

Robustness of Adaptive Measurement of Change to Item Parameter Estimation Error

Supplemental Tables and Figures

Table S1

Classical effect sizes (η^2) from a two-way analysis of variance (ANOVA) on false positive rates, true positive rates, average test length, and change recovery in Simulation 2

Factor	η^2	Sum of Squares	DF
<i>False Positive Rates</i>			
Starting θ (Trait)	0.123	0.007	10
Calibration Size (Size)	0.068	0.004	4
AMC Hypothesis Test (Test)	0.112	0.006	2
Trait \times Size	0.268	0.015	40
Trait \times Test	0.206	0.012	20
Size \times Test	0.027	0.002	8
Residuals	-	0.011	80
<i>True Positive Rates</i>			
Starting θ (Trait)	0.021	0.207	10
Calibration Size (Size)	0.005	0.048	4
AMC Hypothesis Test (Test)	0.131	1.276	2
True Change (Change)	0.641	6.228	8
Trait \times Size	0.066	0.644	40
Trait \times Test	0.039	0.374	20
Trait \times Change	0.018	0.174	80
Size \times Test	0.000	0.004	8
Size \times Change	0.005	0.049	32
Test \times Change	0.010	0.095	16
Residuals	-	0.618	1264
<i>Average Test Length</i>			
Starting θ (Trait)	0.022	229.840	10
Calibration Size (Size)	0.002	20.421	4
AMC Hypothesis Test (Test)	0.096	1018.513	2
True Change (Change)	0.494	5240.393	9
Trait \times Size	0.046	485.211	40
Trait \times Test	0.103	1089.499	20
Trait \times Change	0.020	214.723	90
Size \times Test	0.004	40.999	8
Size \times Change	0.008	85.134	36
Test \times Change	0.068	725.197	18
Residuals	-	1456.837	1412
<i>Change Recovery Index</i>			
Starting θ (Trait)	0.135	3.178	10
Calibration Size (Size)	0.086	2.020	4
AMC Hypothesis Test (Test)	0.021	0.489	2
True Change (Change)	0.023	0.551	9
Trait \times Size	0.082	1.932	40
Trait \times Test	0.253	5.940	20

Trait × Change	0.053	1.243	90
Size × Test	0.022	0.514	8
Size × Change	0.090	2.111	36
Test × Change	0.023	0.533	18
Residuals	-	4.979	1412

Note. Classical effect sizes greater than or equal to 0.02 have been bolded.

Table S2

Classical effect sizes (η^2) from a two-way analysis of variance (ANOVA) on false positive rates, true positive rates, average test length, and change recovery in Simulation 3

Factor	η^2	Sum of Squares	DF
<i>False Positive Rates</i>			
Starting θ (Trait)	0.043	0.032	10
Item Bank (Bank)	0.789	0.590	2
AMC Hypothesis Test (Test)	0.005	0.004	2
Trait \times Bank	0.150	0.112	20
Trait \times Test	0.008	0.006	20
Bank \times Test	0.000	0.000	4
Residuals	-	0.002	40
<i>True Positive Rates</i>			
Starting θ (Trait)	0.087	0.911	10
Item Bank (Bank)	0.092	0.960	2
AMC Hypothesis Test (Test)	0.012	0.125	2
True Change (Change)	0.563	5.896	8
Trait \times Bank	0.121	1.261	20
Trait \times Test	0.010	0.103	20
Trait \times Change	0.044	0.456	80
Bank \times Test	0.000	0.003	4
Bank \times Change	0.021	0.223	16
Test \times Change	0.002	0.019	16
Residuals	-	0.506	712
<i>Average Test Length</i>			
Starting θ (Trait)	0.043	1379.015	10
Item Bank (Bank)	0.168	5388.399	2
AMC Hypothesis Test (Test)	0.002	75.811	2
True Change (Change)	0.630	20149.188	9
Trait \times Bank	0.069	2213.930	20
Trait \times Test	0.013	416.021	20
Trait \times Change	0.020	652.265	90
Bank \times Test	0.000	0.735	4
Bank \times Change	0.007	212.849	18
Test \times Change	0.009	276.738	18
Residuals	-	1232.674	796
<i>Change Recovery Index</i>			
Starting θ (Trait)	0.062	1.041	10
Item Bank (Bank)	0.161	2.699	2
AMC Hypothesis Test (Test)	0.002	0.034	2
True Change (Change)	0.365	6.127	9
Trait \times Bank	0.122	2.049	20
Trait \times Test	0.021	0.345	20

Trait × Change	0.160	2.690	90
Bank × Test	0.000	0.007	4
Bank × Change	0.025	0.421	18
Test × Change	0.004	0.063	18
Residuals	-	1.292	796

Note. Classical effect sizes greater than or equal to 0.02 have been bolded.

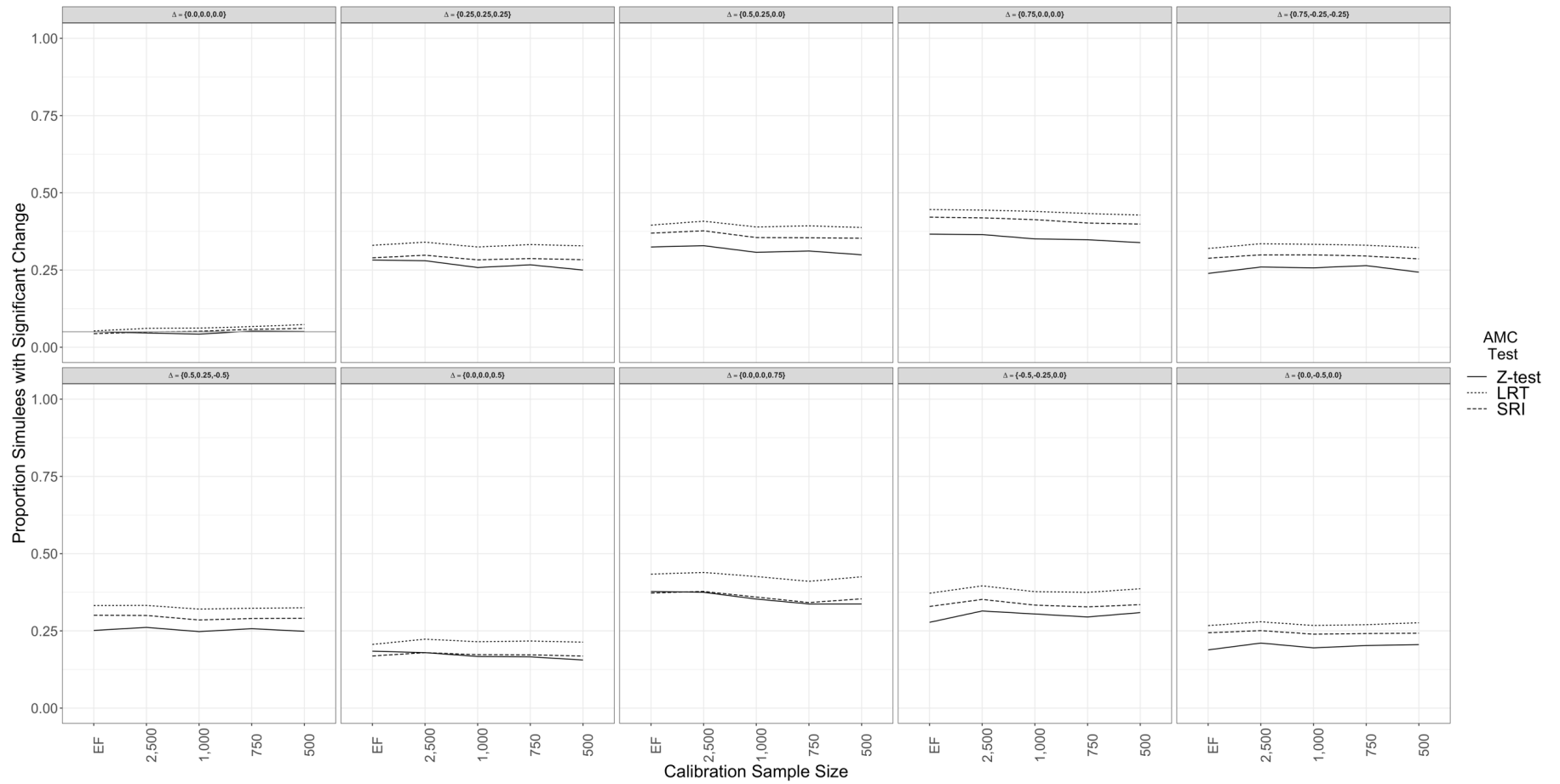


Figure S1. FPRs (Panel 1) and TPRs (Panels 2-10) for three AMC hypothesis tests at the fourth testing occasion across 10 θ change trajectories in Simulation 2 conditioning on calibration sample size.

