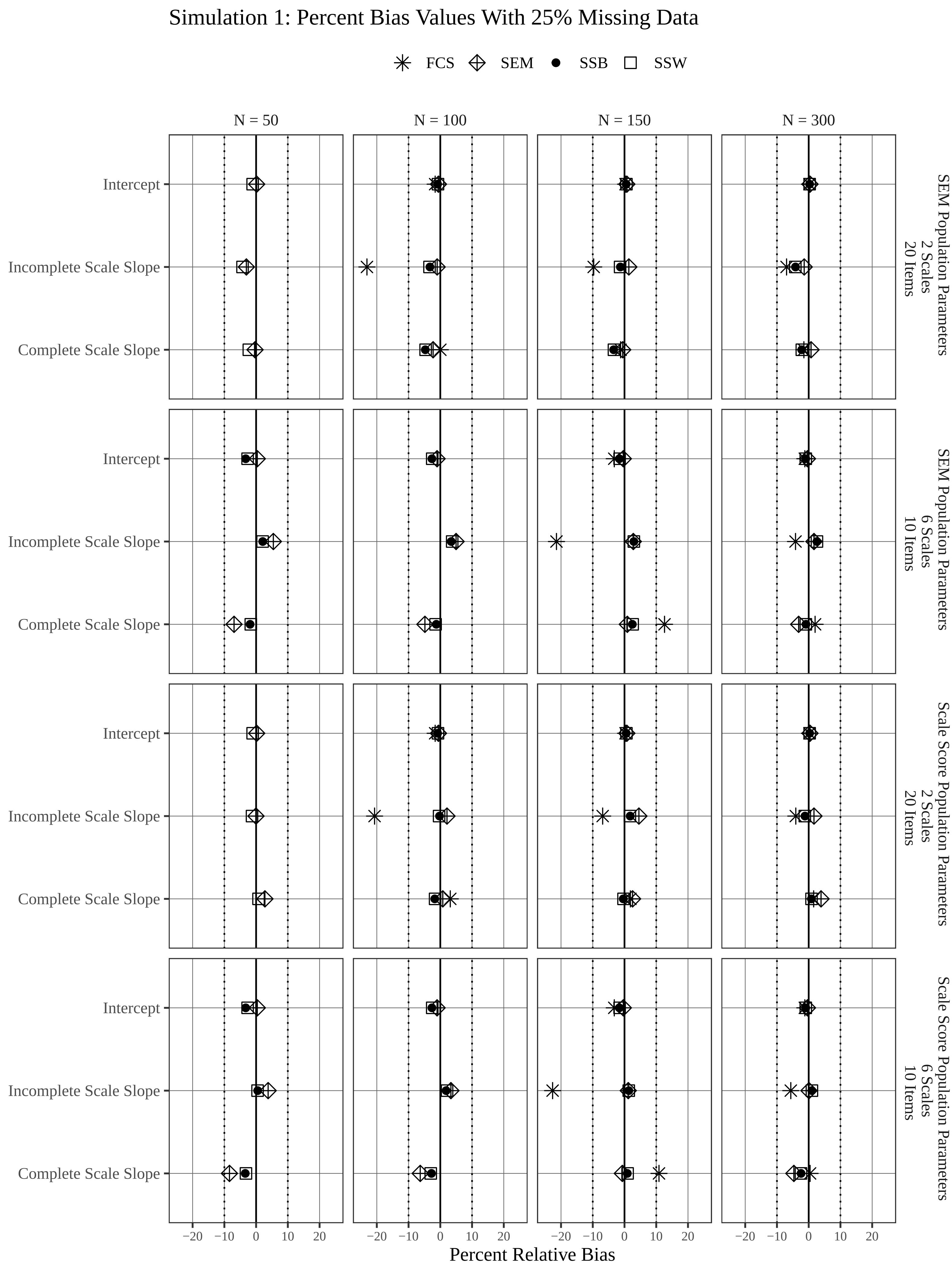
A Factored Regression Model for Composite Scores with Item-Level Missing Data

