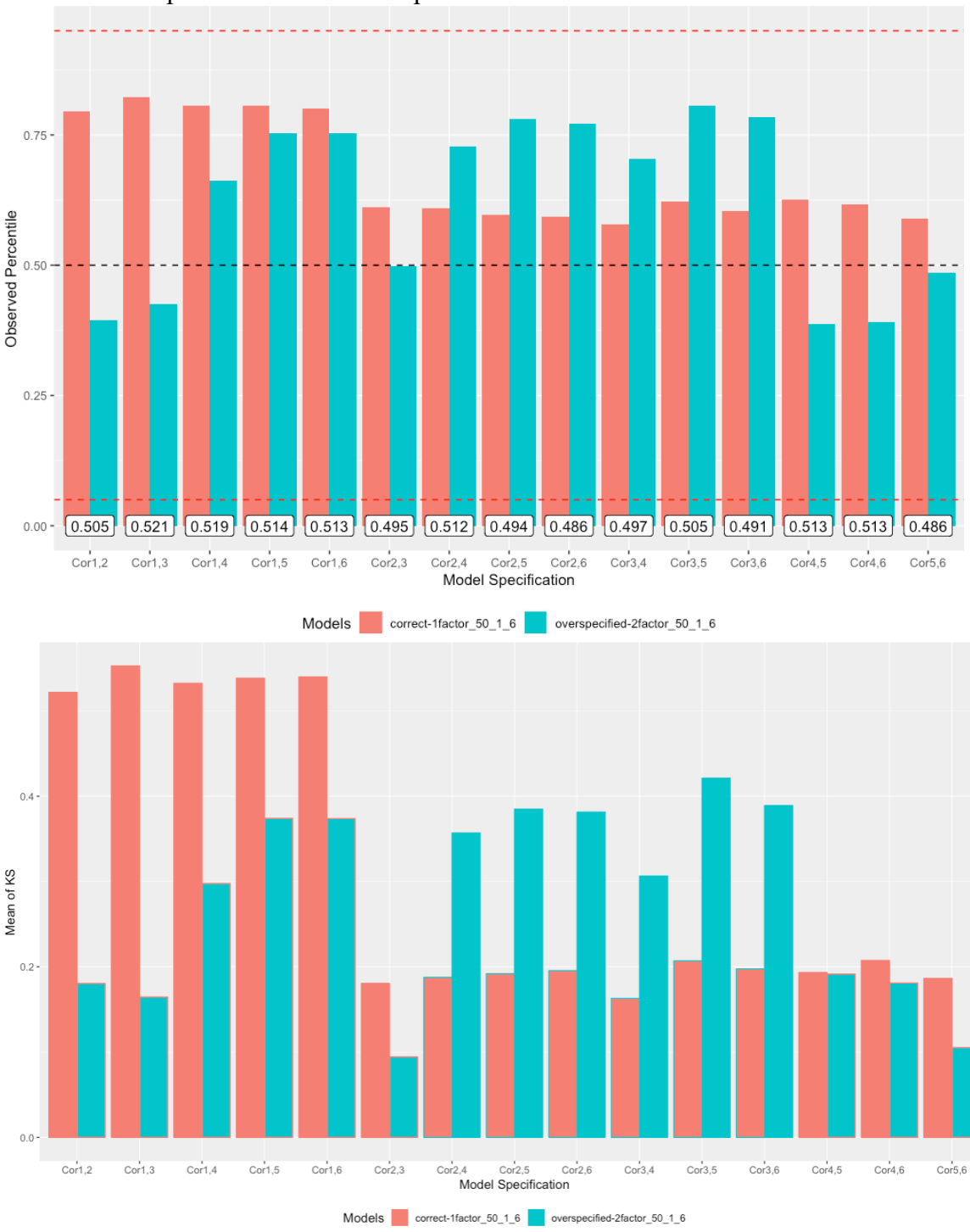
Correct one-factor CFA with N = 500, 12-item test