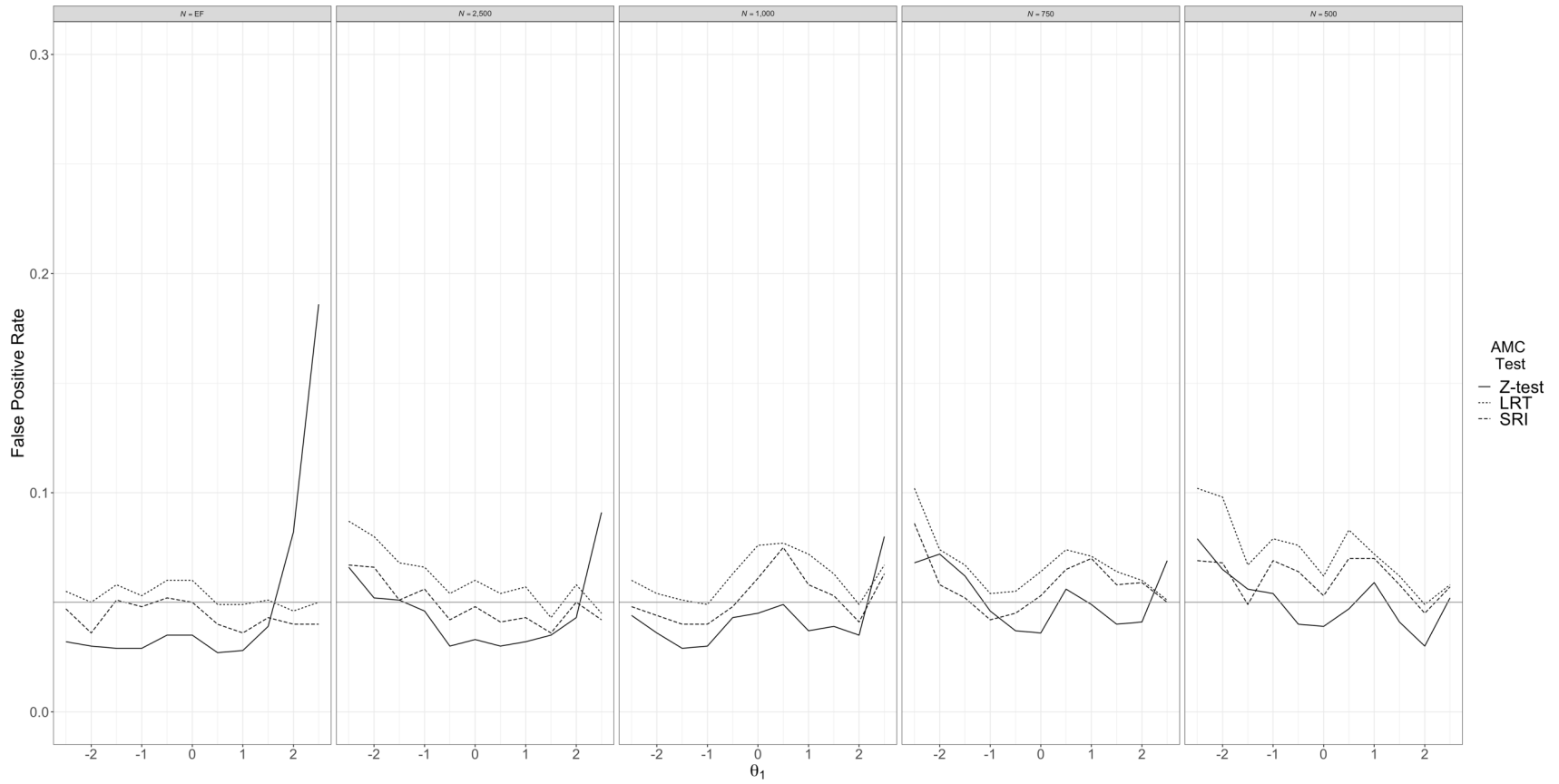


Figure S2. FPRs for three AMC hypothesis tests across decreasing calibration sample sizes when conditioning on θ_1 in Simulation 2.