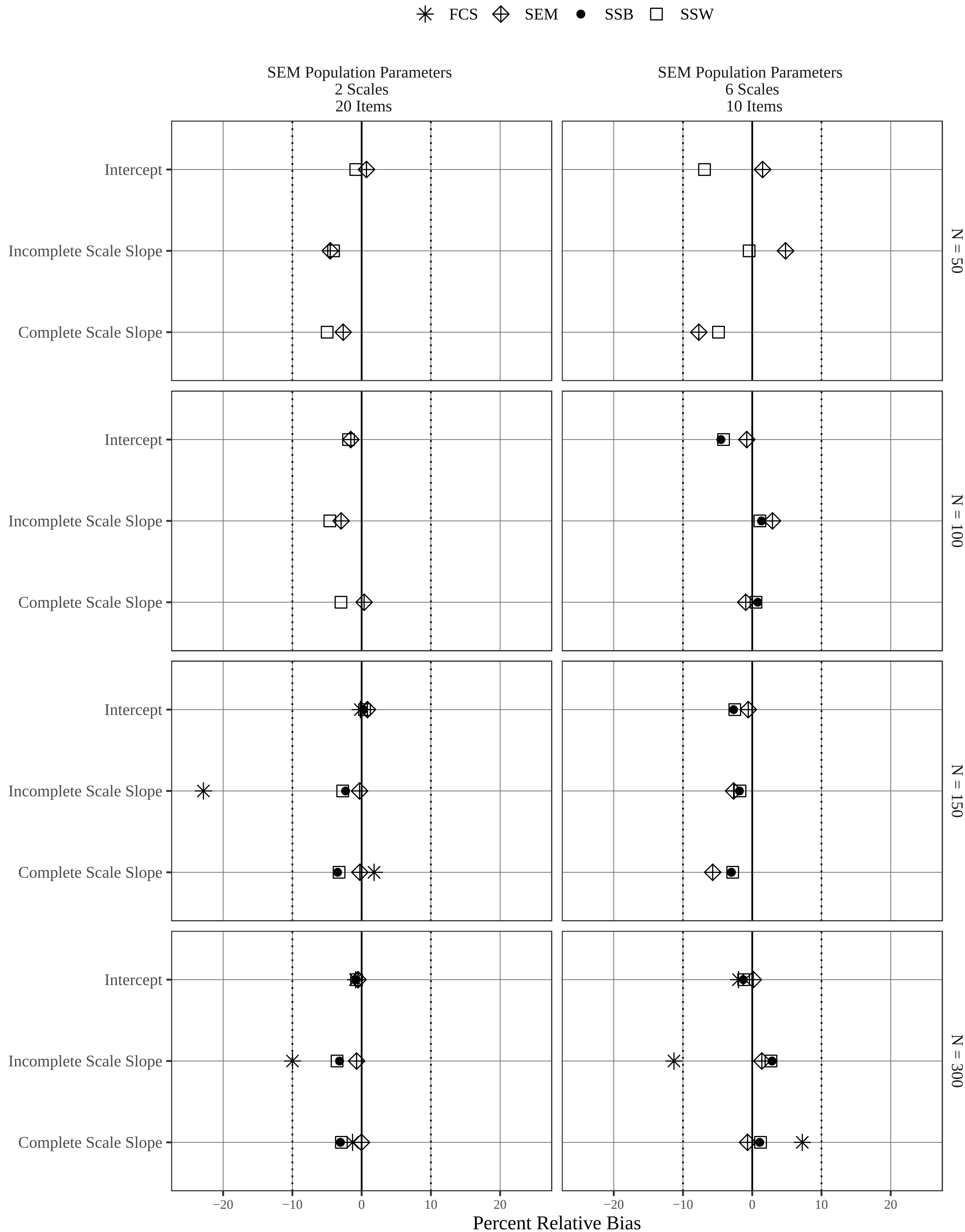
Online Supplemental Materials

Table of Contents of Graphical Summaries for Primary Simulations

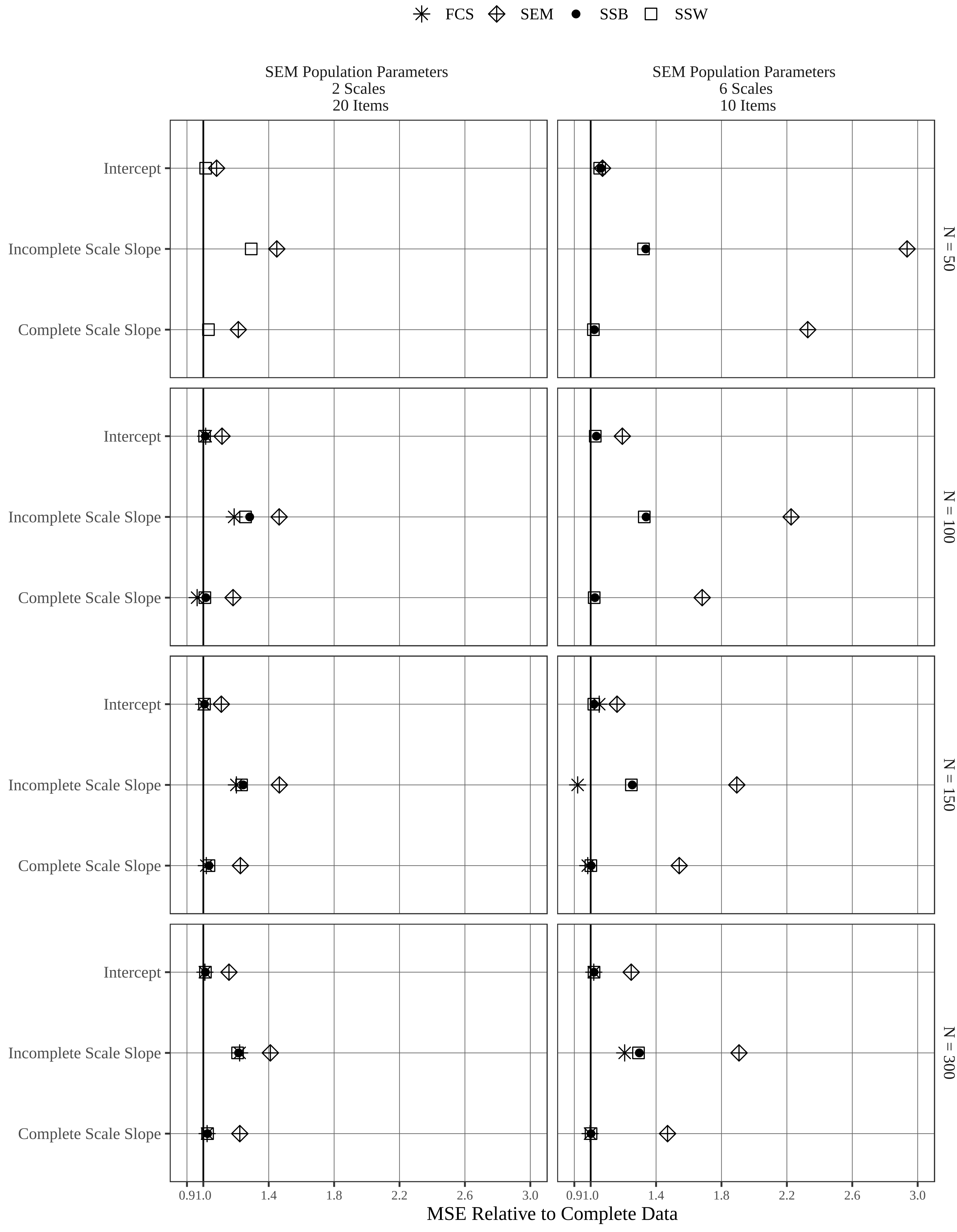


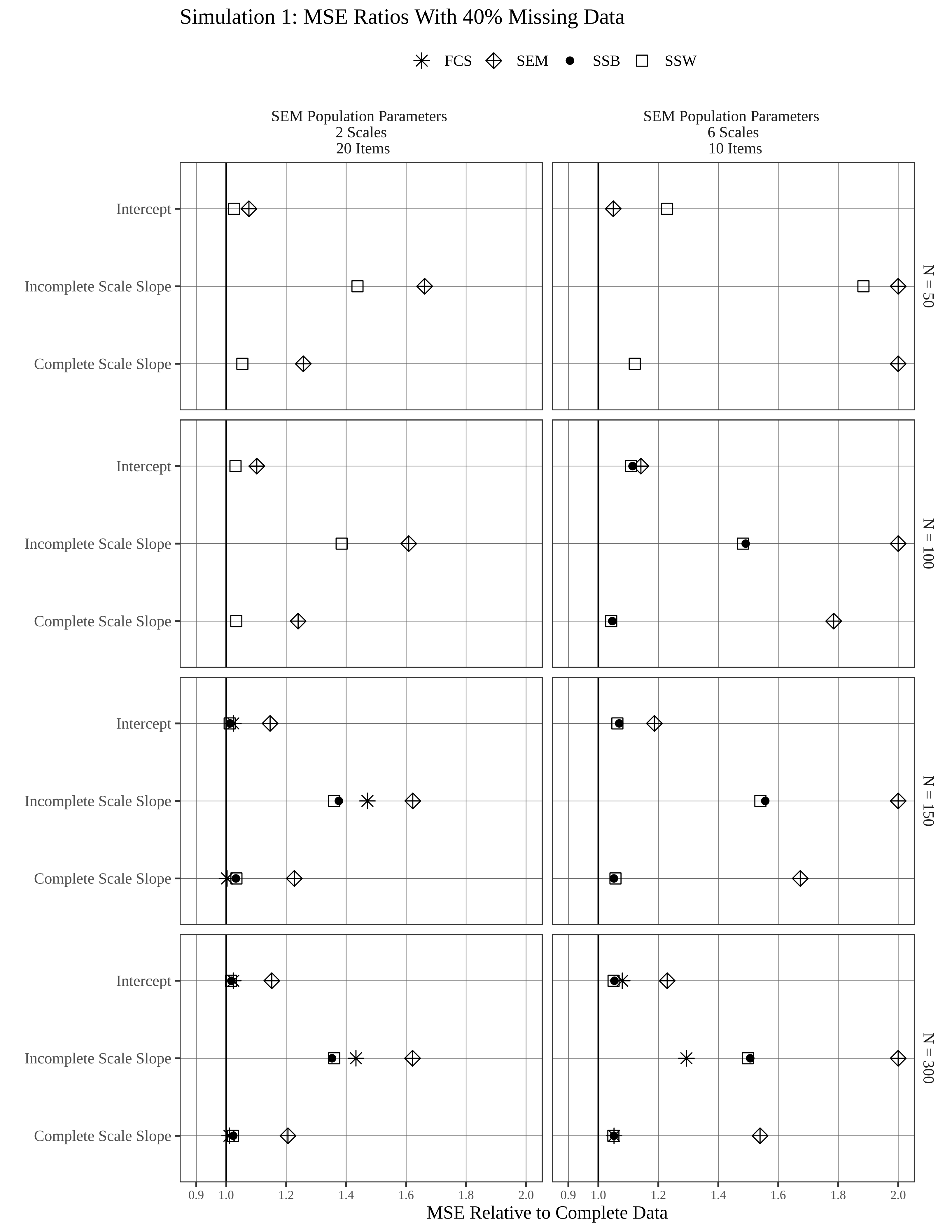
Simulation 1: Percent Bias Values With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
−20 −10 0 10 20
Percent Relative Bias

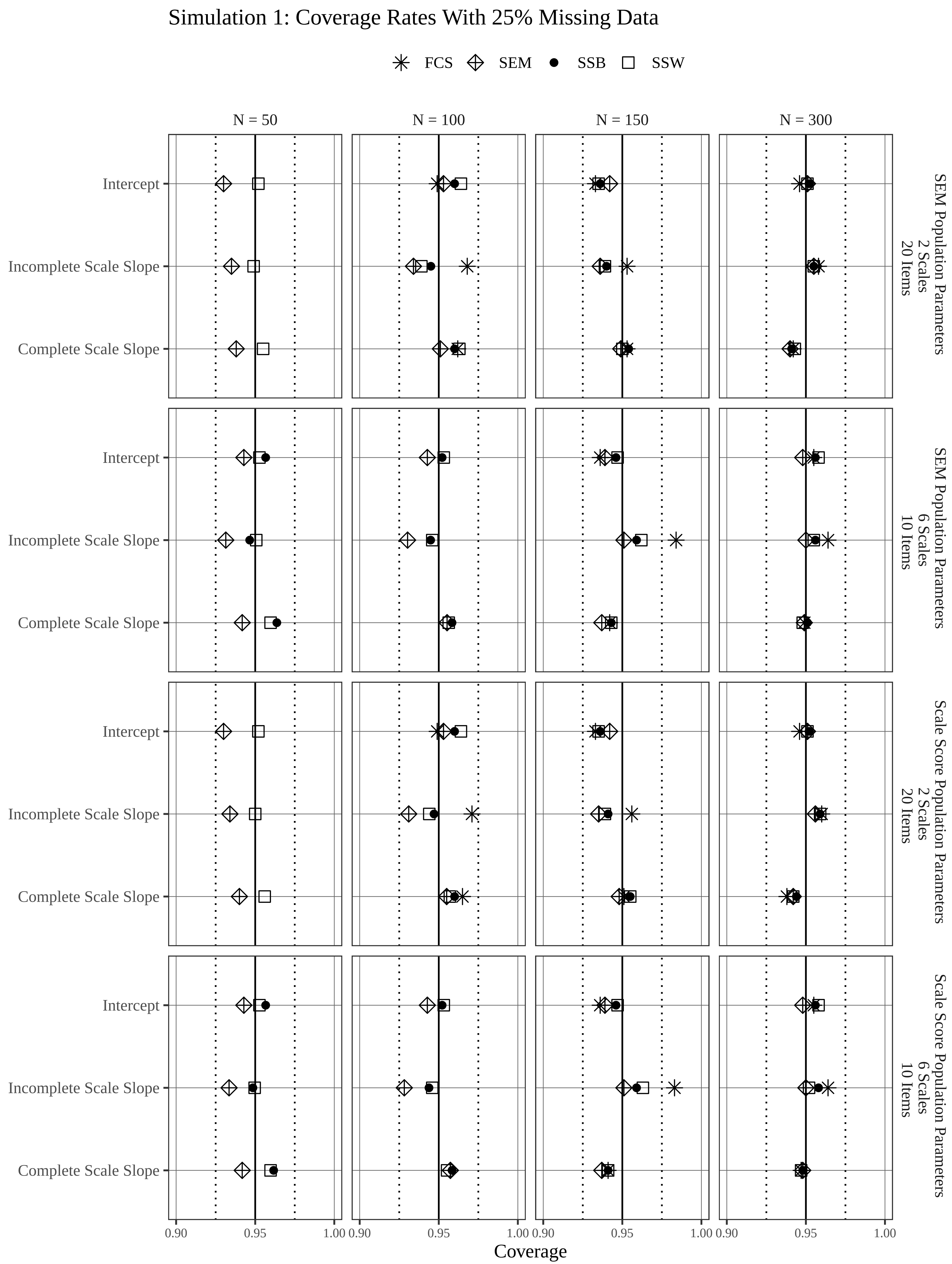
Simulation 1: Percent Bias Values With 40% Missing Data