1.2 Correct One-factor (6 items) Model vs. overspecified Model

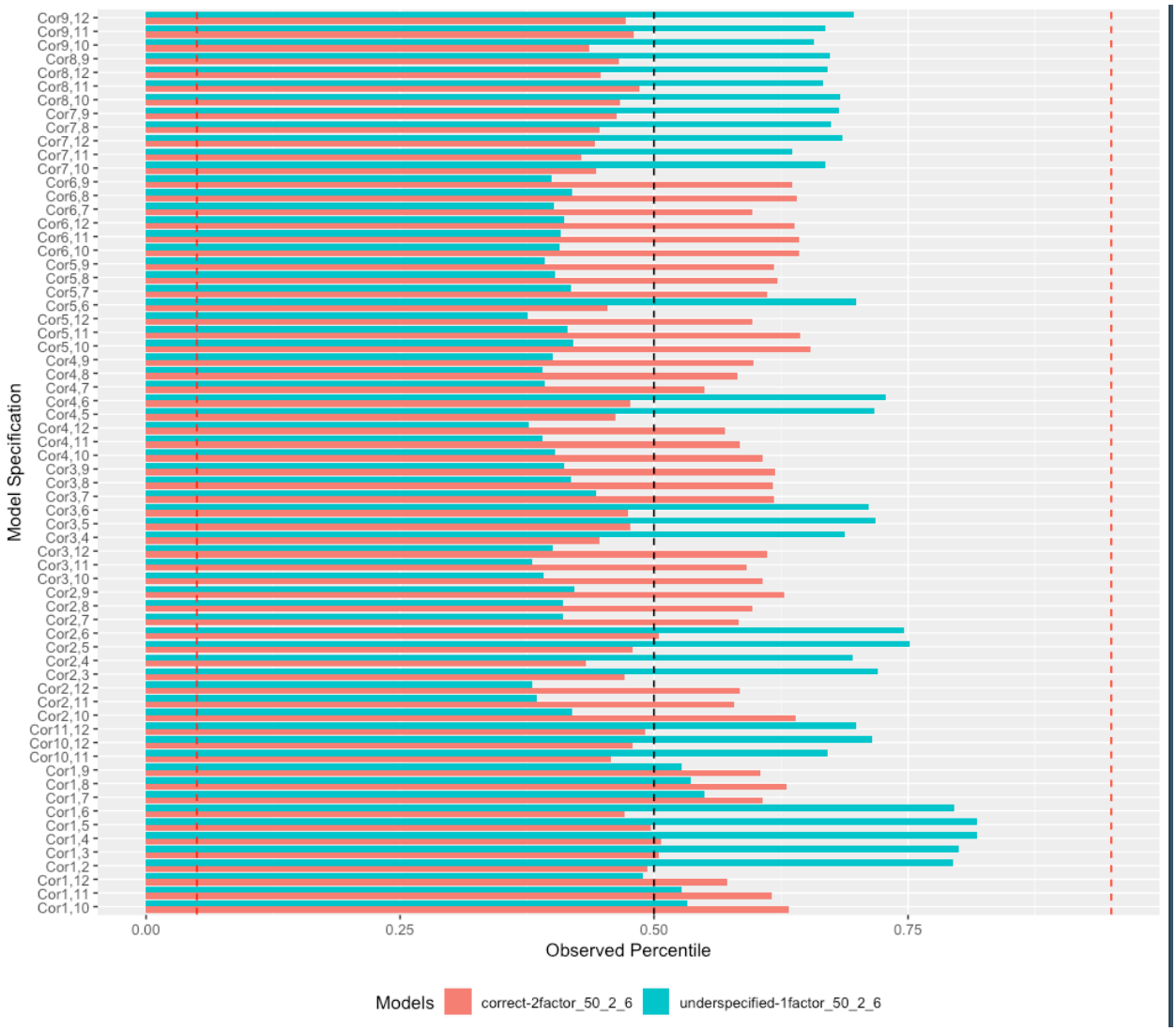
Both Posterior predictive KS test and p-value method were not able to detect the local misfit.



1.2 Correct 2-factor (12 items) test vs. underspecified one-factor model

For correct 2-factor (12 items) test vs. underspecified one-factor model, ppmc p-values was not able to detect local misfit.