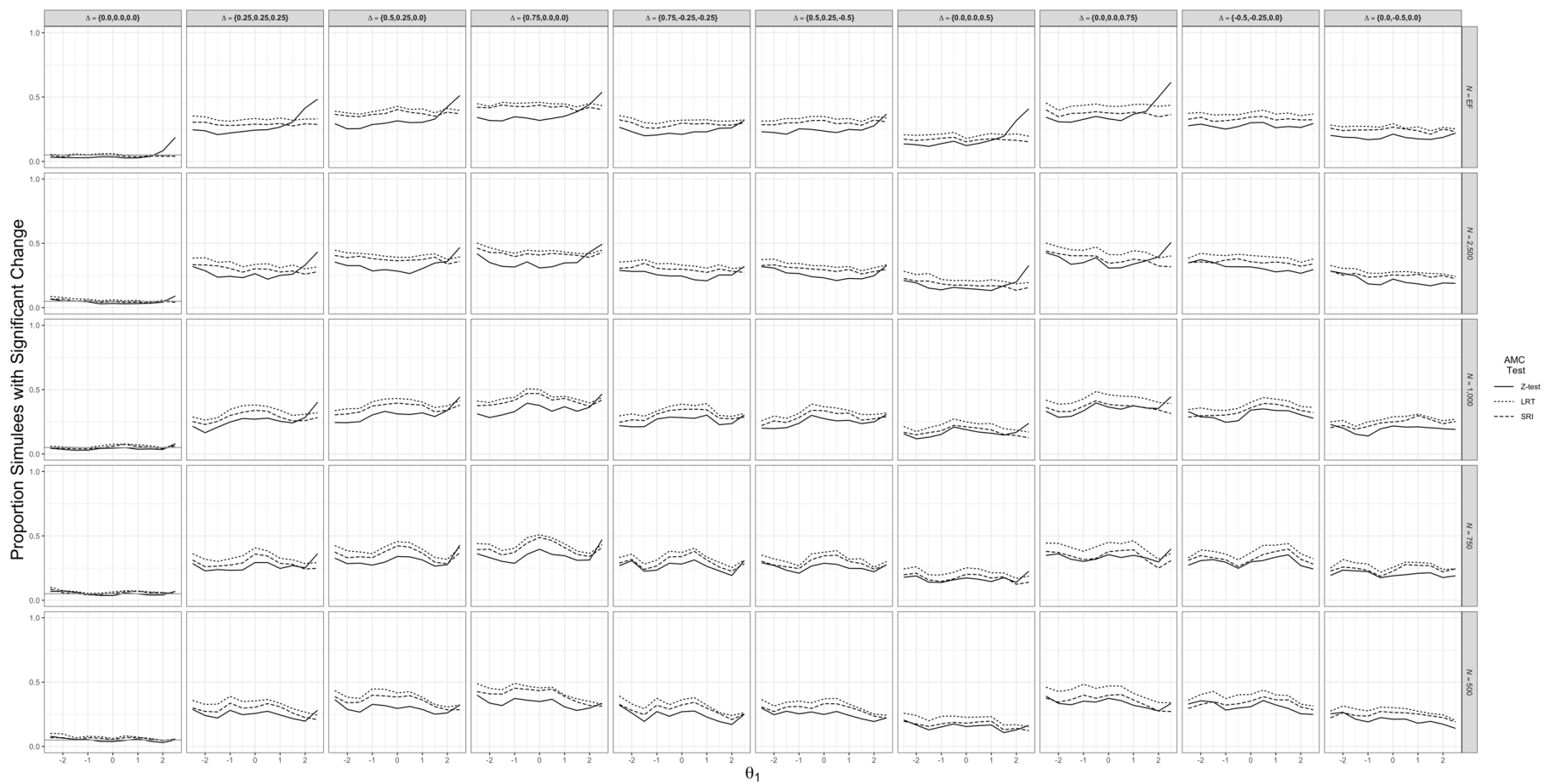


Figure S3. FPRs (Panel 1) and TPRs (Panels 2-10) for three AMC hypothesis tests across 10 θ change trajectories when conditioning on θ_1 in Simulation 2.

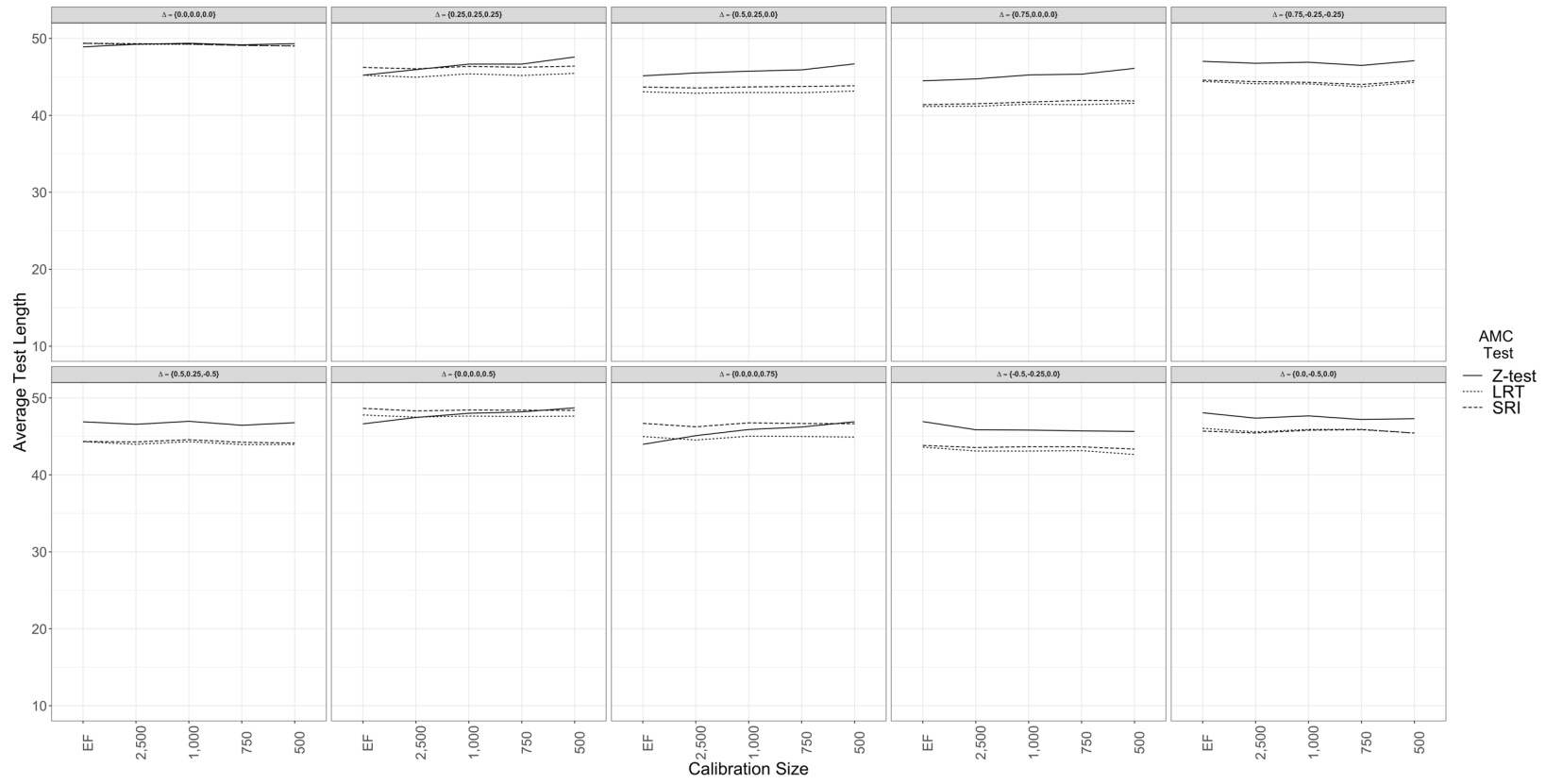


Figure S4. ATL for three AMC hypothesis tests at the fourth testing occasion across 10 θ change trajectories in Simulation 2 conditioning on calibration sample size.