Simulation 1: MSE Ratios With 25% Missing Data

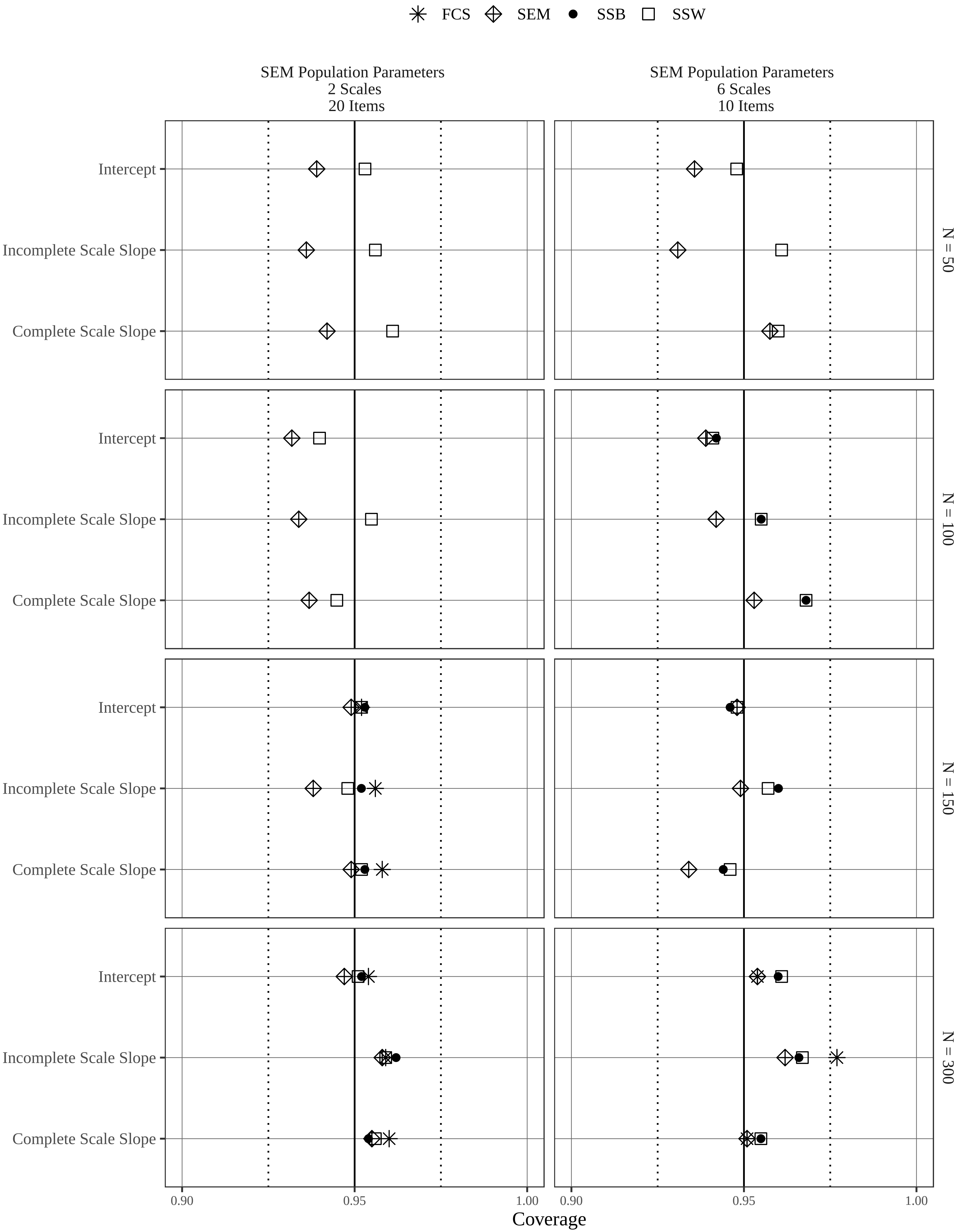


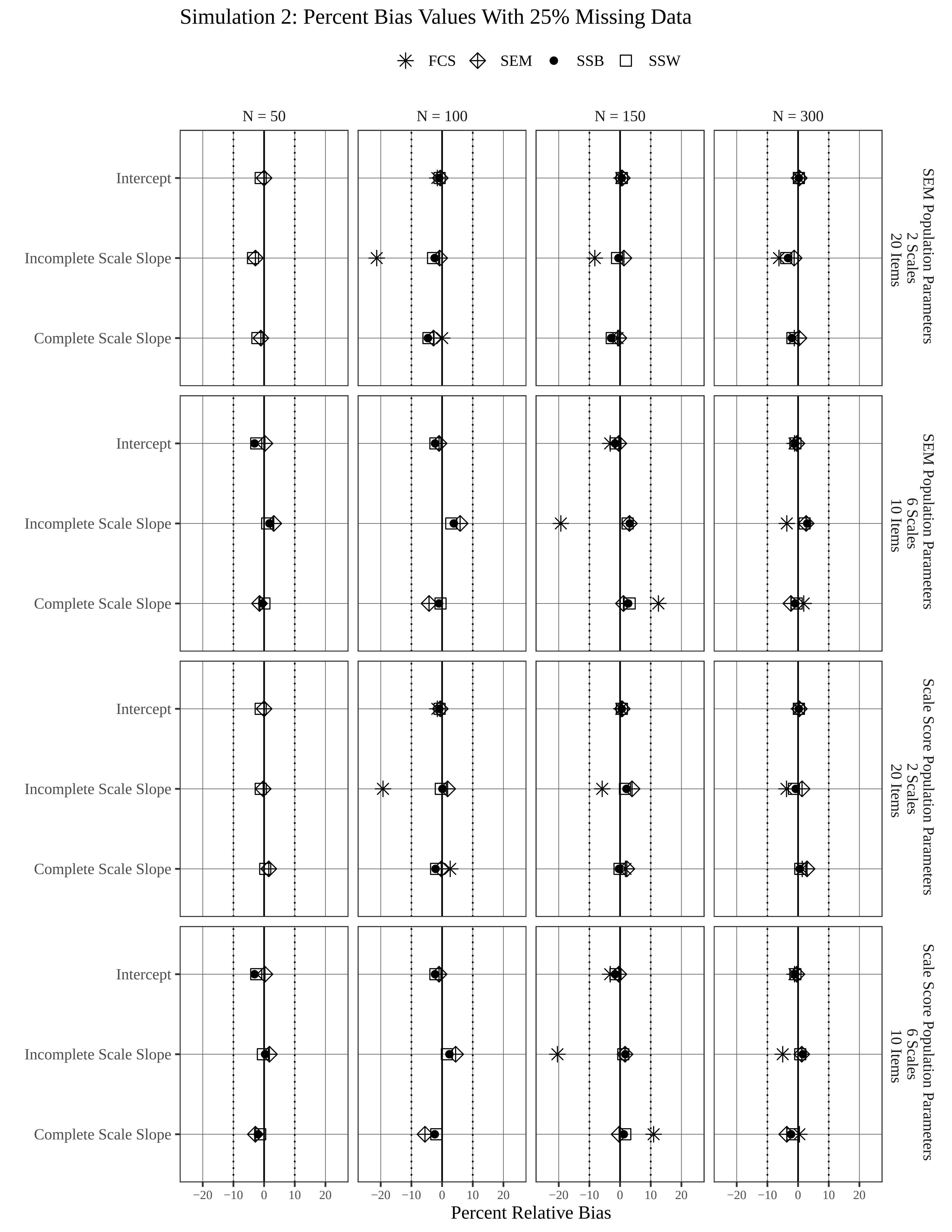

Simulation 1: MSE Ratios With 40% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
0.9
1.0
1.2
1.4
1.6
1.8
2.0
MSE Relative to Complete Data



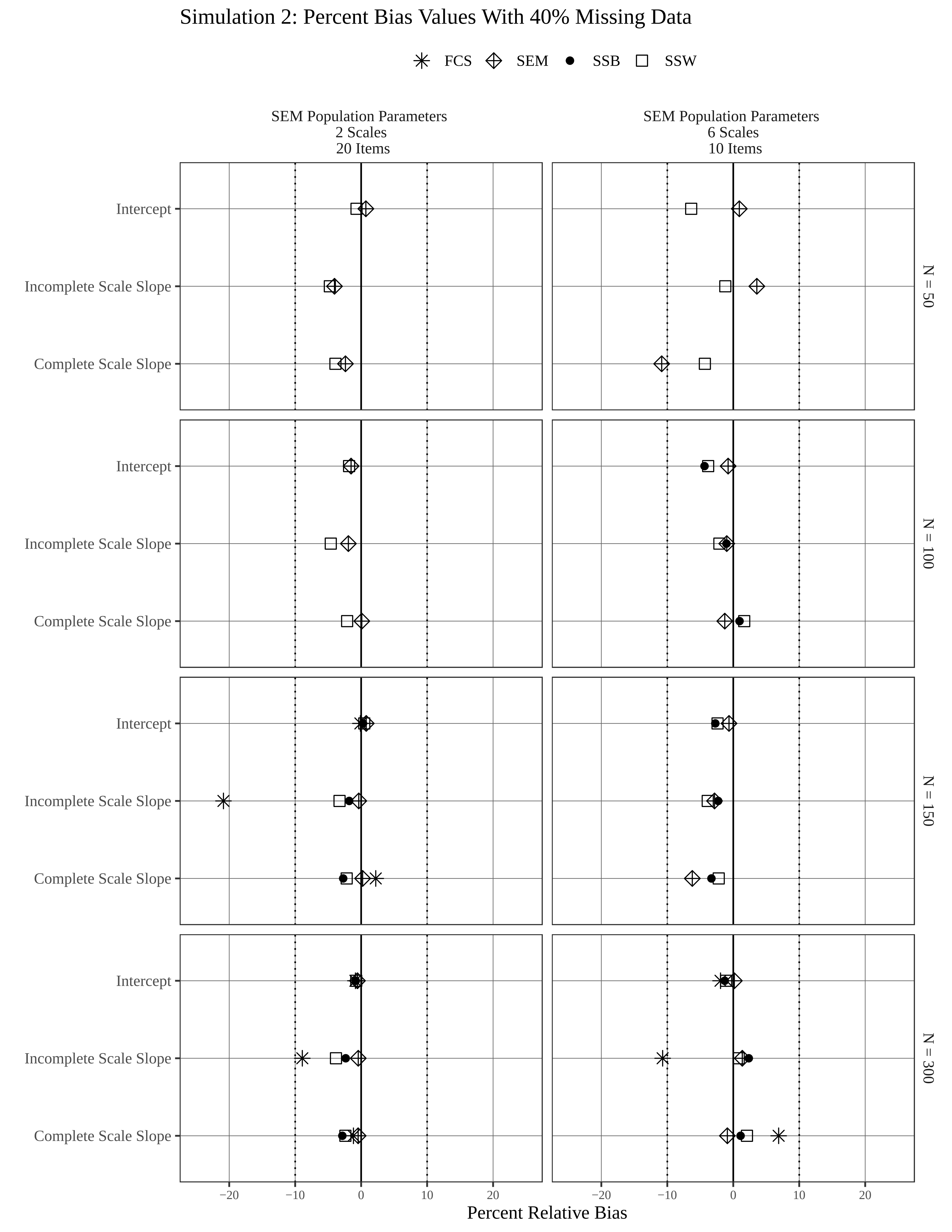
Simulation 1: Coverage Rates With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
0.90
0.95
1.00
Coverage

Simulation 1: Coverage Rates With 40% Missing Data



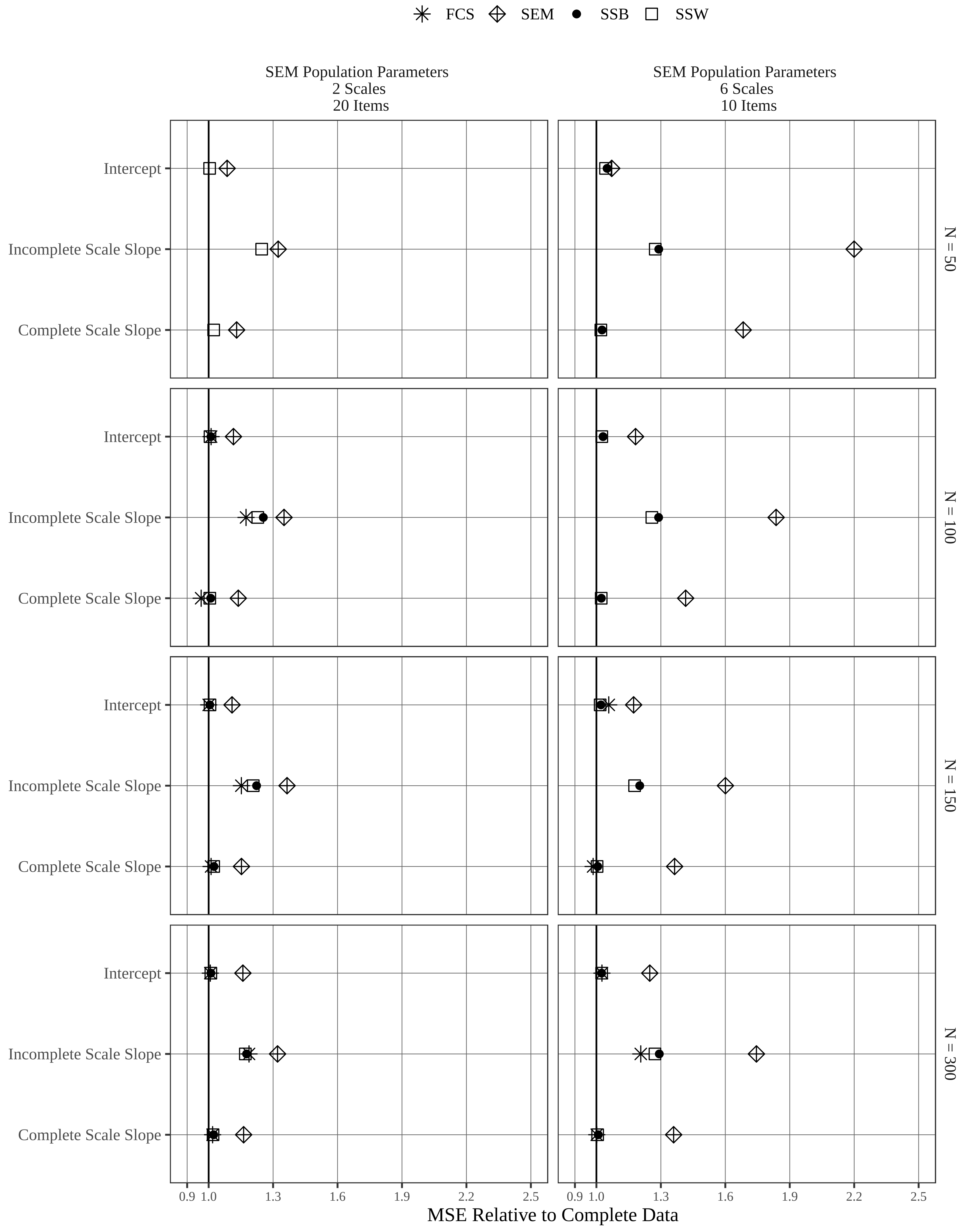


Simulation 2: Percent Bias Values With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
−20 −10 0 10 20
Percent Relative Bias