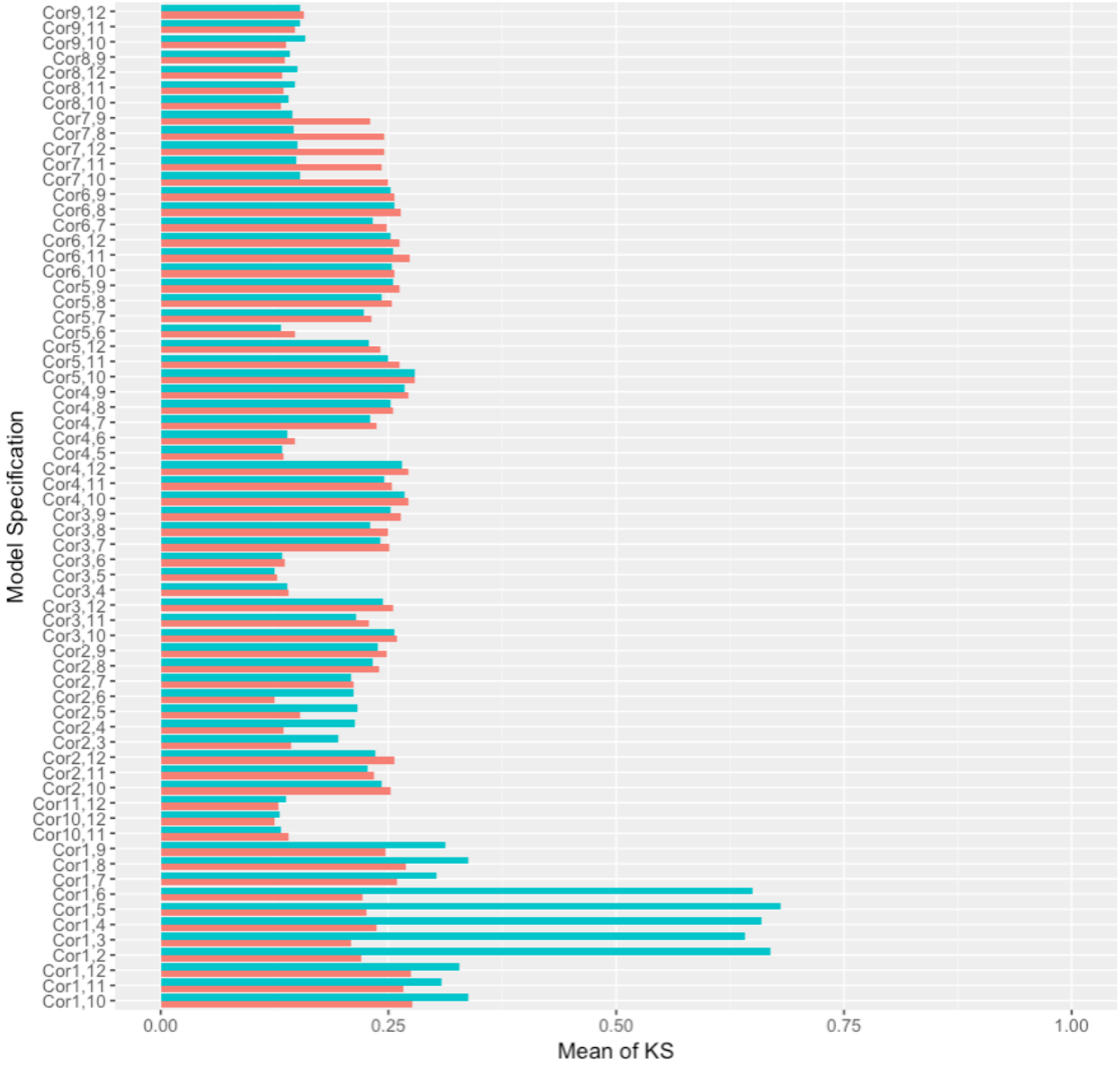
PPMC KS test in Bayesian Framework could detect some local misfit in underspecified with N = 50. For example, the KS means for correlation between item 1 with 5, 1 with 6, 1 with 5 in underspecified one-factor are much higher than those in correct 2-factor model.