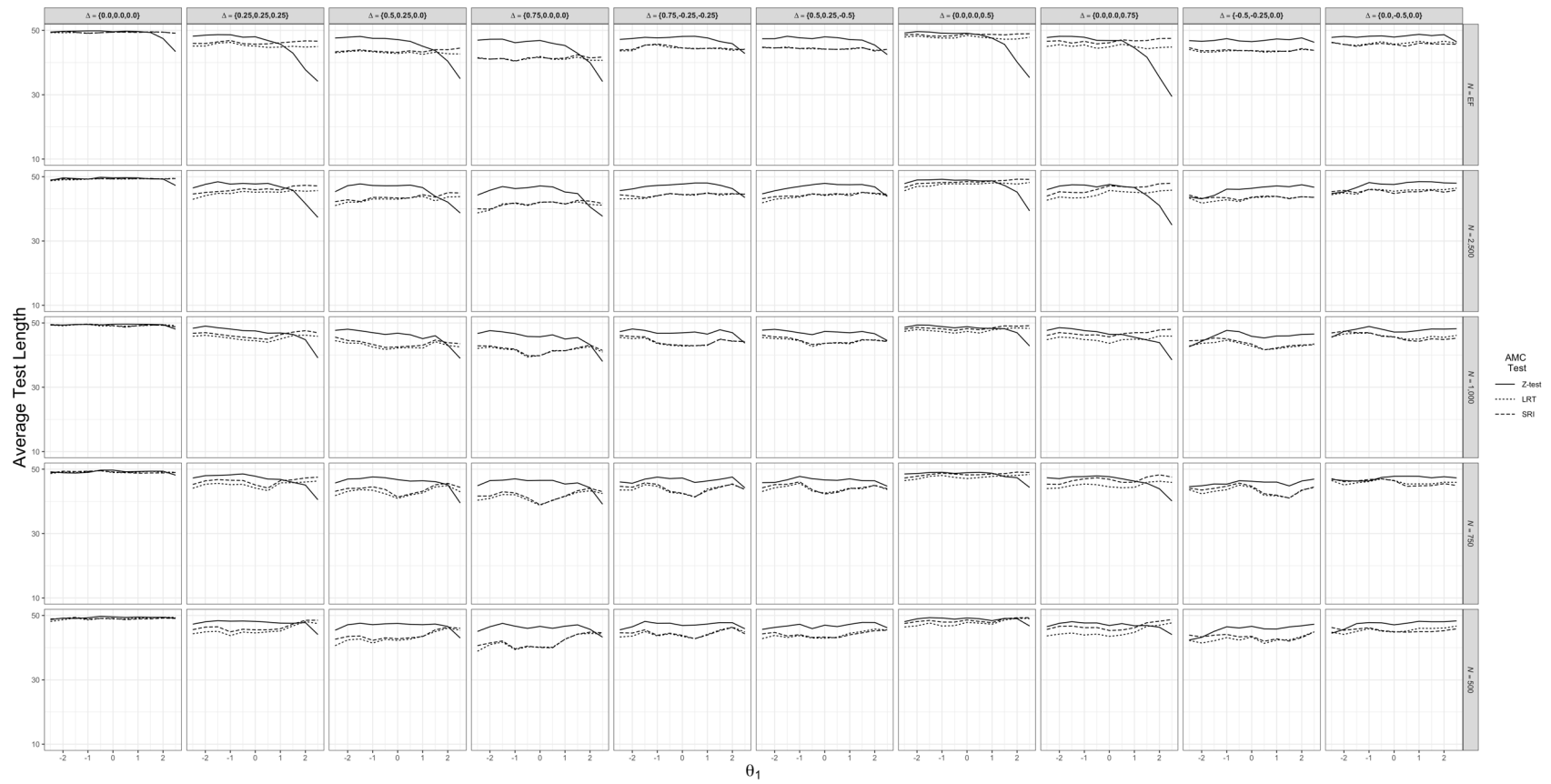


Figure S5. ATL for three AMC hypothesis tests across 10 θ trajectories when conditioning on θ_1 in Simulation 2.

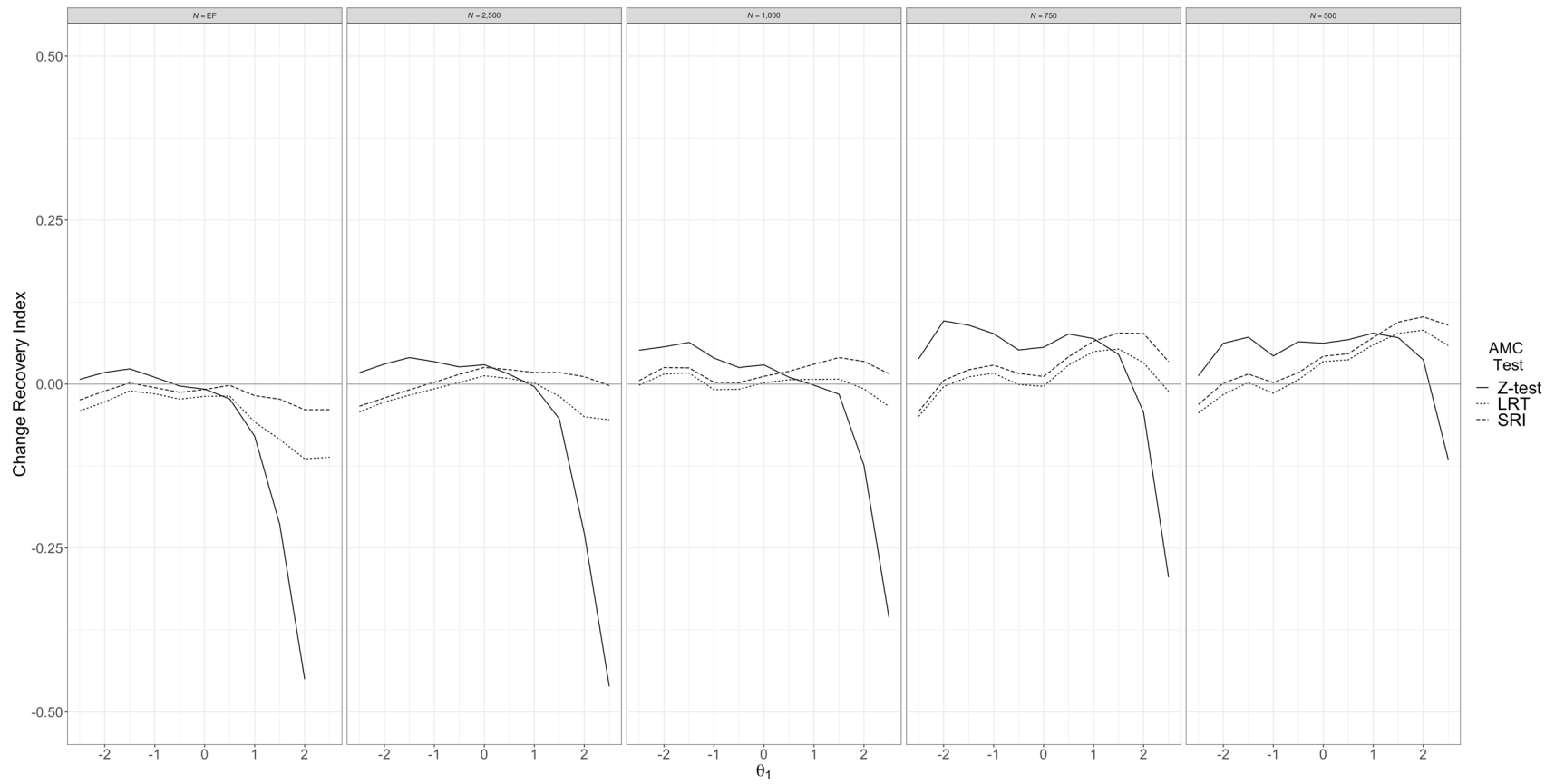


Figure S6. Change recovery index for three AMC hypothesis tests across decreasing calibration sample sizes when conditioning on θ_1 in Simulation 2.