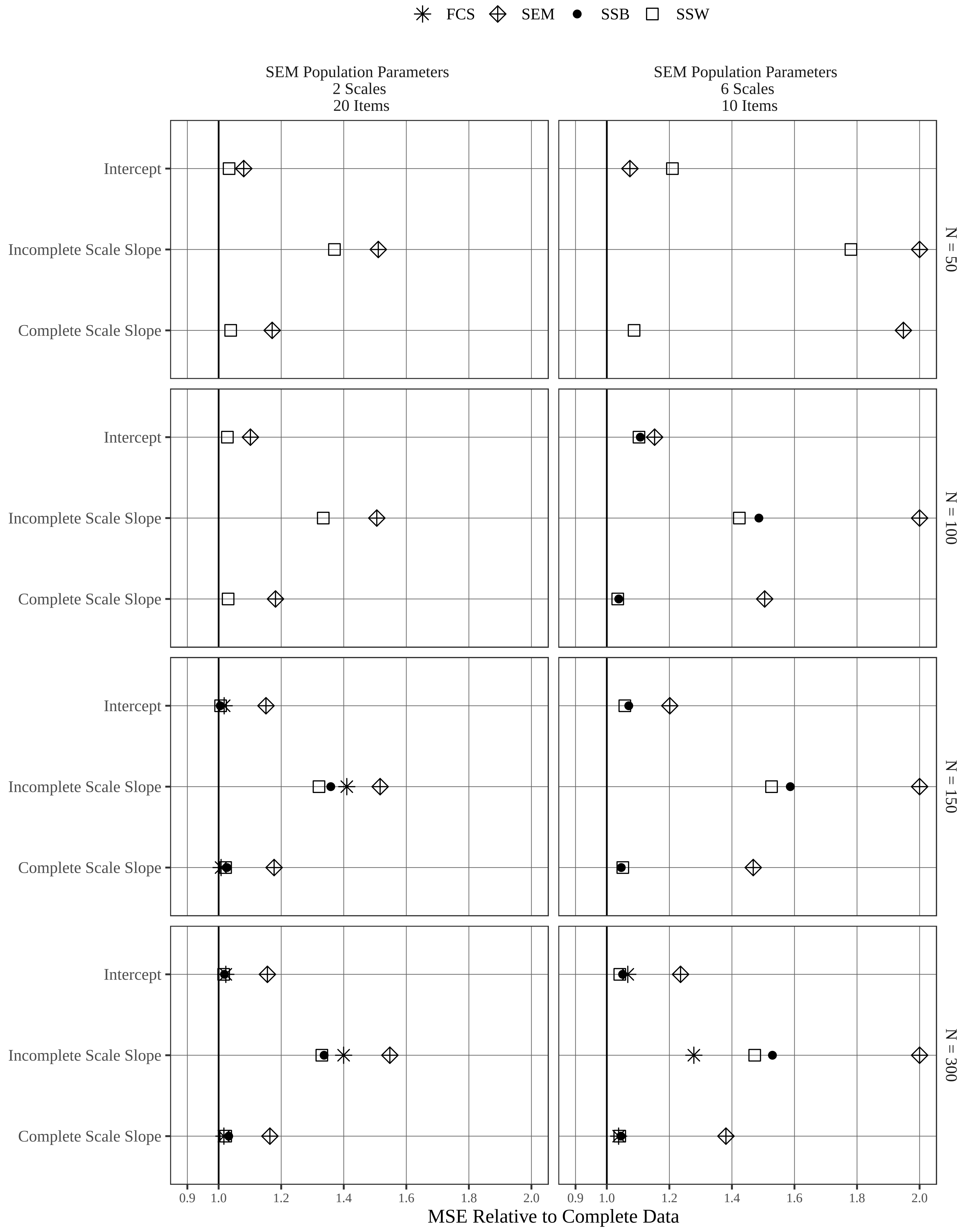


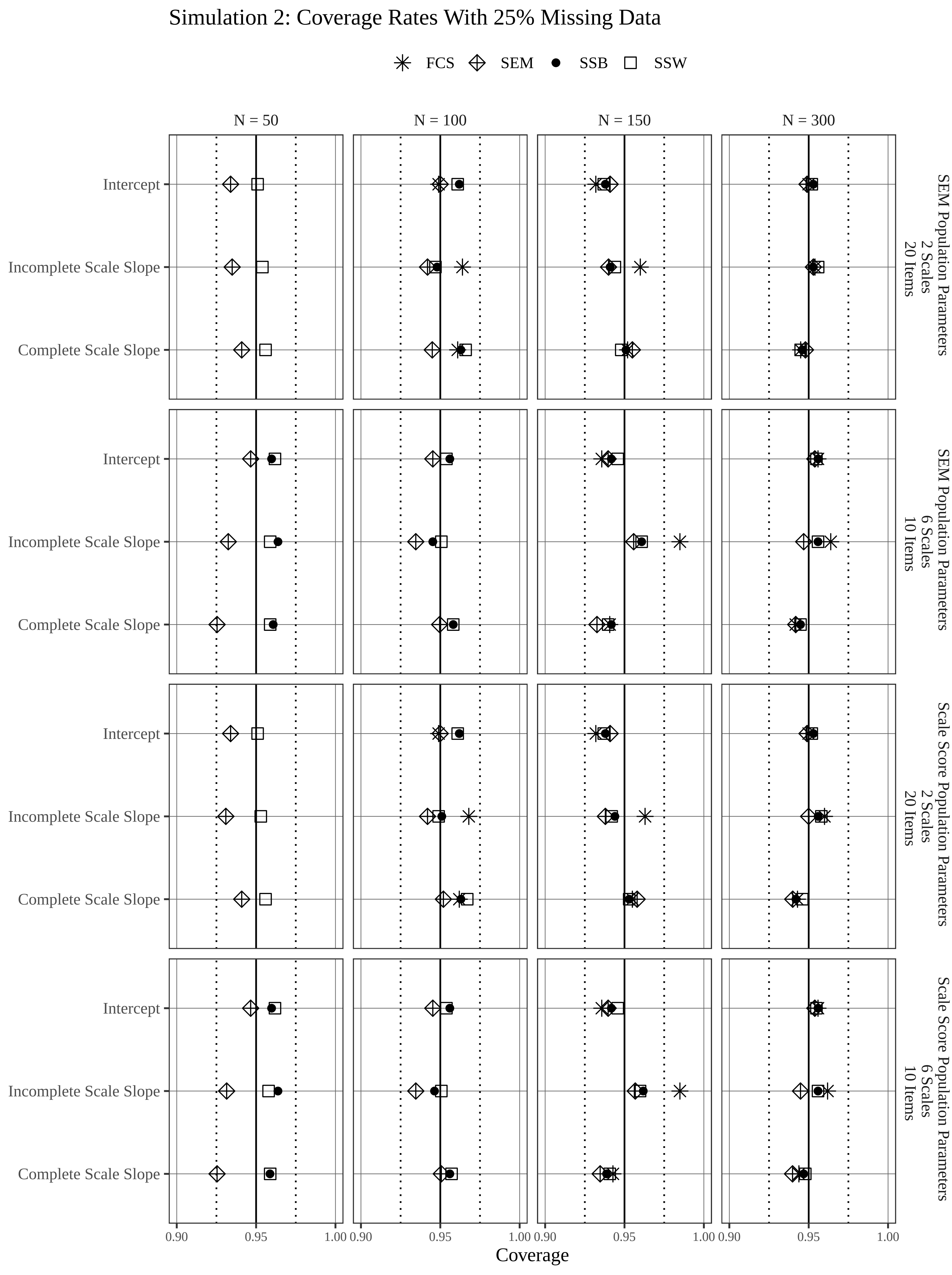
Simulation 2: Percent Bias Values With 40% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
−20
−10
0
10
20
Percent Relative Bias

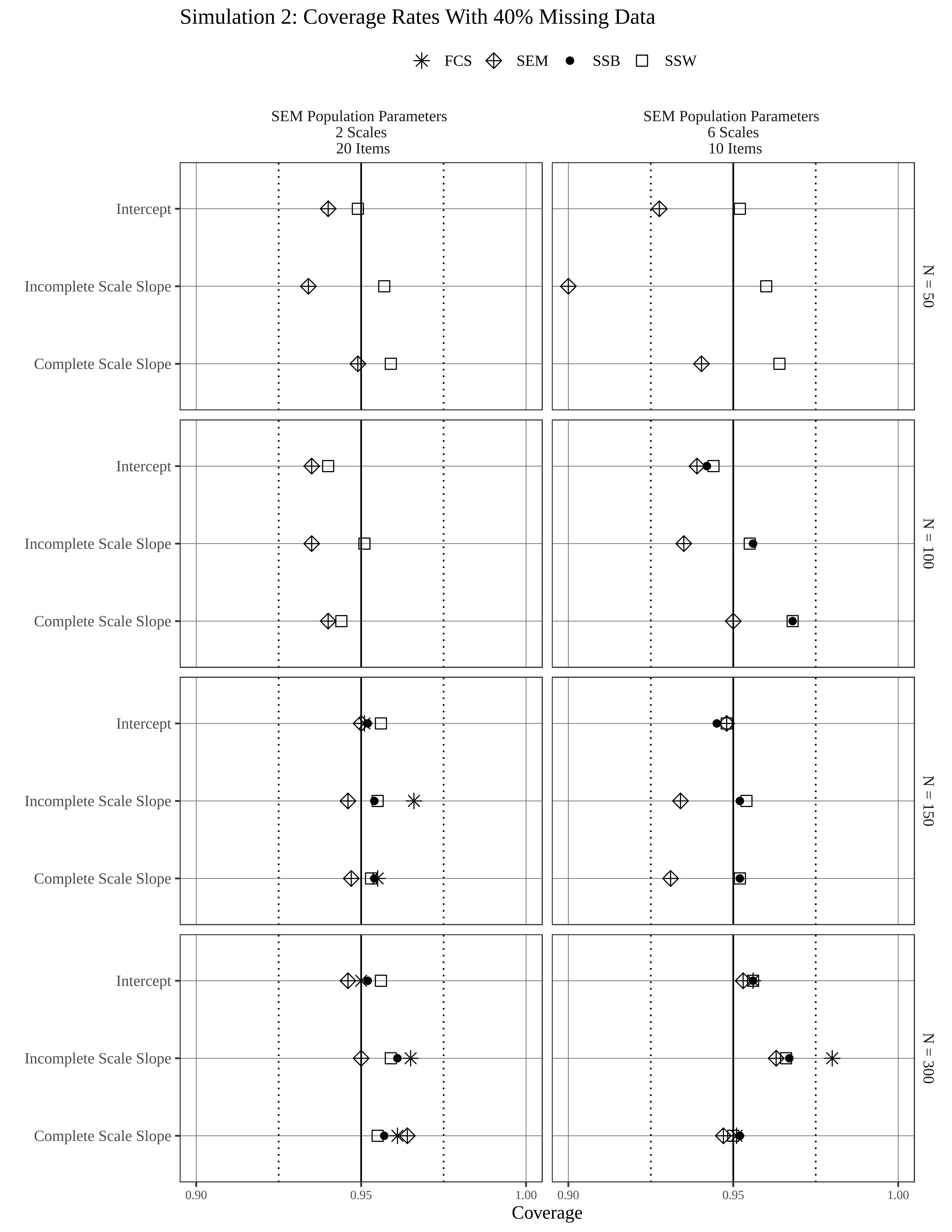
Simulation 2: MSE Ratios With 25% Missing Data



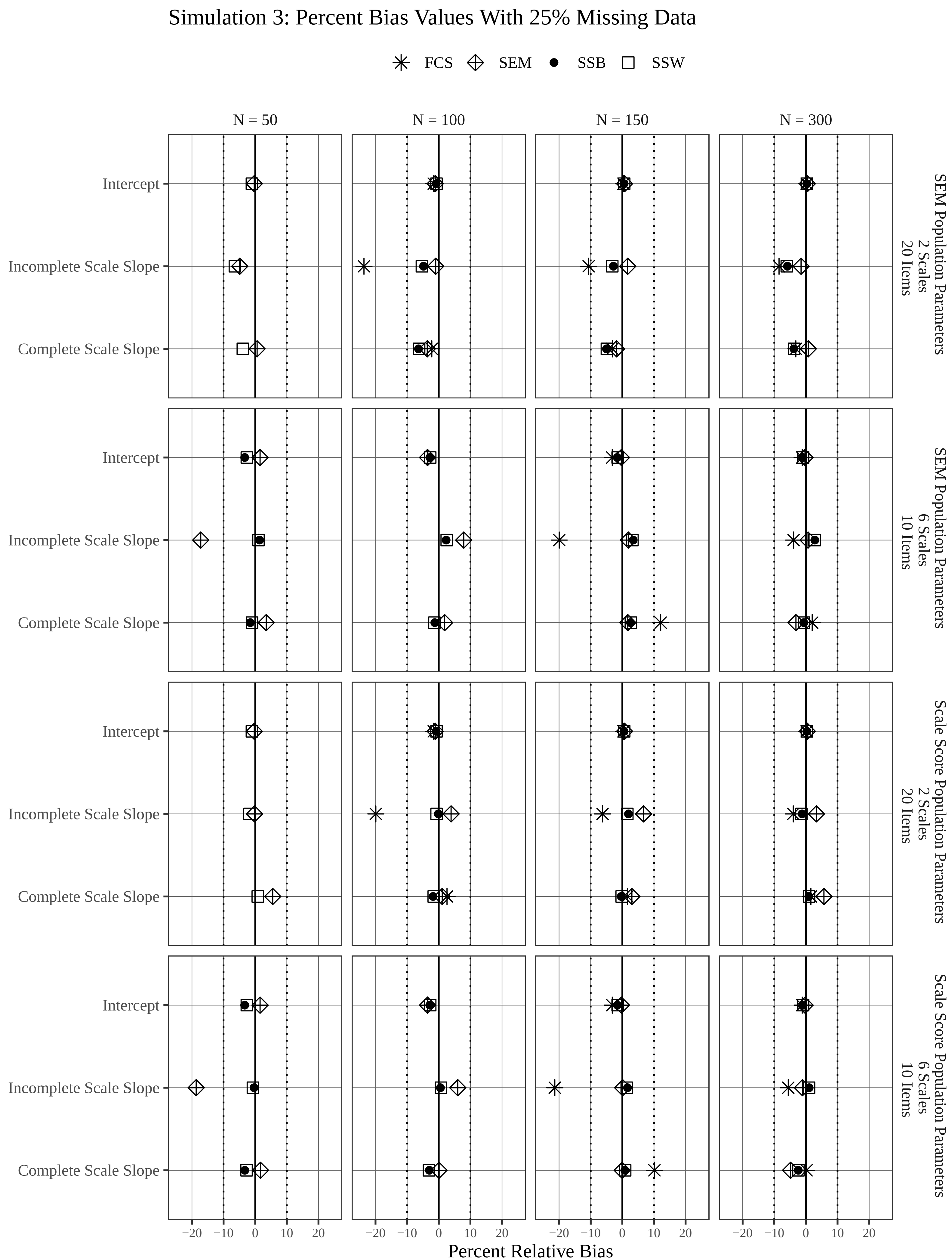
Simulation 2: MSE Ratios With 40% Missing Data



