

1 Appendix 1: Unconstrained Product Indicator Using All Pairs of Product Indicators

Appendix 1: Unconstrained Product Indicator Using All Pairs of Product Indicators

Marsh et al. (2004) demonstrated that all-pair UPI was not preferred due to its lack of substantial improvement, compared to matched-pair UPI, in latent interaction estimation but more more complex model specification. In our preliminary study, we investigated the performance of all-pair UPI on congeneric items, as Marsh et al. (2004) only examined parallel items without varying factor loadings and error variances.

We found that while standardized bias (SB) of interaction estimates produced by all-pair UPI were negligible for parallel items, they were more remarkable for congeneric factor items (with varied factor loadings but consistent error variances) and fully congeneric items (with varied factor loadings and error variances). It was shown that SB increased as sample size increased when $\gamma_{xm} = 0.3$ for all-pair UPI, which was contradictory to the findings of other latent interaction methods in our main study. Moreover, the relative standard error (SE) biases for all-pair UPI unacceptable exceeded the threshold range of [-10% to 10%], which were consistently downward. The results indicated that all-pair UPI, though similar to matched-pair UPI, tended to underestimate standard error of interaction effects particularly under the conditions of low and medium reliability, which were consistent with Marsh et al. (2004). Considering the pattern of SB and relative SE bias, a likely explanation for the increased SB could be that, although raw biases (RB) were very small across all the conditions, the corresponding SE estimates systematically decreased as sample size increased and thus amplified the calculation of SB.

Another issue we found with all-pair UPI was that its coverage rates, especially for low and medium item reliability, did not reach to the threshold of 91%, which was unsatisfactory. The low coverage rates indicated that all-pair UPI, potentially due to downward biased SE estimates, could not capture the true effects of latent interaction when observed items had certain amount of measurement errors, while the performance improved substantially as the reliability level increased.

28 These results suggested that all-pair UPI did not perform as well as matched-pair
29 UPI on congeneric items, particularly when the item reliability level was not high enough,
30 in addition to model complexity.

Table 1

2S-PA-IN

Evaluation Criteria of Zero Latent Interaction Effect ($\gamma_{xm} = 0$) for All-Pair UPI Across 2,000 Replications.

N	$Corr(\xi_x, \xi_m)$	Bias			Relative Standard Error Bias			Coverage Rate			RMSE			Type I Error Rate		
		$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$
100	0	-0.01 (-0.01)	-0.01 (-0.00)	0.07 (0.01)	-10.71 (9.35)	-9.01 (6.40)	-4.41 (3.20)	97.75	96.70	95.00	0.67	0.16	0.11	0.05	0.05	0.06
	0.3	-0.01 (-0.00)	0.02 (0.00)	0.05 (0.01)	-12.28 (10.65)	-12.07 (6.55)	-4.86 (1.85)	97.15	96.75	95.45	0.39	0.15	0.10	0.05	0.05	0.06
	0.6	-0.02 (-0.01)	0.01 (0.00)	0.04 (0.00)	-11.31 (9.10)	-2.66 (5.20)	-3.57 (1.55)	96.40	96.25	95.65	0.36	0.12	0.08	0.06	0.06	0.05
250	0	0.01 (0.00)	0.00 (0.00)	-0.00 (-0.00)	-4.98 (6.85)	-7.37 (4.25)	-5.83 (1.50)	95.80	95.00	94.15	0.12	0.08	0.07	0.05	0.06	0.06
	0.3	0.06 (0.01)	0.00 (0.00)	-0.01 (-0.00)	-5.29 (6.40)	-6.43 (3.25)	-2.56 (1.25)	96.20	94.50	94.00	0.13	0.07	0.06	0.05	0.06	0.06
	0.6	0.02 (0.00)	0.00 (0.00)	-0.01 (-0.00)	-4.71 (5.90)	-2.06 (3.30)	-2.36 (1.40)	95.60	94.45	94.00	0.08	0.06	0.05	0.06	0.07	0.07
500	0	0.00 (0.00)	-0.01 (-0.00)	-0.01 (-0.00)	-3.98 (4.65)	-2.89 (1.95)	-1.04 (1.00)	95.80	94.90	94.70	0.06	0.05	0.04	0.05	0.05	0.05
	0.3	0.00 (0.00)	-0.00 (-0.00)	-0.01 (-0.00)	-4.21 (3.80)	-1.32 (1.85)	-1.63 (1.05)	95.30	94.90	95.20	0.06	0.05	0.04	0.05	0.05	0.05
	0.6	0.00 (0.00)	-0.00 (-0.00)	-0.01 (-0.00)	-4.14 (3.05)	-2.03 (1.50)	2.65 (1.20)	95.30	95.00	94.55	0.05	0.04	0.04	0.05	0.06	0.06