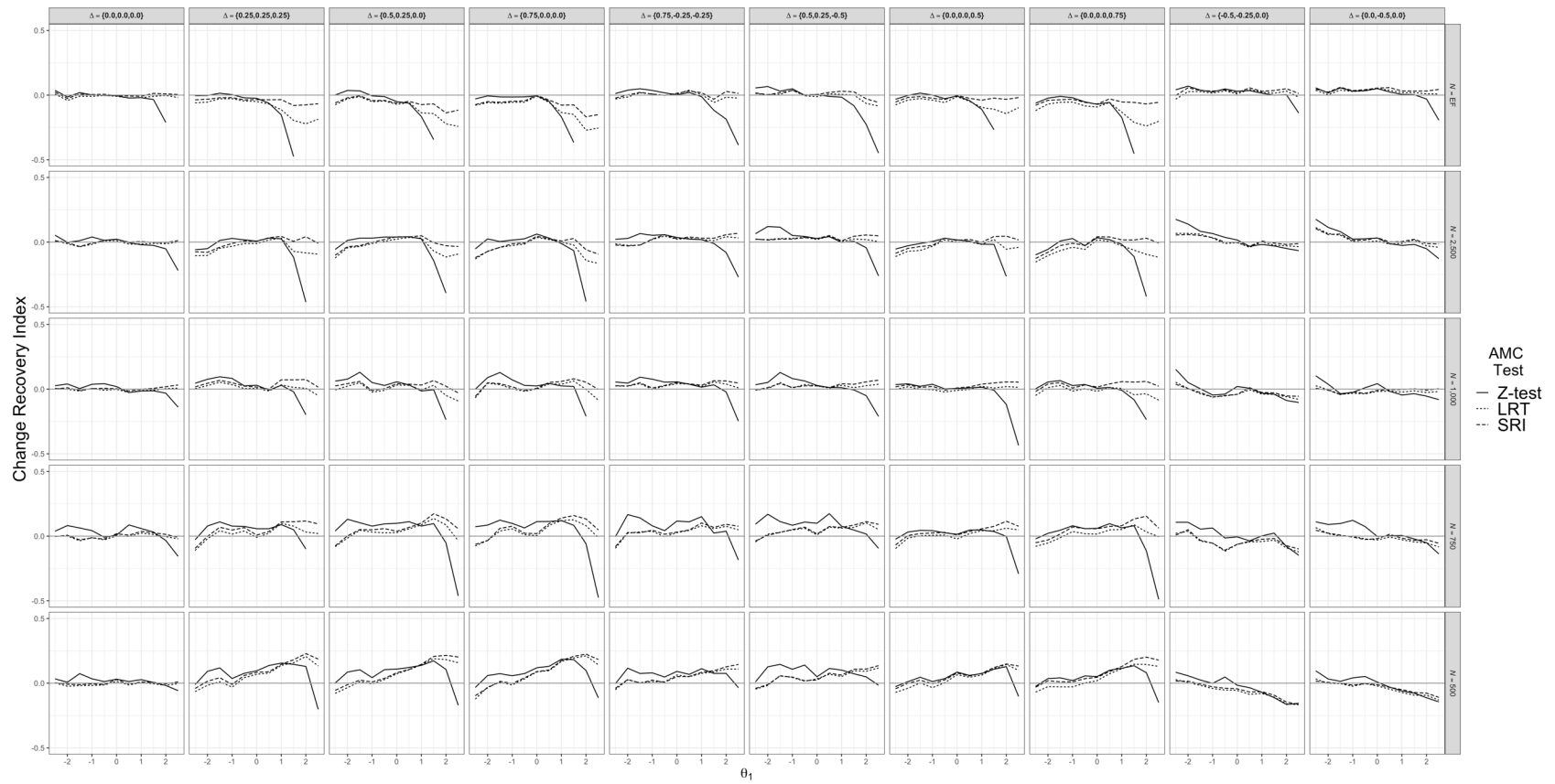


Figure S7. Change recovery index for three AMC hypothesis tests across 10 θ trajectories when conditioning on θ_1 in Simulation 2.

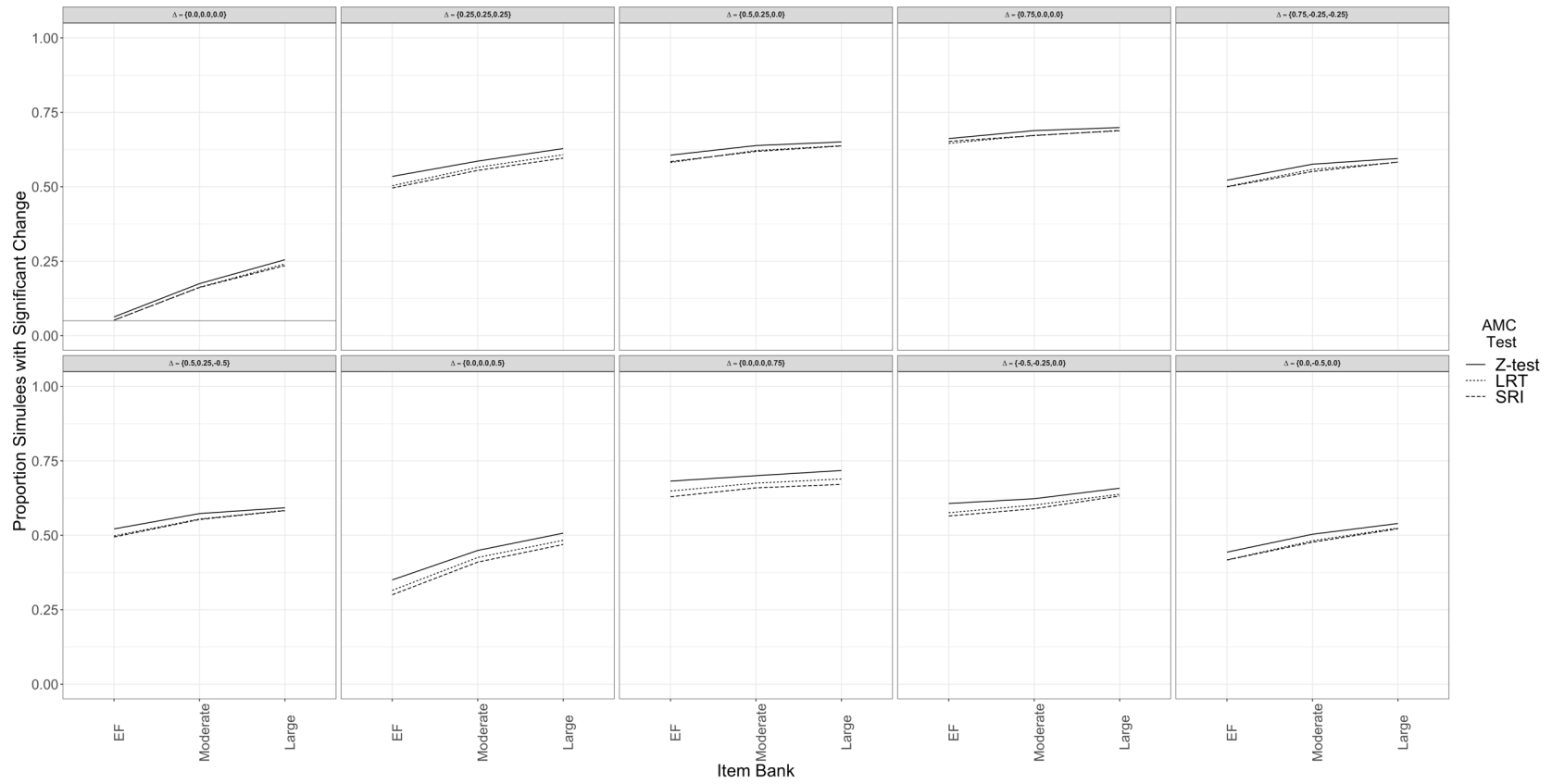


Figure S8. FPRs (Panel 1) and TPRs (Panels 2-10) for three AMC hypothesis tests at the fourth testing occasion across 10 θ change trajectories in Simulation 3 for item banks with no (EF), moderate, or large calibration error.