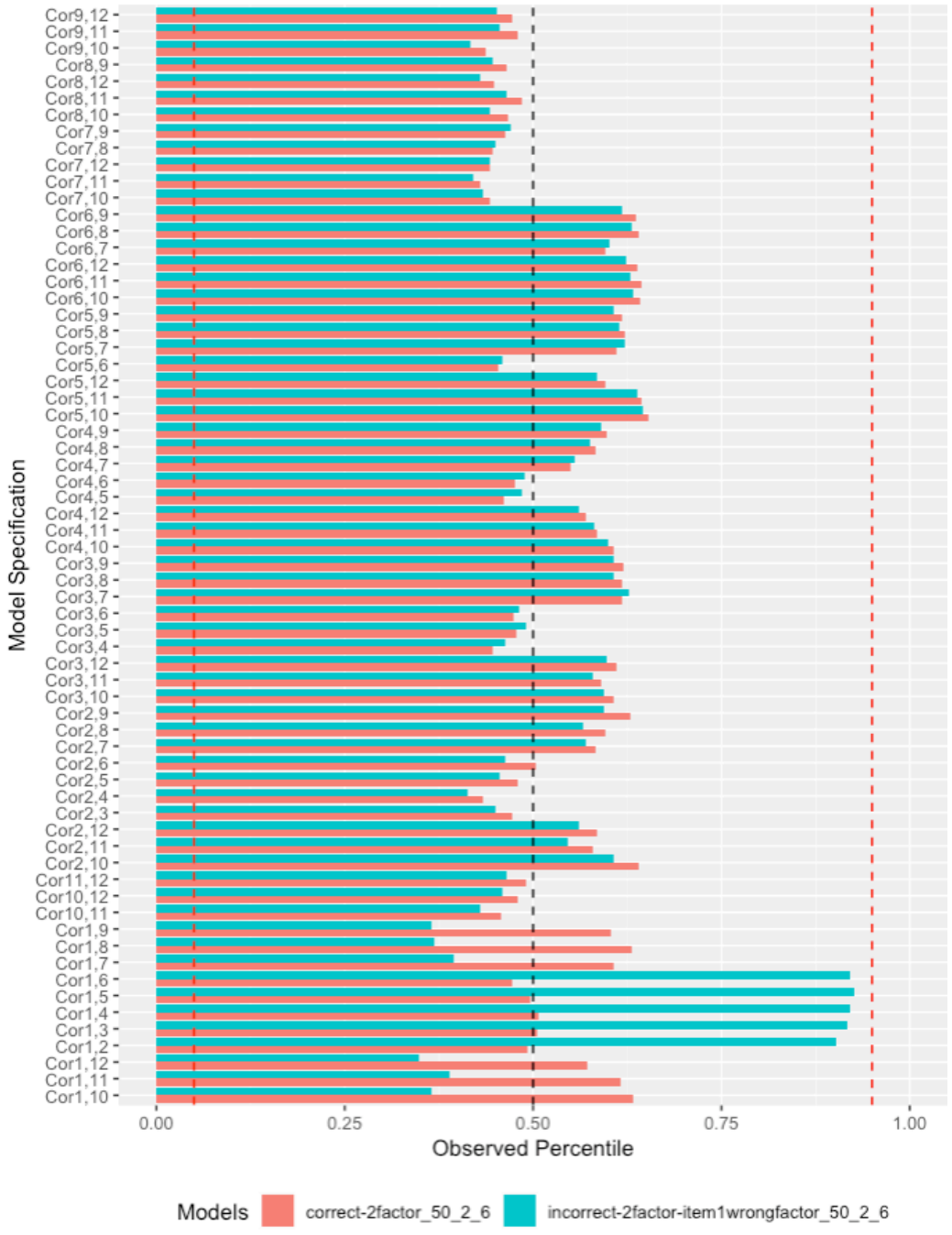


Maximum-Likelihood based p-values do not enough power to detect local misfit for underspecified model with N = 50.

1.3 Correct 2 factor (12-item test) vs. 2-factor cross-loading model

PP KS mean could detect the local misfit of the correlation item 1 with other items. PPP method don't have power to reject H0 hypothesis.