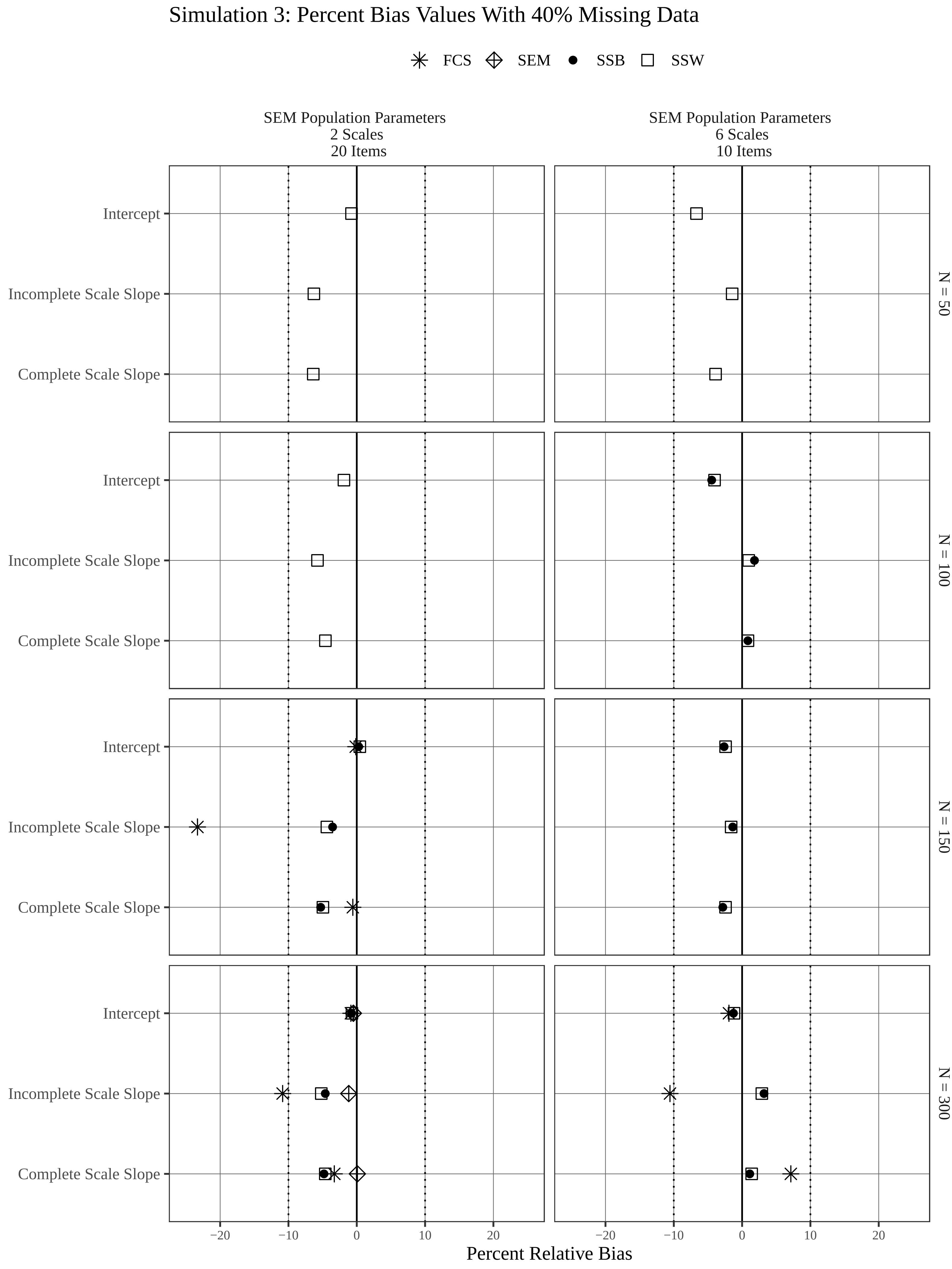

Simulation 2: Coverage Rates With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
0.90
0.95
1.00
Coverage

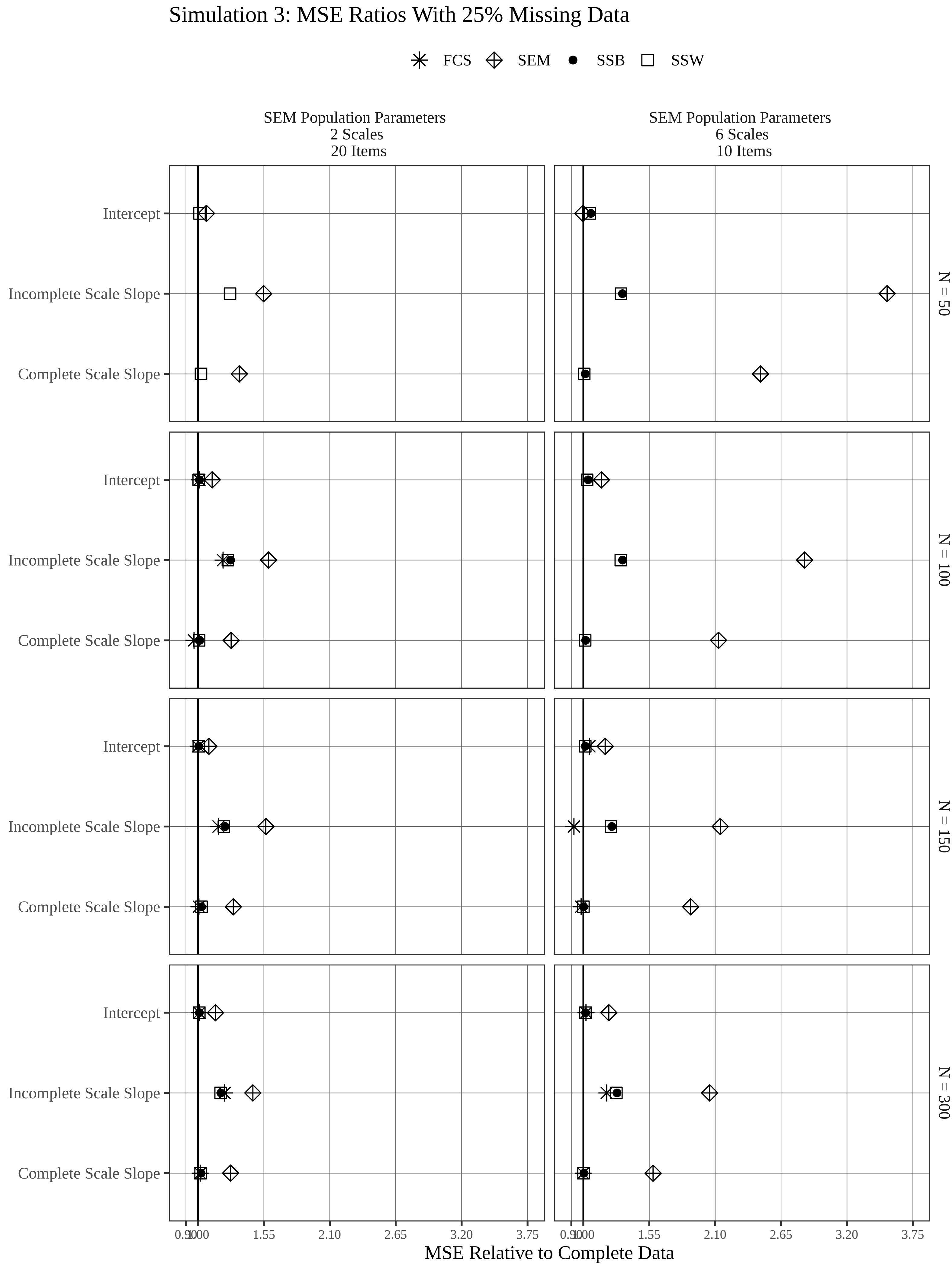


Simulation 2: Coverage Rates With 40% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
0.90
0.95
1.00
Coverage