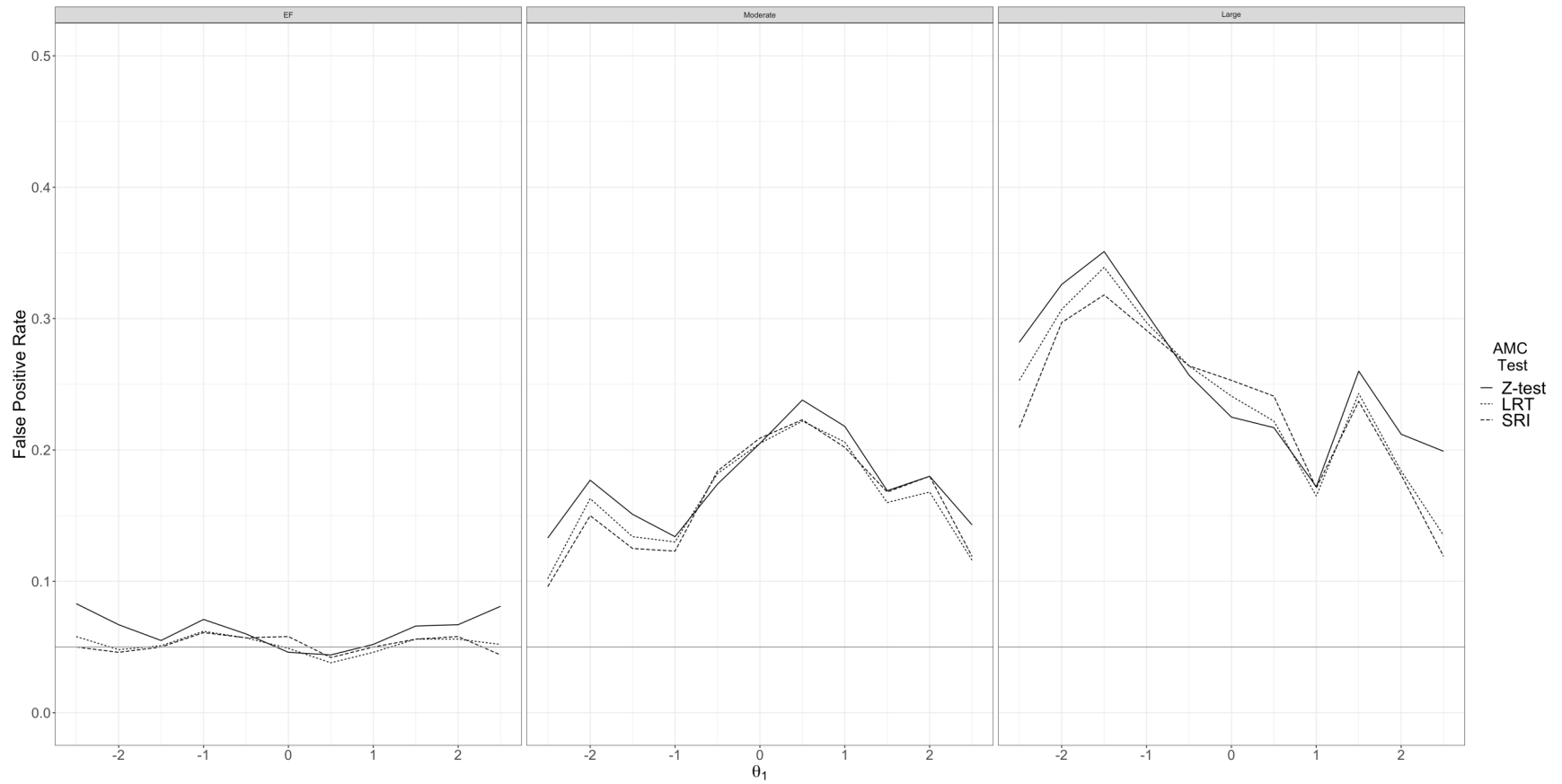


Figure S9. FPRs for three AMC hypothesis tests across item banks with increasing amounts of item parameter estimation error when conditioning on θ_1 in Simulation 3.

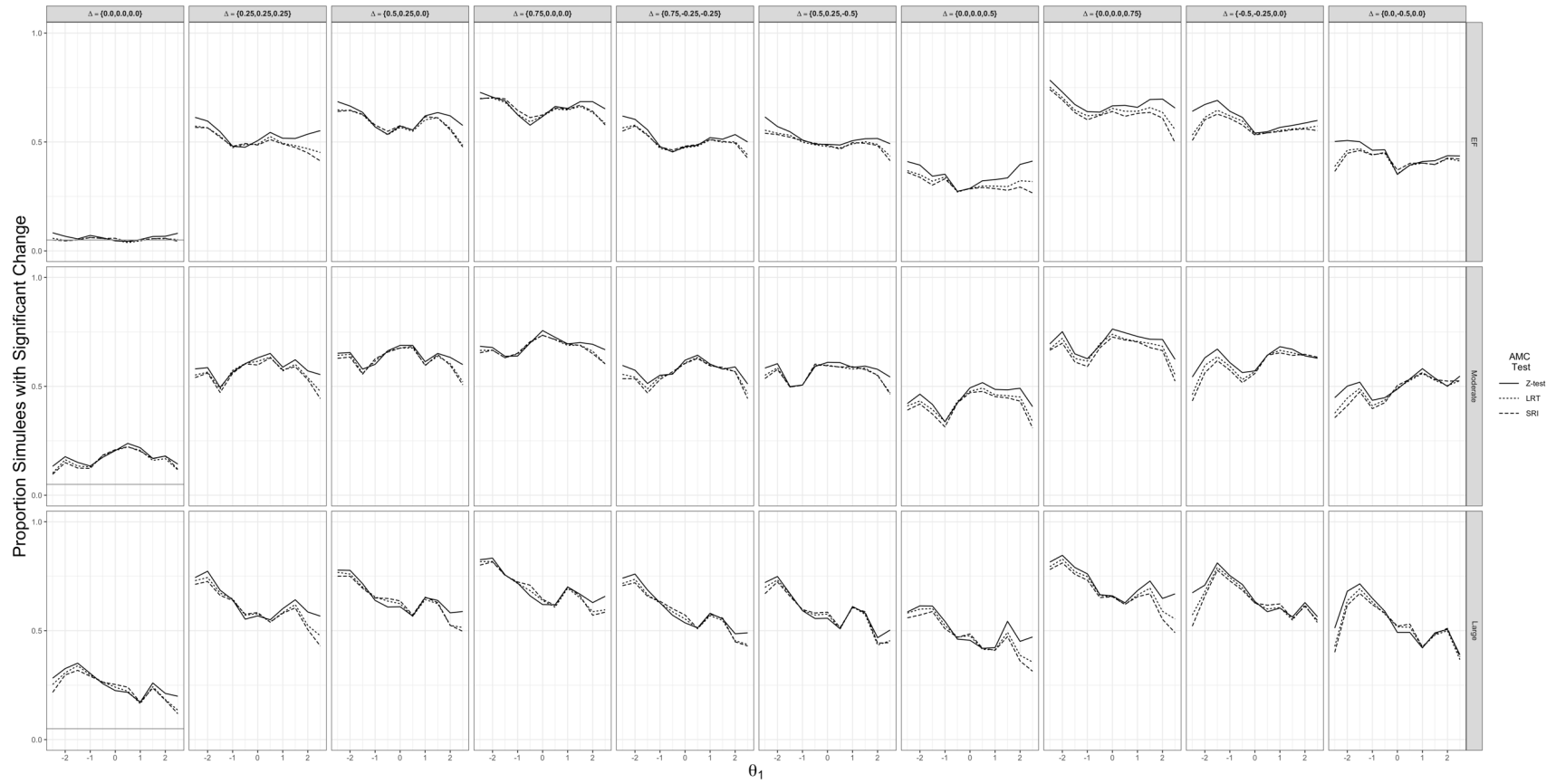


Figure S10. FPRs (Panel 1) and TPRs (Panels 2-10) for three AMC hypothesis tests across 10 θ change trajectories when conditioning on θ_1 in Simulation 3.