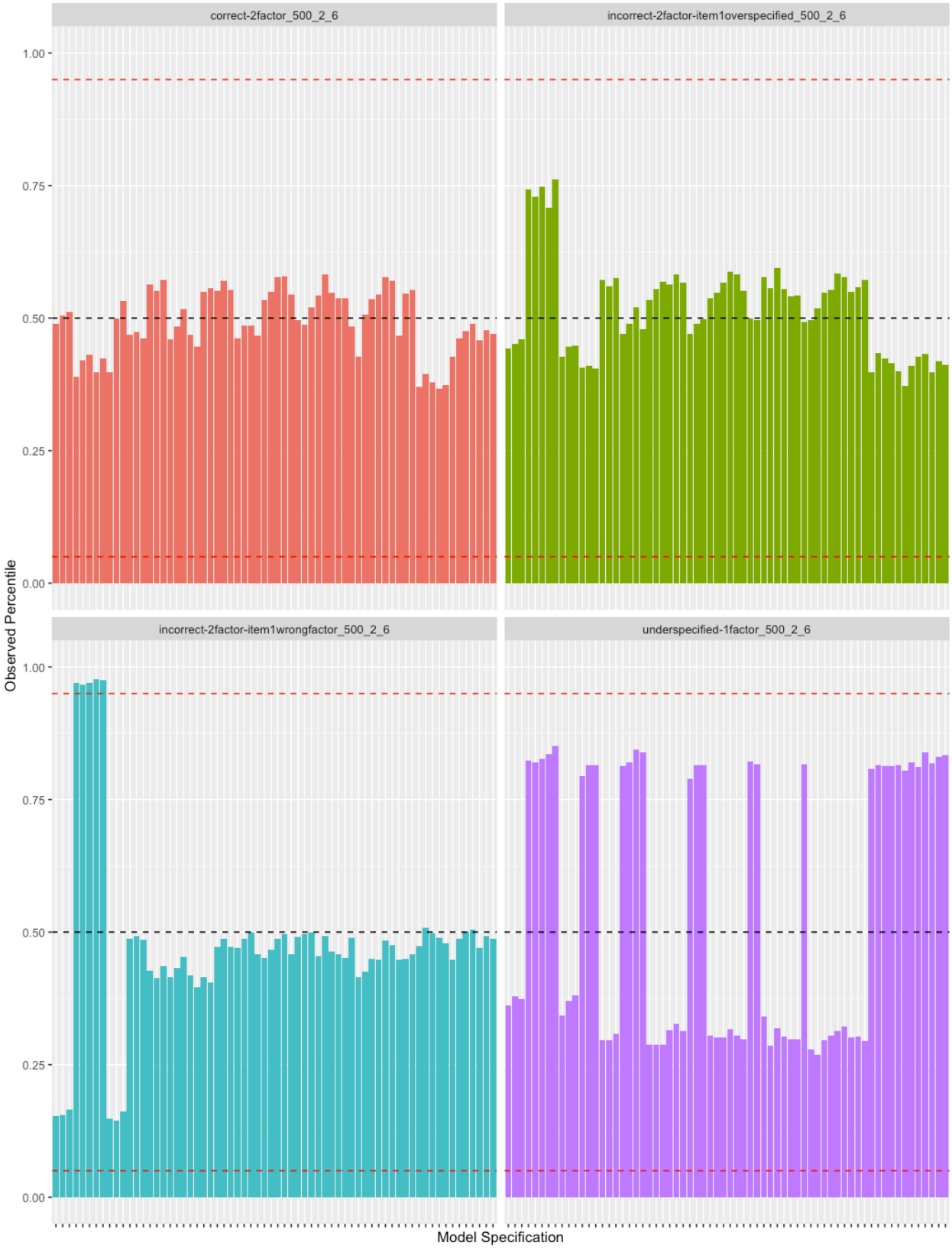



Model Specification
Cor9,12
Cor9,11
Cor9,10
Cor8,9
Cor8,12
Cor8,11
Cor8,10
Cor7,9
Cor7,8
Cor7,12
Cor7,11
Cor7,10
Cor6,9
Cor6,8
Cor6,7
Cor6,12
Cor6,11
Cor6,10
Cor5,9
Cor5,8
Cor5,7
Cor5,6
Cor5,12
Cor5,11
Cor5,10
Cor4,9
Cor4,8
Cor4,7
Cor4,6
Cor4,5
Cor4,12
Cor4,11
Cor4,10
Cor3,9
Cor3,8
Cor3,7
Cor3,6
Cor3,5
Cor3,4
Cor3,12
Cor3,11
Cor3,10
Cor2,9
Cor2,8
Cor2,7
Cor2,6
Cor2,5
Cor2,4
Cor2,3
Cor2,12
Cor2,11
Cor2,10
Cor11,12
Cor10,12
Cor10,11
Cor1,9
Cor1,8
Cor1,7
Cor1,6
Cor1,5
Cor1,4
Cor1,3
Cor1,2
Cor1,12
Cor1,11
Cor1,10
0.00
0.25
0.50
0.75
1.00
Observed Percentile
Models
correct-2factor_50_2_6
incorrect-2factor-item1wrongfactor_50_2_6