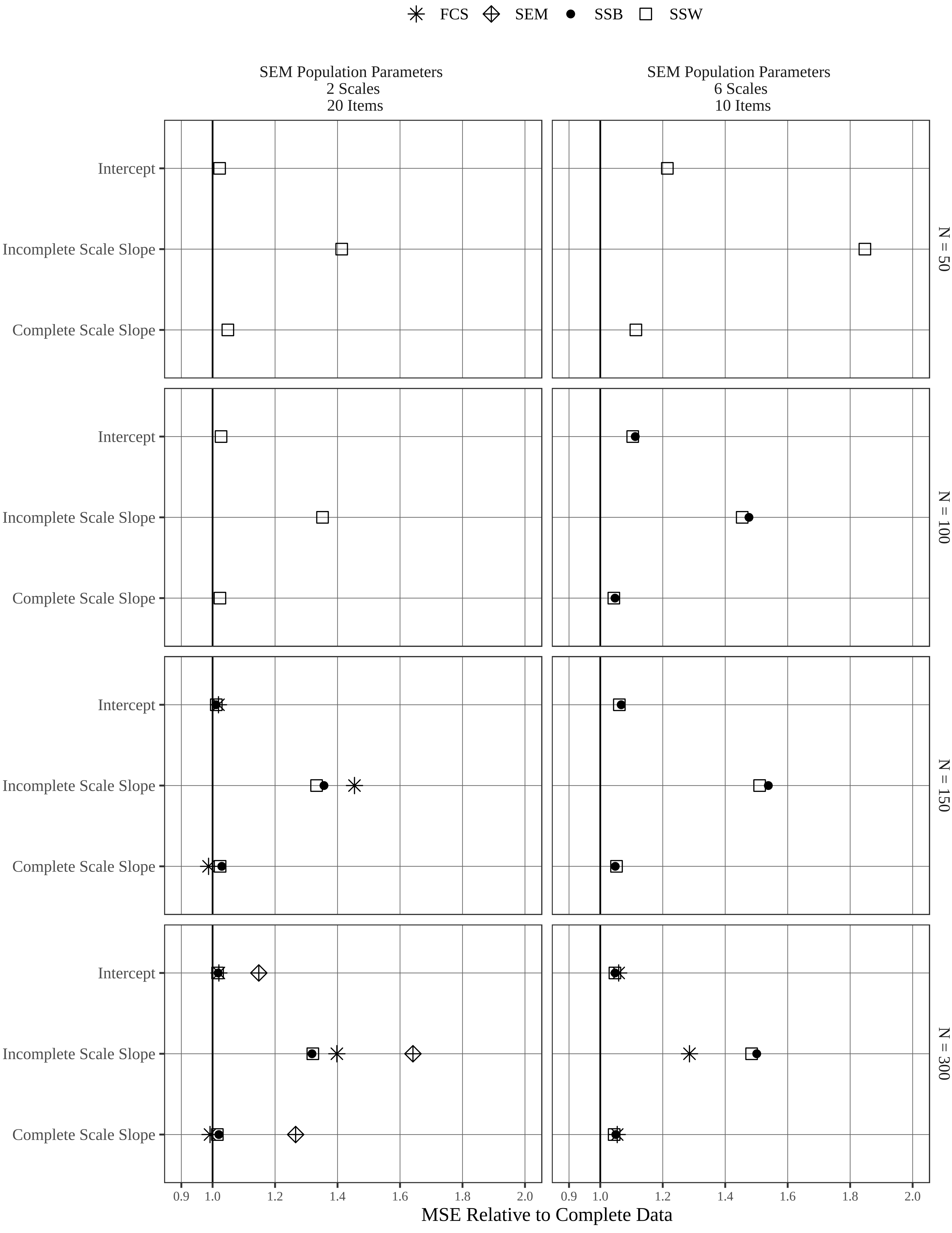


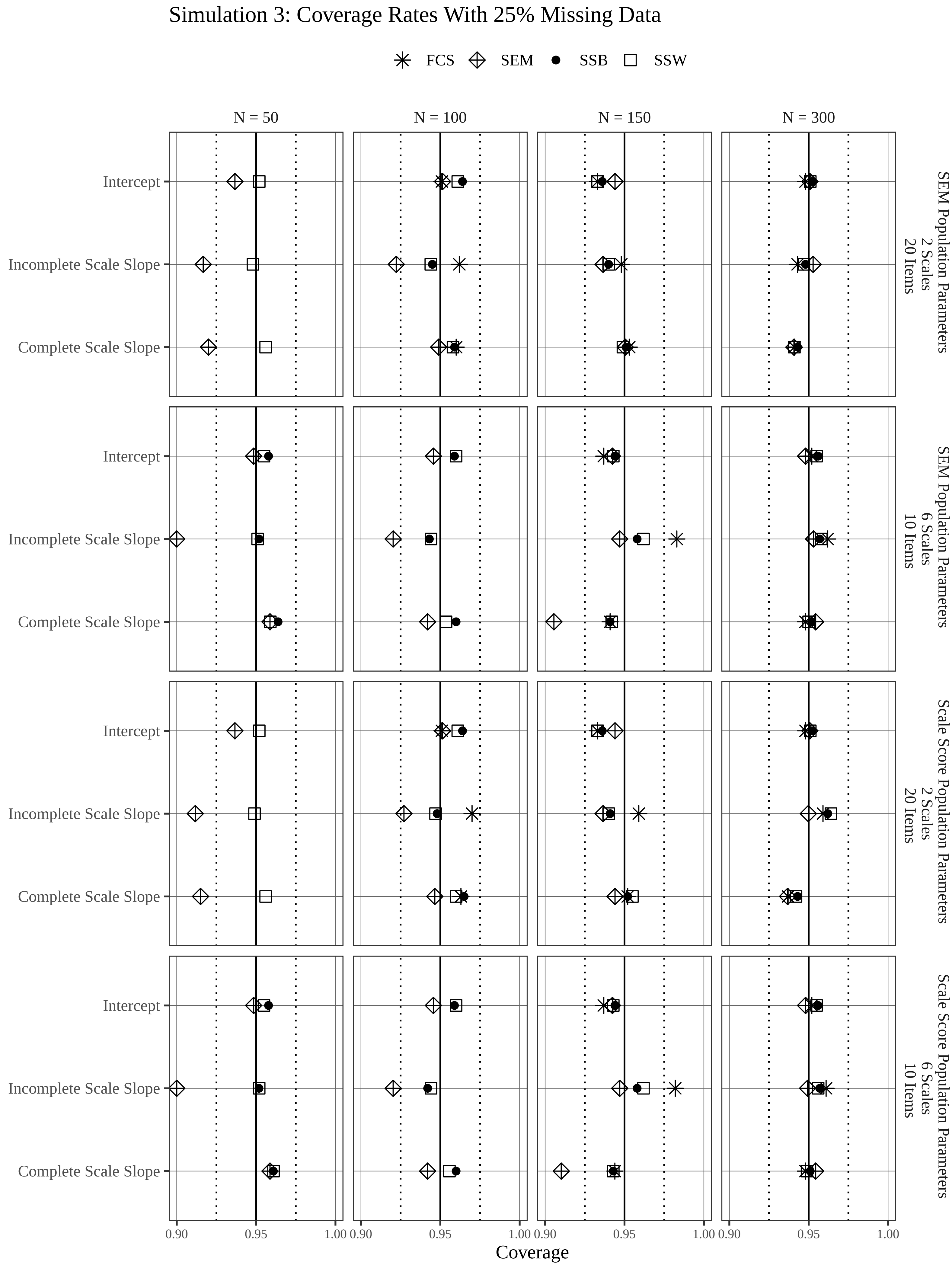
Simulation 3: Percent Bias Values With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
−20
−10
0
10
20
Percent Relative Bias

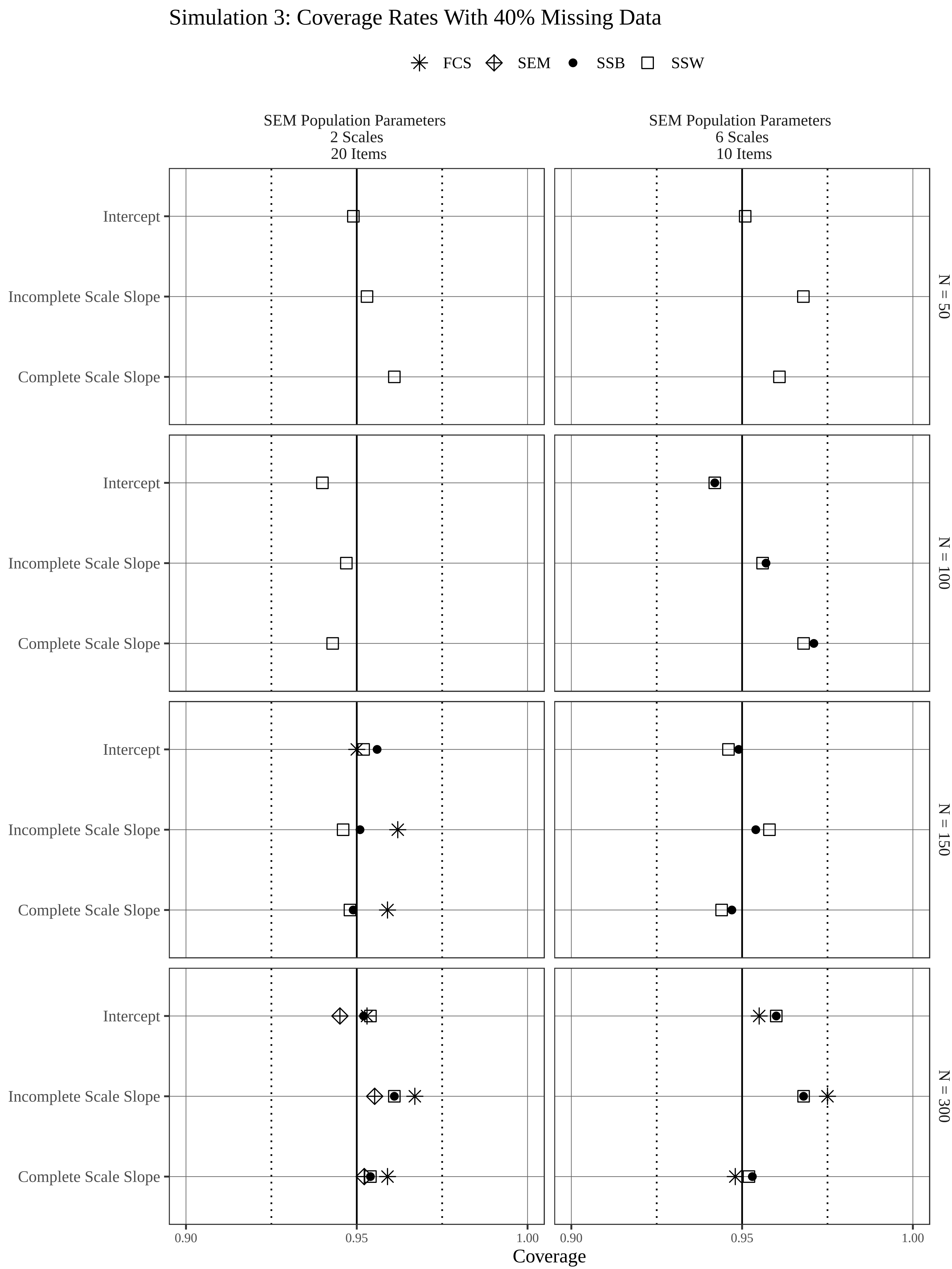

Simulation 3: Percent Bias Values With 40% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
−20
−10
0
10
20
Percent Relative Bias


Simulation 3: MSE Ratios With 25% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
0.90
1.00
1.55
2.10
2.65
3.20
3.75
MSE Relative to Complete Data

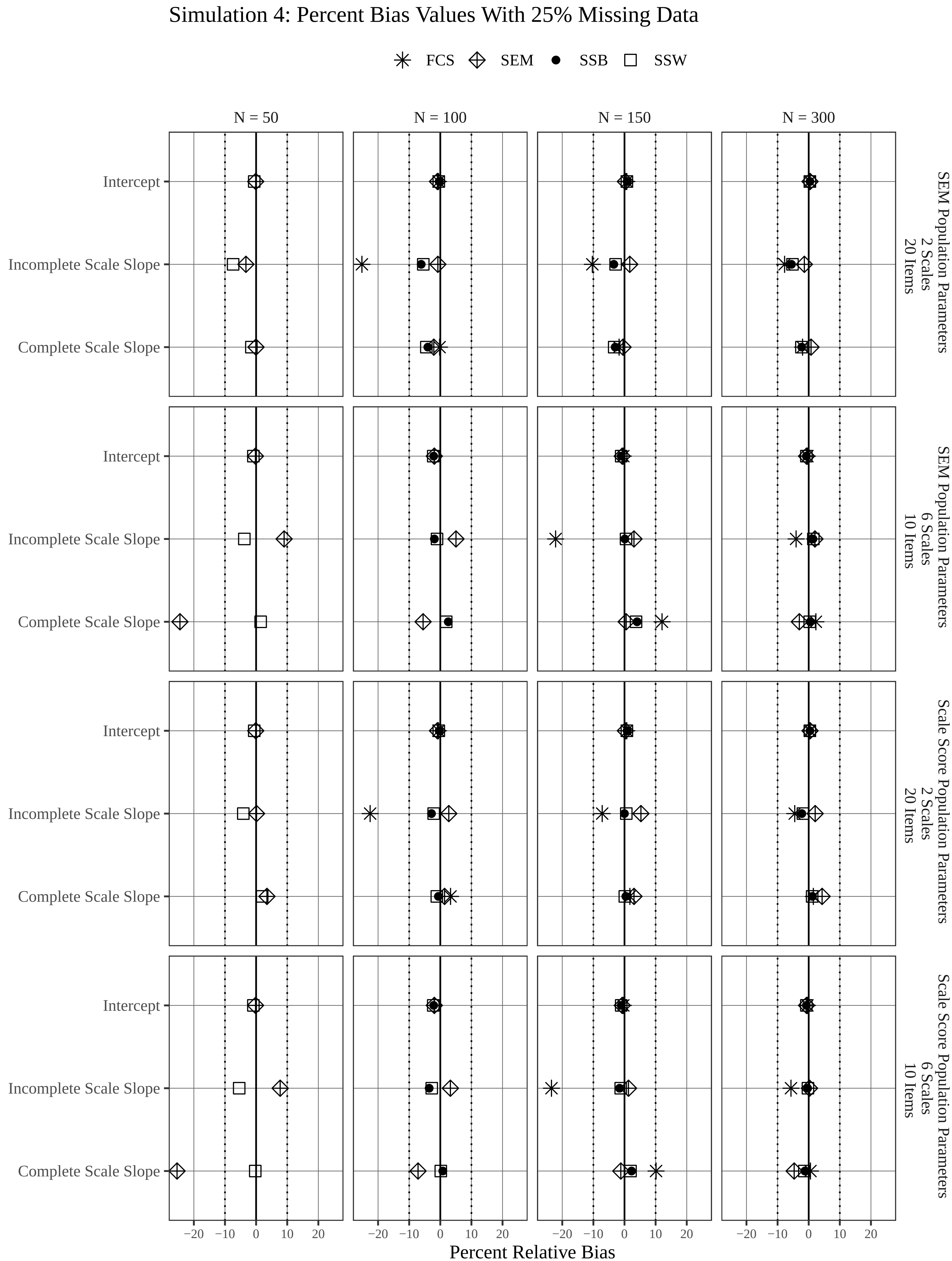
Simulation 3: MSE Ratios With 40% Missing Data