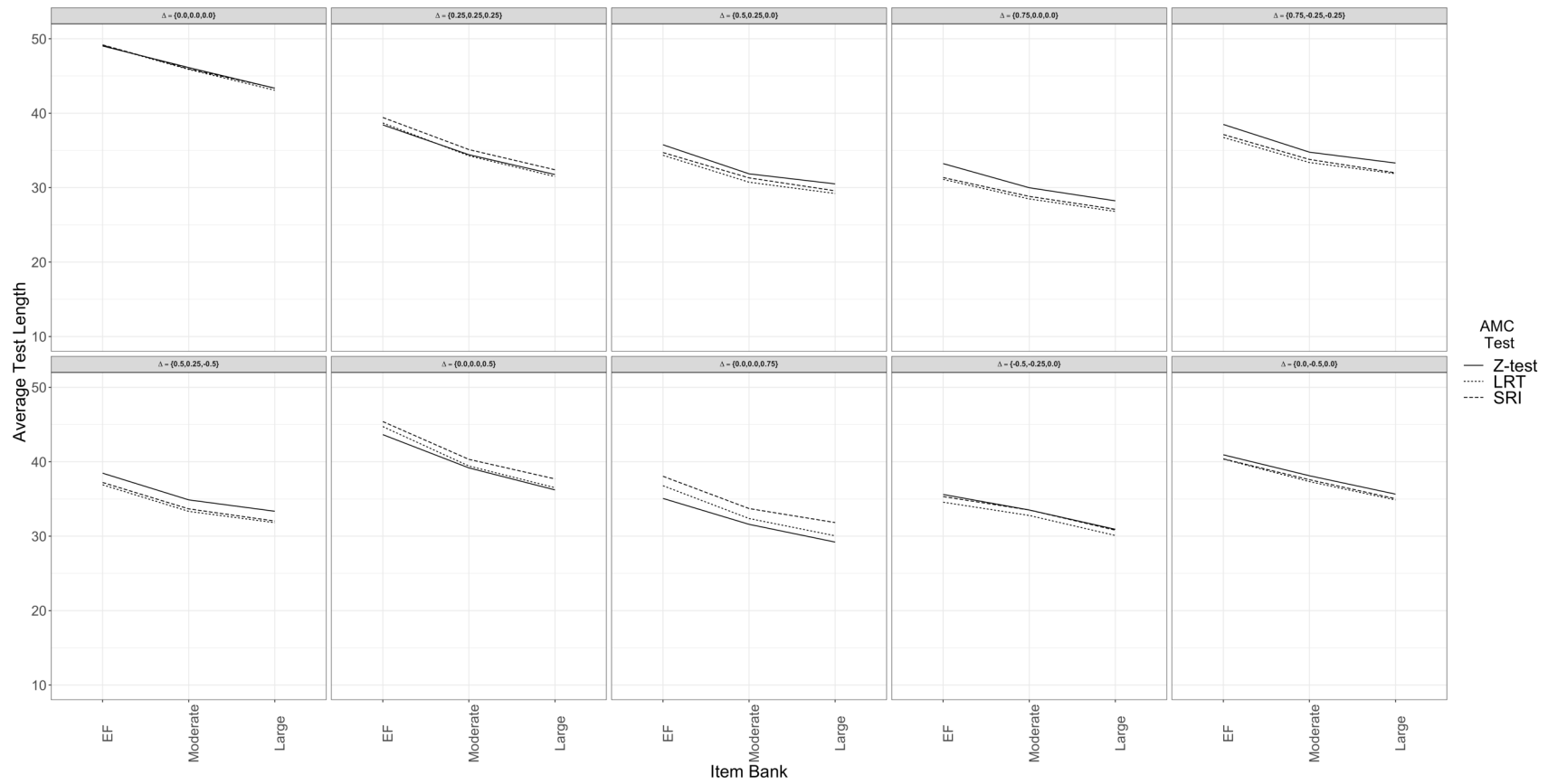


Figure S11. ATL for three AMC hypothesis tests at the fourth testing occasion across 10 θ change trajectories in Simulation 3 for item banks with no (EF), moderate, or large calibration error.

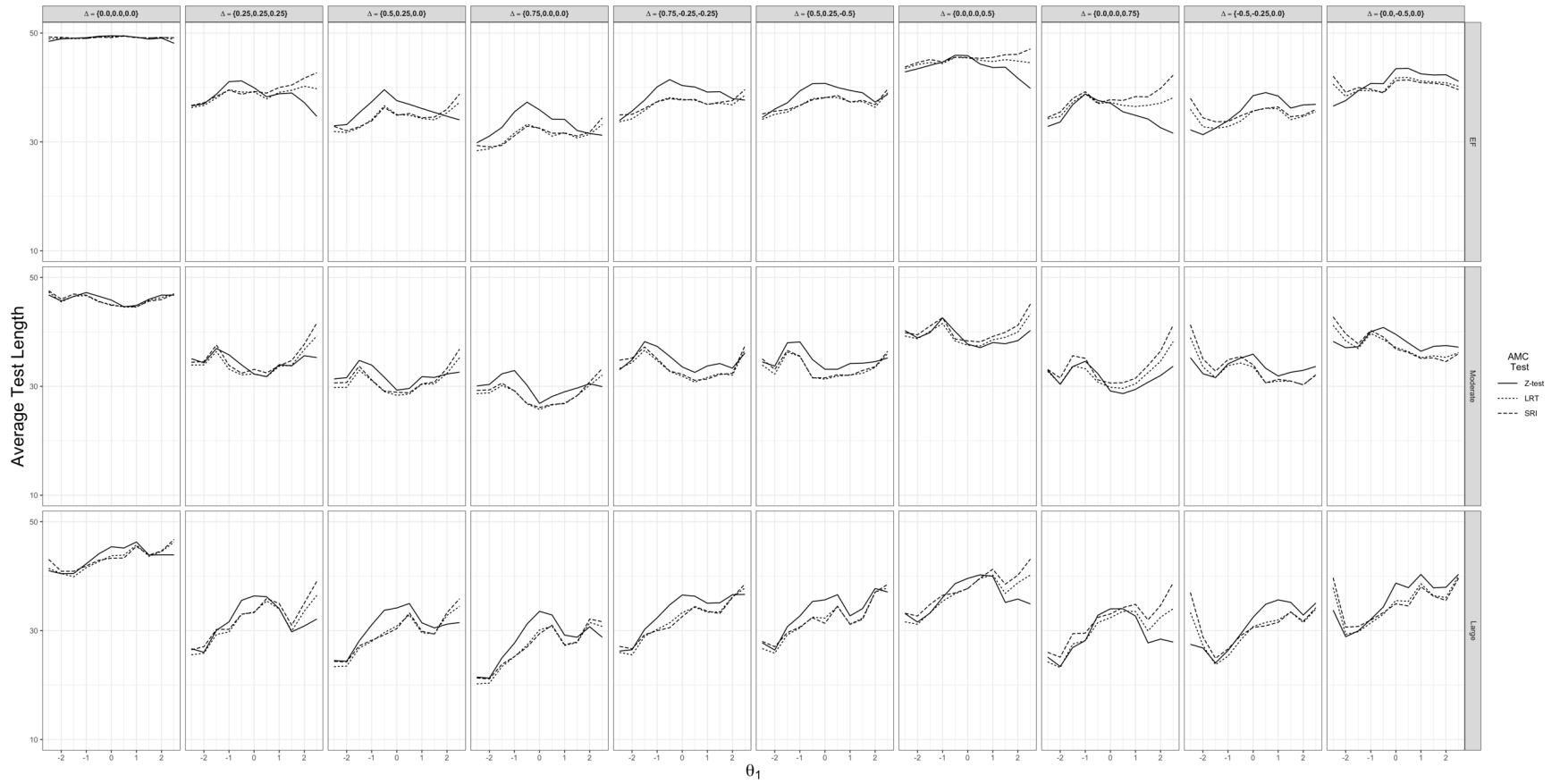


Figure S12. ATL for three AMC hypothesis tests across 10 θ trajectories when conditioning on θ_1 in Simulation 3.