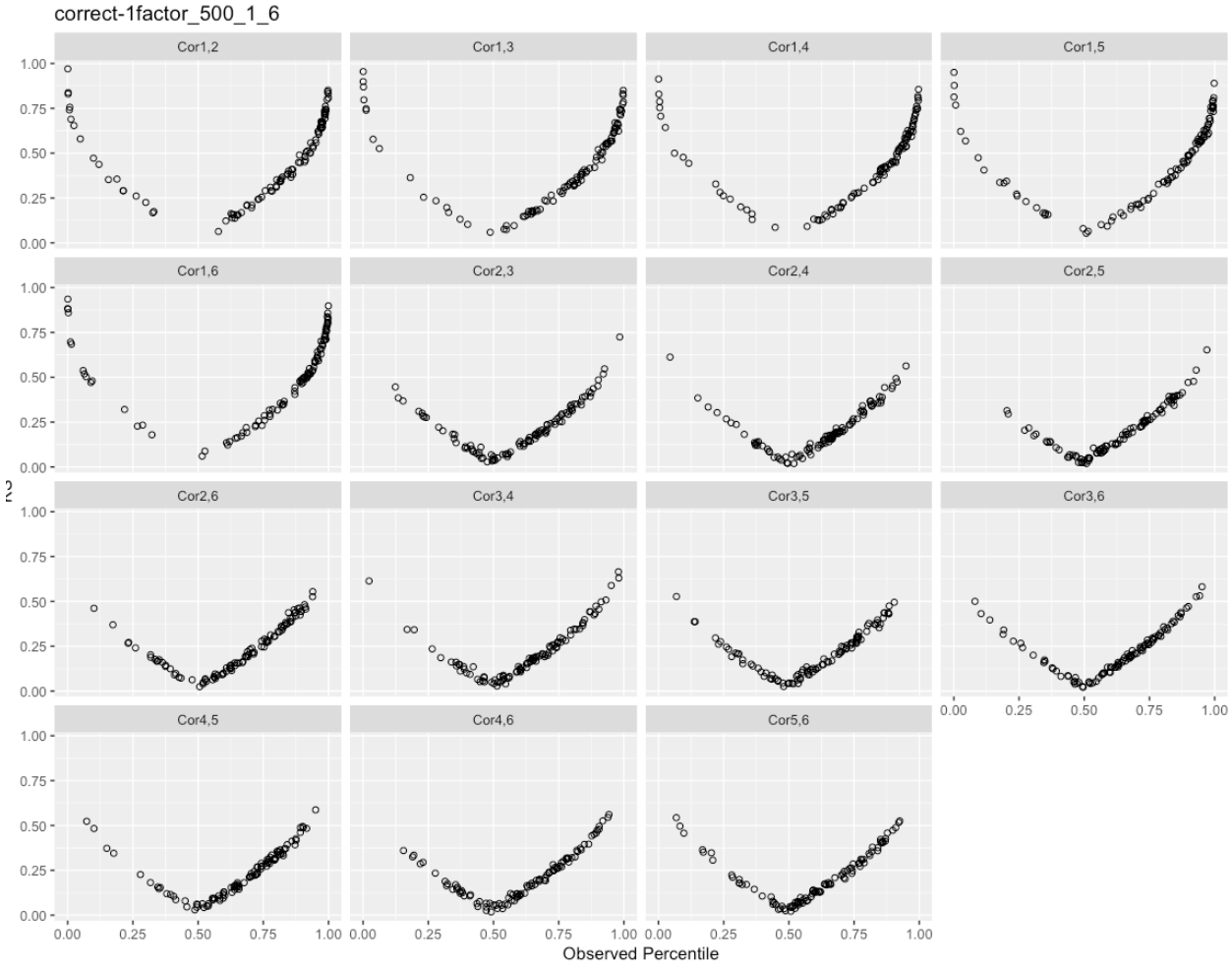
2. Large Sample Size with different types of misspecification