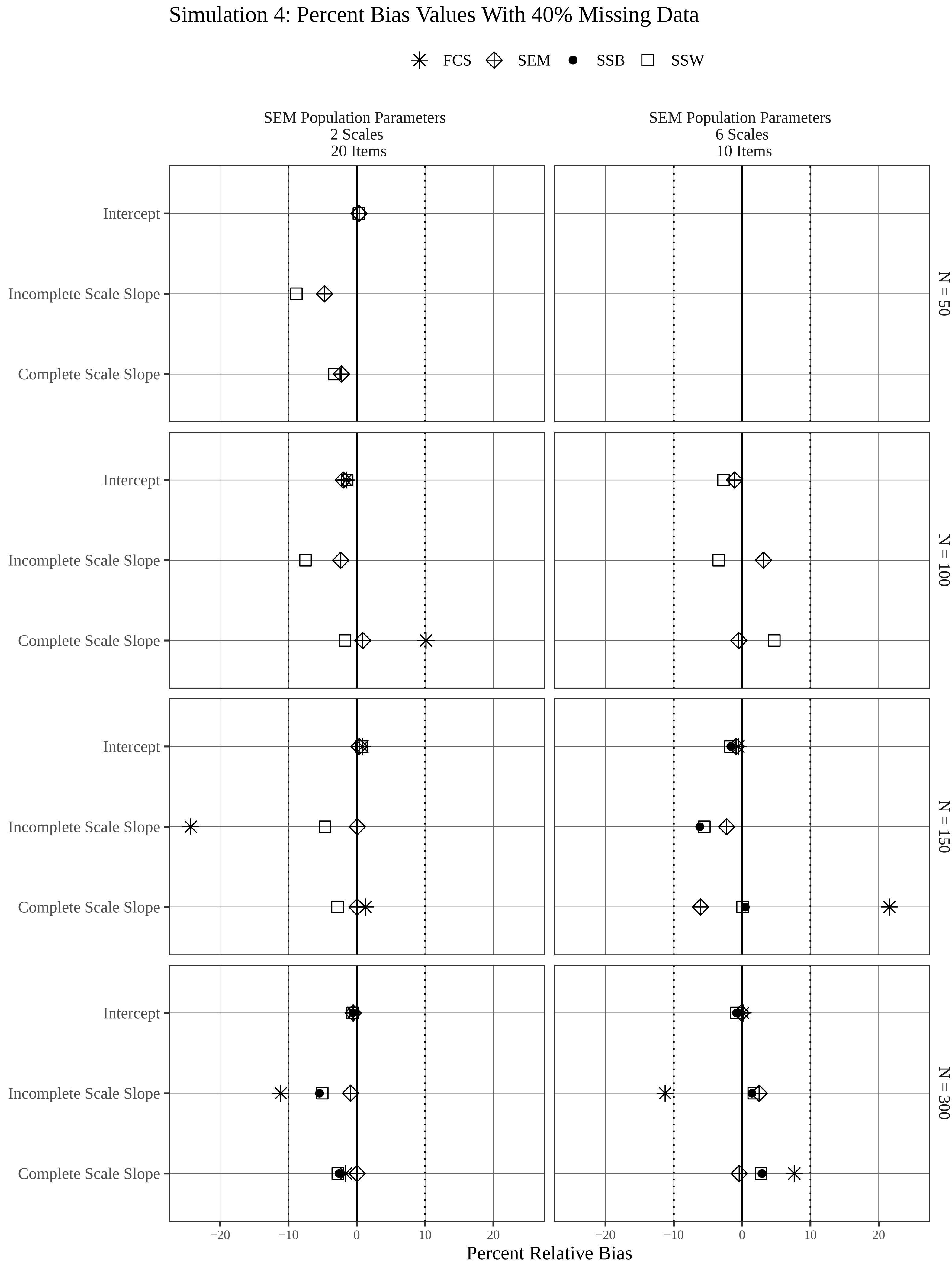

Simulation 3: Coverage Rates With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
0.90
0.95
1.00
Coverage

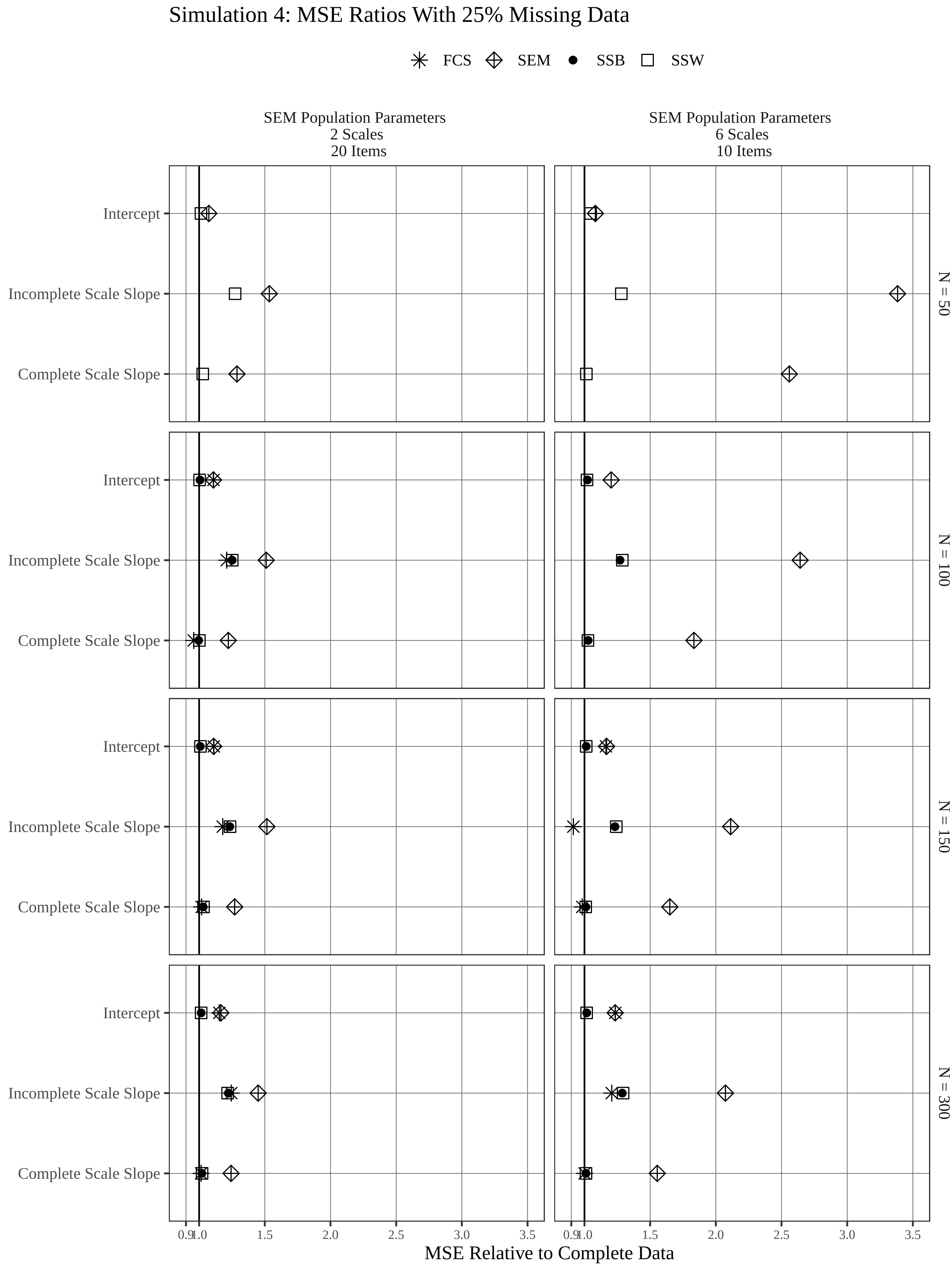


Simulation 3: Coverage Rates With 40% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
0.90
0.95
1.00
Coverage



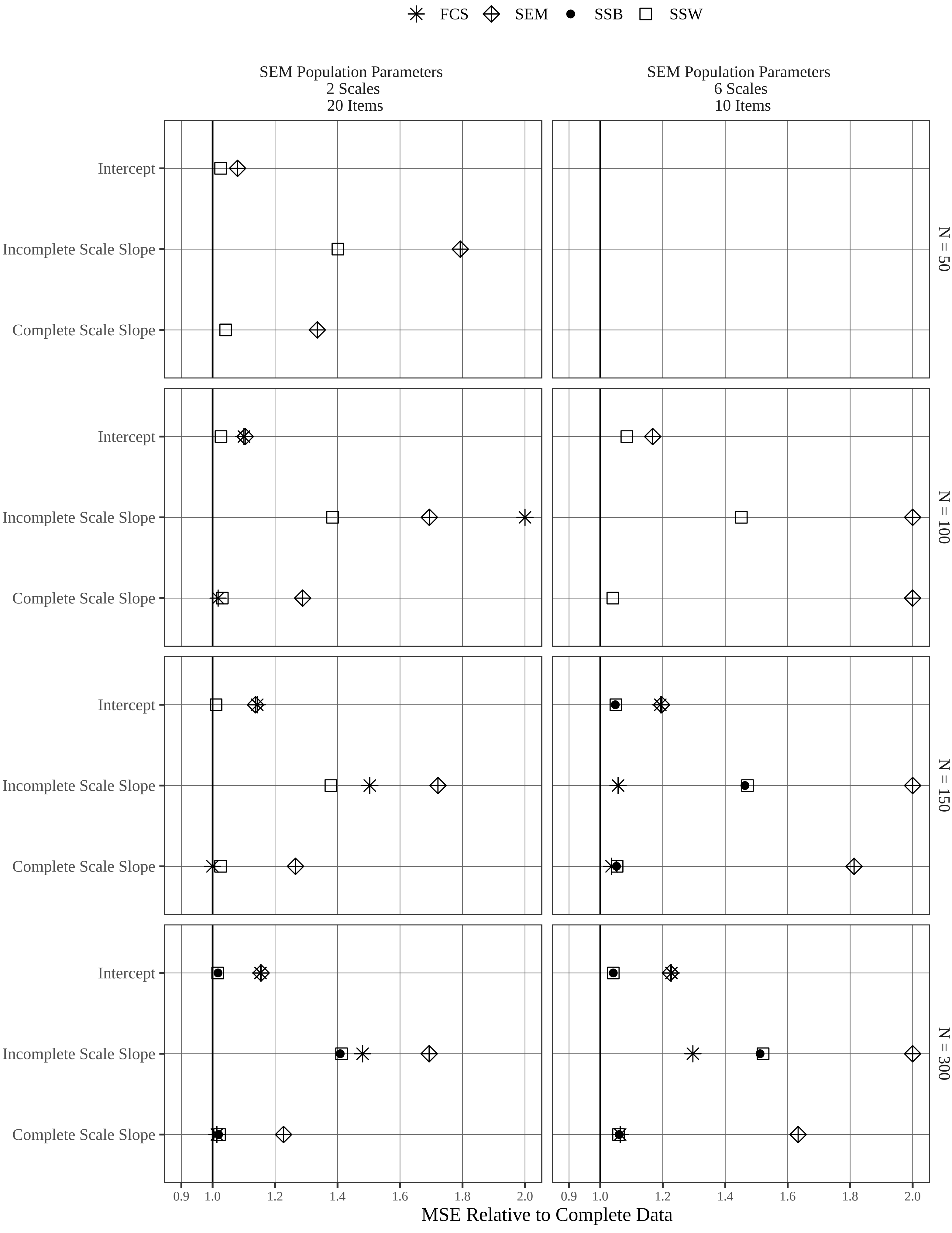
Simulation 4: Percent Bias Values With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
−20
−10
0
10
20
Percent Relative Bias