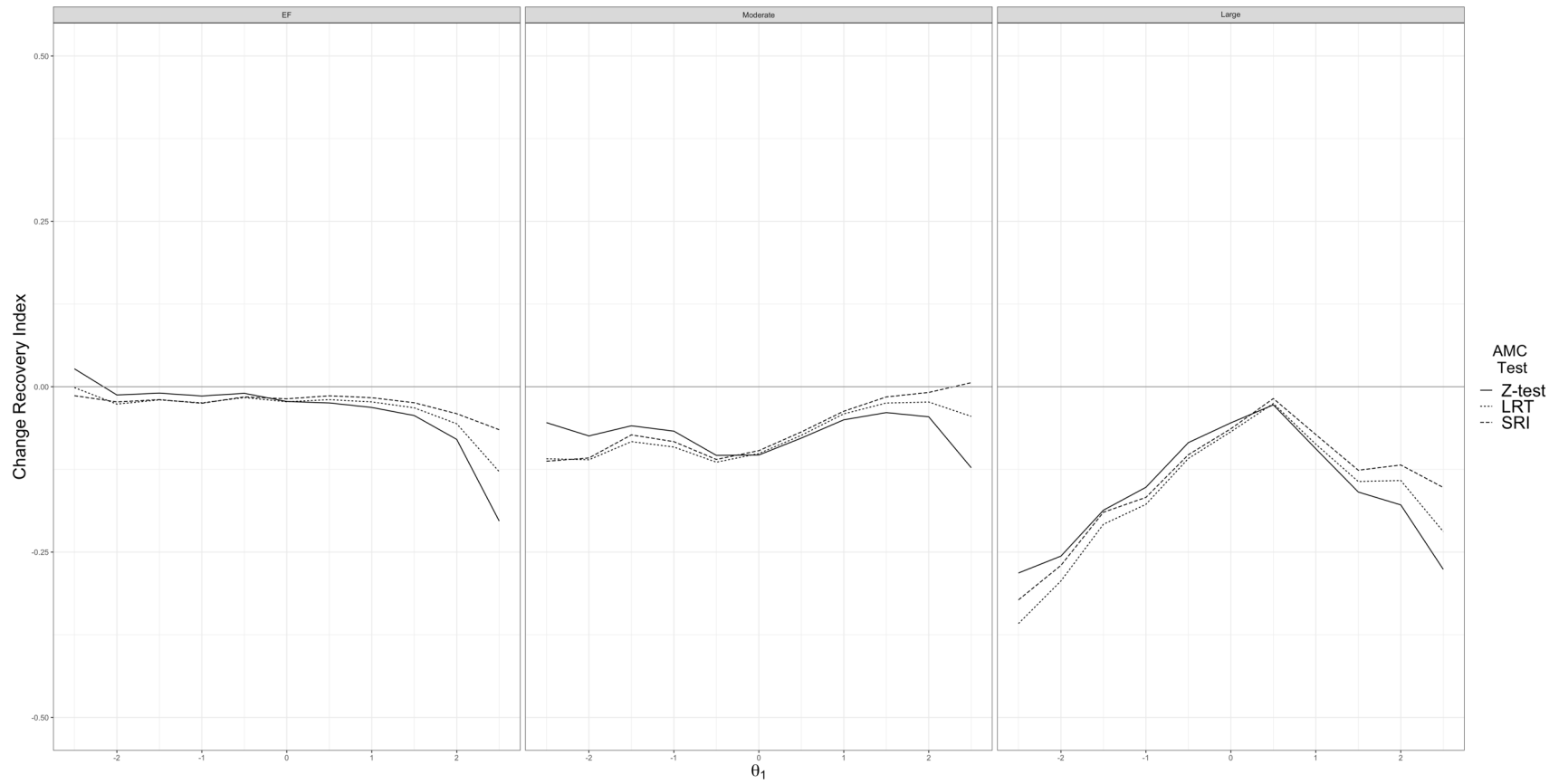


Figure S13. Change recovery index for three AMC hypothesis tests across decreasing calibration sample sizes when conditioning on θ_1 in Simulation 3.

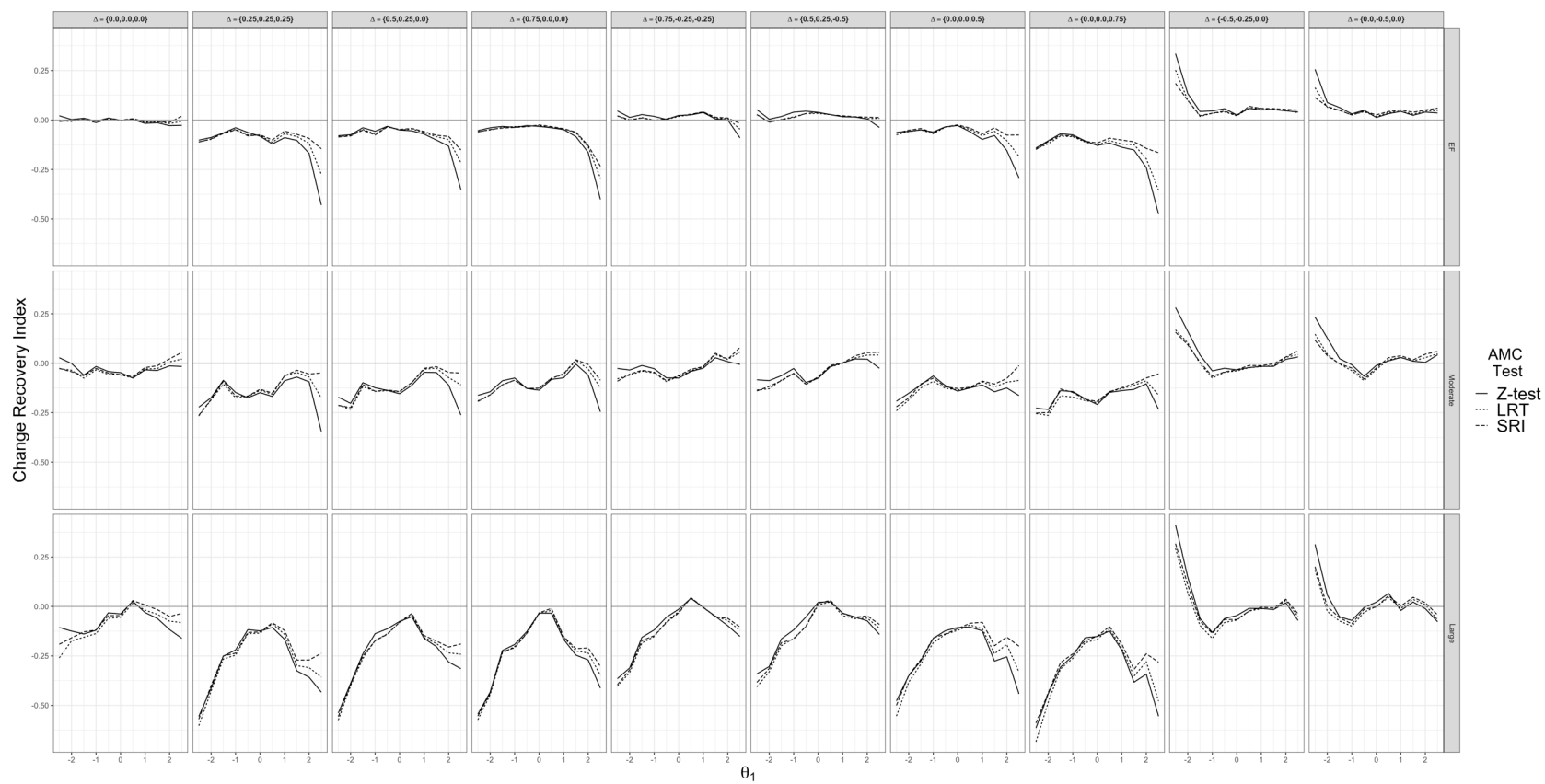


Figure S14. Change recovery index for three AMC hypothesis tests across 10 θ trajectories when conditioning on θ_1 in Simulation 3.