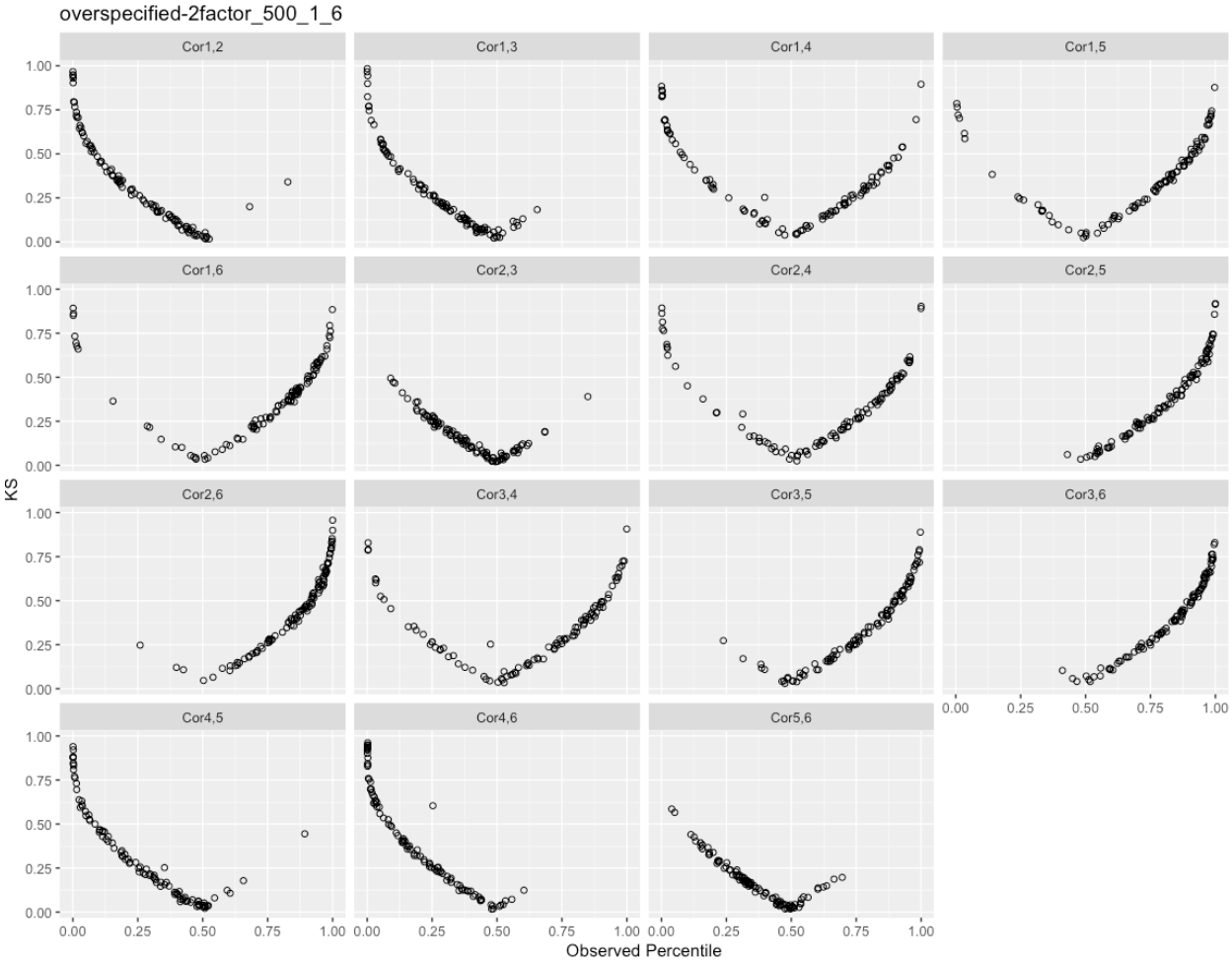
KS-plot for 2-factor (12-item test, N = 500)