Note. N = sample size; $Corr(\xi_x, \xi_m)$ = correlation between ξ_x and ξ_m ; ρ = reliability level; Bias = standardized bias of latent interaction effect (with raw bias shown in parentheses); Relative standard error bias = relative standard error bias of estimated standard errors (with outlier proportions of standard errors shown in parentheses); Coverage rate = coverage rate of 95% confidence interval (CI) of latent interaction effect; RMSE = root mean square error of latent interaction effect; Type I error rate = empirical error rate of falsely rejecting the null hypothesis of zero latent interaction effect. All numerical values are rounded to two decimal places for consistency. Note that values close to zero are displayed as 0.00, with negative signs maintained to indicate the direction of bias. Besides, values exceeding the recommended threshold (0.40) are bolded.

Table 2

Evaluation Criteria of Non-Zero Latent Interaction Effect ($\gamma_{xm} = 0$) for All-Pair UPI Across 2,000 Replications.

N	$Corr(\xi_x, \xi_m)$	Bias			Relative Standard Error Bias			Coverage Rate			RMSE			Statistical Power		
		$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$	$\rho = .70$	$\rho = .80$	$\rho = .90$
100	0	-0.03 (-0.01)	-0.10 (-0.02)	-0.09 (-0.01)	-25.16 (9.45)	-18.52 (6.00)	-3.74 (3.00)	75.35	85.15	91.95	0.43	0.19	0.12	0.38	0.62	0.82
	0.3	0.04 (0.03)	-0.01 (-0.00)	-0.10 (-0.01)	-25.96 (10.20)	-14.40 (6.30)	-5.81 (2.35)	77.05	86.95	92.80	0.86	0.34	0.11	0.45	0.71	0.87
	0.6	0.01 (0.00)	-0.03 (-0.01)	-0.11 (-0.01)	-22.18 (10.35)	-10.76 (5.30)	-1.69 (1.45)	80.85	87.60	92.50	0.42	0.23	0.09	0.58	0.83	0.93
250	0	-0.12 (-0.02)	-0.25 (-0.02)	-0.21 (-0.02)	-21.97 (7.40)	-11.64 (4.25)	-9.20 (1.45)	79.75	86.20	91.05	0.19	0.10	0.07	0.83	0.96	0.99
	0.3	-0.09 (-0.02)	-0.25 (-0.02)	-0.21 (-0.01)	-20.09 (6.75)	-9.27 (3.40)	-7.33 (1.15)	79.55	87.20	92.00	0.23	0.09	0.07	0.88	0.98	1.00
	0.6	-0.16 (-0.02)	-0.23 (-0.02)	-0.19 (-0.01)	-12.95 (5.45)	-4.74 (2.90)	-5.38 (1.20)	83.85	89.35	92.40	0.12	0.08	0.06	0.96	1.00	1.00
500	0	-0.40 (-0.04)	-0.43 (-0.03)	-0.32 (-0.02)	-19.01 (4.25)	-9.74 (2.25)	-3.09 (0.90)	79.00	85.70	91.25	0.10	0.07	0.05	0.99	1.00	1.00
	0.3	-0.39 (-0.03)	-0.40 (-0.02)	-0.30 (-0.01)	-19.43 (4.30)	-10.84 (2.50)	-3.60 (1.15)	79.60	86.20	92.15	0.09	0.06	0.05	1.00	1.00	1.00
	0.6	-0.36 (-0.03)	-0.35 (-0.02)	-0.26 (-0.01)	-19.25 (3.55)	-11.94 (1.95)	-4.90 (0.90)	82.80	88.10	92.00	0.08	0.06	0.04	1.00	1.00	1.00

Note. N = sample size; $Corr(\xi_x, \xi_m)$ = correlation between ξ_x and ξ_m ; ρ = reliability level; Bias = standardized bias of latent interaction effect (with raw bias shown in paratheses); Relative standard error bias = relative standard error bias of estimated standard errors (with outlier proportions of standard errors shown in parentheses); Coverage rate = coverage rate of 95% confidence interval (CI) of latent interaction effect; RMSE = root mean square error of latent interaction effect; Statistical power = rate of correctly rejecting the null hypothesis of zero latent interaction effect. All numerical values are rounded to two decimal places for consistency. Note that values close to zero are displayed as 0.00, with negative signs maintained to indicate the direction of bias. Besides, values exceeding the recommended threshold (0.40) are bolded.