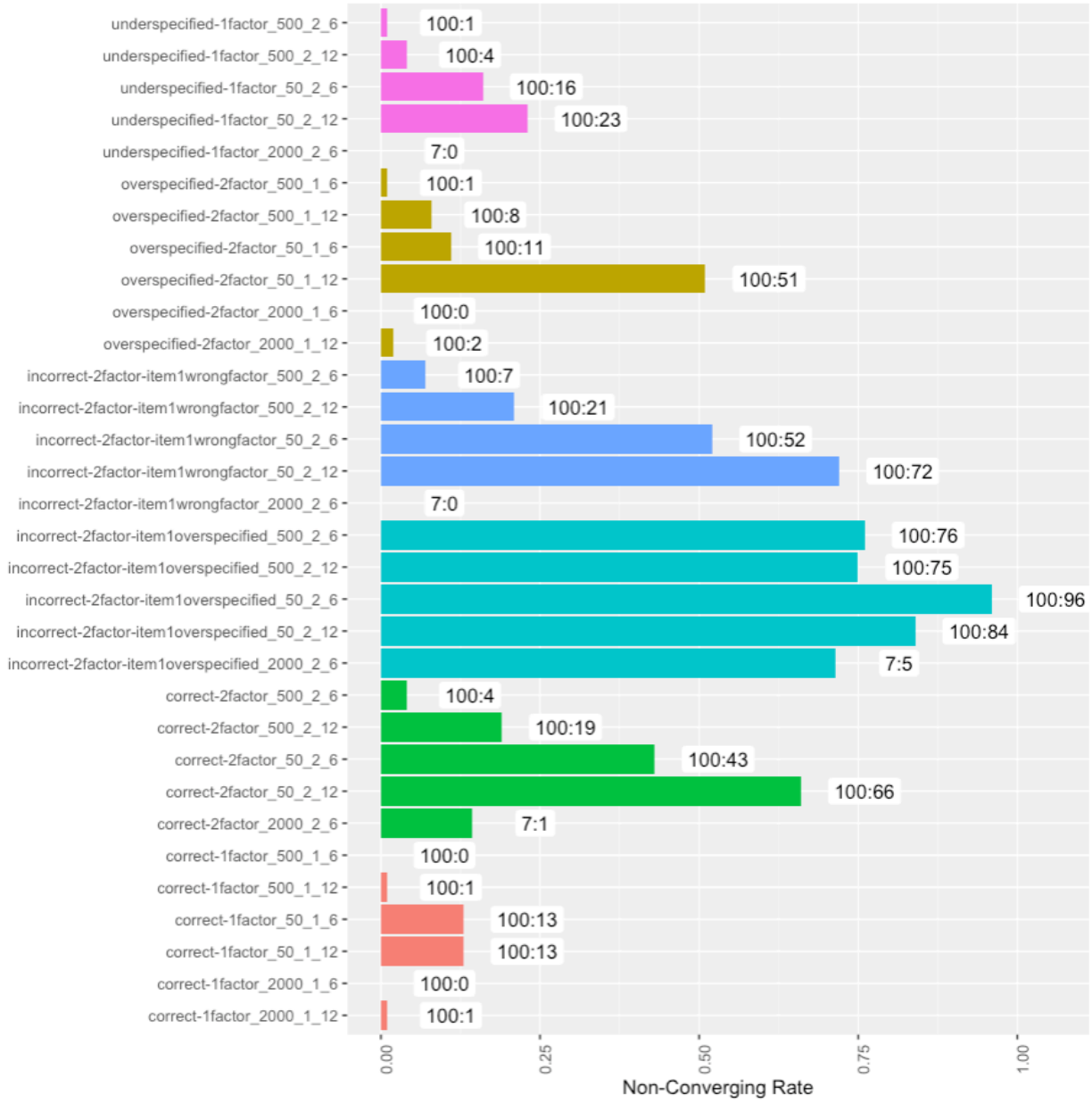
PPP value for 2-factor (12-item test, N=500)



The pattern between KS and PPP values follow a “V” shape line.




overspecified-2factor_500_1_6
Cor1,2
Cor1,3
Cor1,4
Cor1,5
Cor1,6
Cor2,3
Cor2,4
Cor2,5
Cor2,6
Cor3,4
Cor3,5
Cor3,6
Cor4,5
Cor4,6
Cor5,6
KS
Observed Percentile
0.00
0.25
0.50
0.75
1.00

underspecified-1factor_500_2_6
Cor1,10
Cor1,11
Cor1,12
Cor1,2
Cor1,3
Cor1,4
Cor1,5
Cor1,6
Cor1,7
Cor1,8
Cor1,9
Cor10,11
Cor10,12
Cor11,12
Cor2,10
Cor2,11
Cor2,12
Cor2,3
Cor2,4
Cor2,5
Cor2,6
Cor2,7
Cor2,8
Cor2,9
Cor3,10
Cor3,11
Cor3,12
Cor3,4
Cor3,5
Cor3,6
Cor3,7
Cor3,8
Cor3,9
Cor4,10
Cor4,11
Cor4,12
Cor4,5
Cor4,6
Cor4,7
Cor4,8
Cor4,9
Cor5,10
Cor5,11
Cor5,12
Cor5,6
Cor5,7
Cor5,8
Cor5,9
Cor6,10
Cor6,11
Cor6,12
Cor6,7
Cor6,8
Cor6,9
Cor7,10
Cor7,11
Cor7,12
Cor7,8
Cor7,9
Cor8,10
Cor8,11
Cor8,12
Cor8,9
Cor9,10
Cor9,11
Cor9,12
0.00
0.25
0.50
0.75
1.00
0.000.250.500.751.00
KS
Observed Percentile

3. Convergence Rate