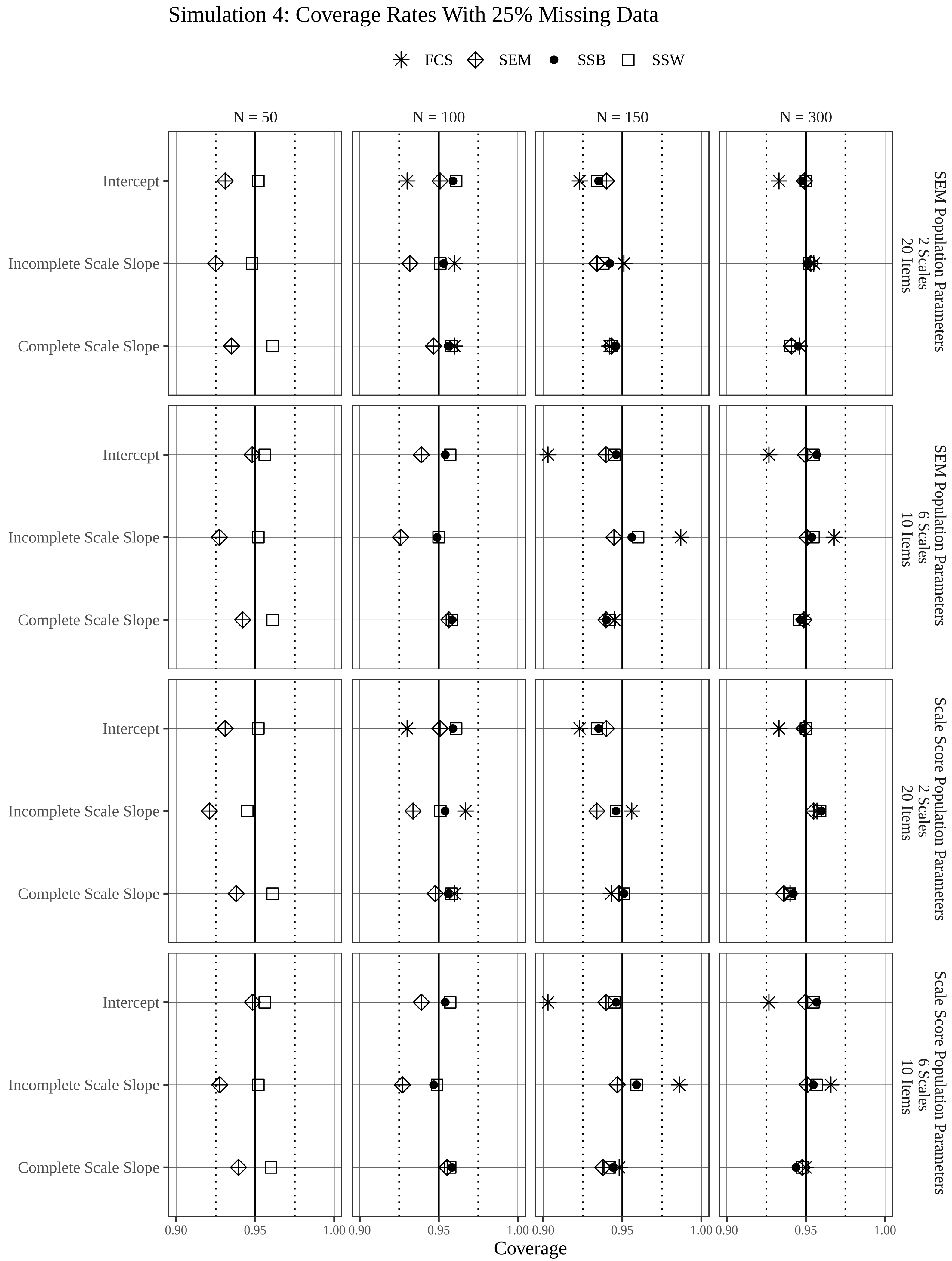

Simulation 4: Percent Bias Values With 40% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
−20
−10
0
10
20
Percent Relative Bias

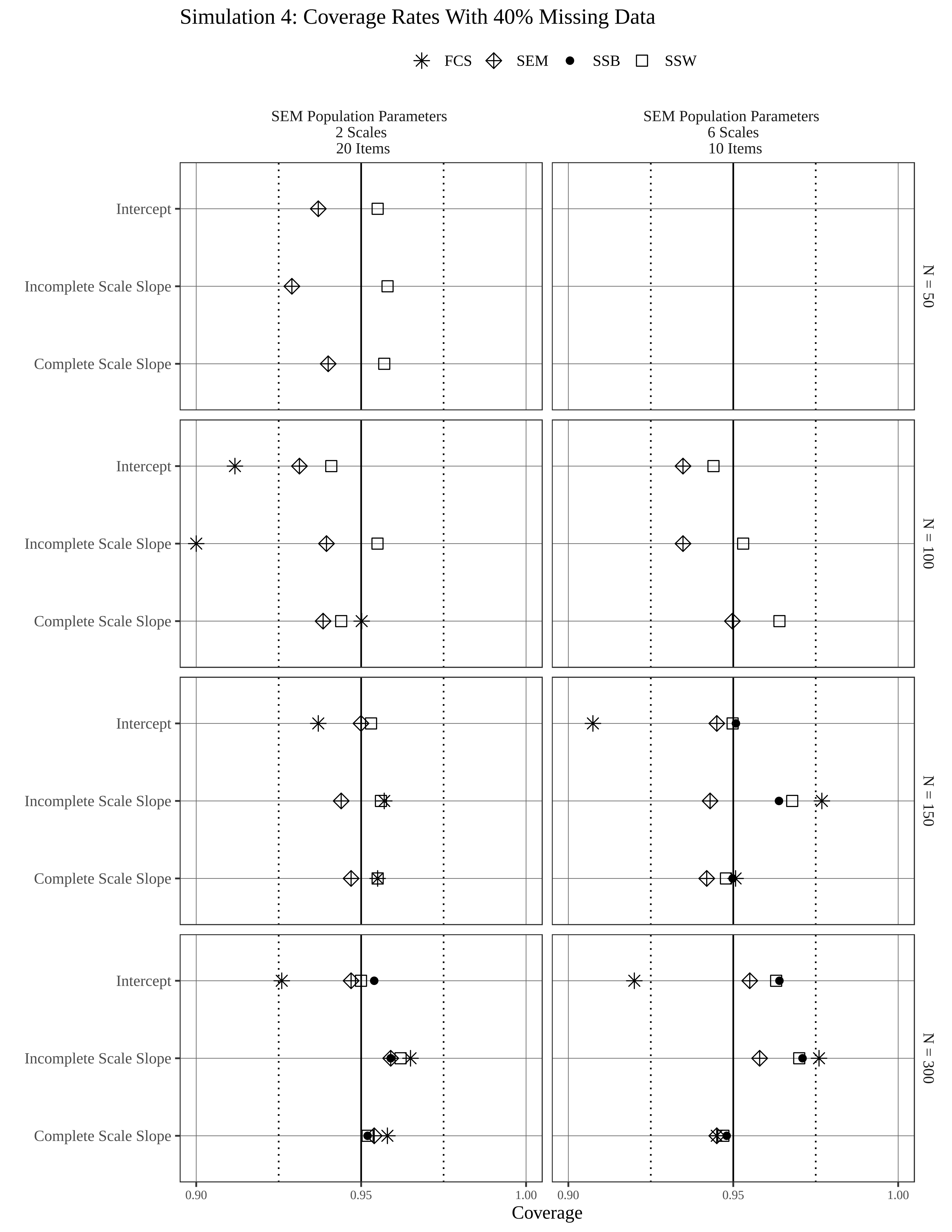


Simulation 4: MSE Ratios With 25% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
0.9
1.0
1.5
2.0
2.5
3.0
3.5
MSE Relative to Complete Data

Simulation 4: MSE Ratios With 40% Missing Data




Simulation 4: Coverage Rates With 25% Missing Data
FCS
SEM
SSB
SSW
N = 50
N = 100
N = 150
N = 300
Intercept
Incomplete Scale Slope
Complete Scale Slope
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Scale Score Population Parameters
2 Scales
20 Items
Scale Score Population Parameters
6 Scales
10 Items
0.90
0.95
1.00
Coverage



Simulation 4: Coverage Rates With 40% Missing Data
FCS
SEM
SSB
SSW
SEM Population Parameters
2 Scales
20 Items
SEM Population Parameters
6 Scales
10 Items
Intercept
Incomplete Scale Slope
Complete Scale Slope
N = 50
N = 100
N = 150
N = 300
0.90
0.95
1.00
Coverage

Supplemental Simulation Study 1

In addition to the simulations reported in the body text, we examined the performance of factored regression missing data handling using a series of computer simulation studies. The first examines the method with between-scale constraints imposed on the predictor items; the second considers both within- and between-scale constraints; the third also examines within- and between-scale constraints but does so in a context where the item-level covariance structure could be at odds with arbitrary within-scale restrictions; and the fourth simulation introduces categorical indicators. In the interest of space, we present graphical summaries that illustrate the main points, and the online supplement reports all simulation results.

The first simulation study implemented a full factorial design comprised of one within-subject factor (two missing data methods; factored regression Bayesian missing data handling versus fully conditional specification) and four between-subject factors: (a) sample size (N = 50, 100, 150, 200, 250, or 300), (b) the item-level missingness probability (P_{miss} = .10, .25, or .40), (c) number of scales (N_{scale} = two or six), and (d) number of items per scale (N_{item} = five, 10, or 20). This resulted in a total of 108 simulated conditions, and we generated 1000 artificial datasets within each between-subjects design cell.

Prior to computing the three evaluation criteria, we assessed the convergence rates of the missing data handling methods in each condition. Table S1 displays the convergence rates for all conditions. As expected, we can see that the convergence rates for the factored regression Bayesian method were better in some cases especially when sample sizes are relatively small (e.g., fully conditional specification failed to converge in cells where the combination of the number of scales and items were large and when sample size was small). Going forward, we report results only for design cells where convergence exceeded 90%.

Figures S1 through S18 give trellis plots displaying the simulation results. First, consider percent bias values for a subset of conditions with 25% missing data rate and sample sizes equal to 50, 100, and 150. Panels with no circle or asterisk symbols indicate that factored regression Bayesian estimation or fully conditional specification, respectively, had insufficient convergence rates or failed completely. Moreover, if an estimate's percent bias was greater than ±15%, then its bias value was plotted at the minimum or maximum on the horizontal axis. This only was an issue for FCS, which exhibited severe bias in some combinations with many scales and items; for example, the slope bias in the first column panel and second row panel (N = 50, two scales with 10 items each) was -32.03%. Importantly, the fully Bayesian estimates were approximately unbiased, with all relative bias values lower than ± 10%. The results in the figure are representative of other conditions, which are reported in the online supplement.

The trellis plots also display frequentist confidence interval and Bayesian credible interval coverage values; the dashed vertical lines at .925 and .975 correspond to Bradley's (1978) so-called liberal criterion. Following the bias plot, a missing black circle or asterisk indicates that a method did not reliably converge. All methods generally produced adequate coverage values, and there were no conditions in the figure where the factored regression Bayesian procedure produced coverage values outside Bradley's bounds. Interestingly, the complete-data coverage values were somewhat worse (lower) than those of Bayes estimation with missing data (e.g., conditions crossing six scales and sample sizes of $N = 50$ and 100).

Finally, a set of trellis plots display the mean-squared error (MSE) ratios that express missing data MSEs relative to the complete-data MSE. A ratio larger than unity indicates that the complete data estimates were more precise, as would be expected. A missing black circle or triangle indicates that a method did not reliably converge. For the scale with incomplete items, the FCS slope estimates were generally more precise than those of Bayesian estimation, especially as sample size decreased. This finding suggests that, in situations where item-level FCS converges, it may produce estimates that are closer, on average, to the true values. However, the absence of triangle symbols in some panels again emphasizes that Bayesian estimation converges in situations where FCS does not. The figure also shows that FCS multiple imputation estimates were, in some cases, more precise than those from a complete-data analysis. We explored this counterintuitive finding and found that, when the sample size was small, the FCS regression equations produced large R-square values that were substantially larger than those of the complete data. We interpret this as a symptom of overfitting, such that FCS is infusing too little noise variation into the imputations, thereby producing inappropriately narrow sampling distributions. Kernel density plots of the sampling distributions support this conclusion.

Table S1
Convergence Values for Conditions with Convergence Rates that are not 100% in Simulation Study 1

Condition	Convergence FB	Convergence FCS
N_{scales}= 2, N_{items}= 10, P_{miss}= .25, N = 50	99.9%	99.5%
N_{scales} = 2, N_{items} = 10, P_{miss} = .40, N = 50	57.83%	71.3%
N_{scales} = 2, N_{items} = 20, P_{miss} = .10, N = 50	100%	0%
N_{scales} = 2, N_{items} = 20, P_{miss} = .25, N = 50	8.77%	1.60%
N_{scales} = 2, N_{items} = 20, P_{miss} = .40, N = 50	0%	0%
N_{scales} = 2, N_{items} = 20, P_{miss} = .40, N = 100	36.82%	84.76%
N_{scales} = 2, N_{items} = 20, P_{miss} = .40, N = 150	99.9%	100%
N_{scales} = 6, N_{items} = 5, P_{miss} = .10, N = 50	100%	0%
N_{scales} = 6 , N_{items} = 5, P_{miss} = .25, N = 50	100%	0%
N_{scales} = 6 , N_{items} = 5, P_{miss} = .40, N = 50	90.31%	14.93%
N_{scales} = 6 , N_{items} = 10, P_{miss} = .10, N = 50	100%	0%
N_{scales} = 6 , N_{items} = 10, P_{miss} = .10, N = 100	100%	0%
N_{scales} = 6 , N_{items} = 10, P_{miss} = .25, N = 50	92.38%	0%
N_{scales} = 6 , N_{items} = 10, P_{miss} = .25, N = 100	100%	0%
N_{scales} = 6 , N_{items} = 10, P_{miss} = .40, N = 50	0.138%	0%
N_{scales} = 6 , N_{items} = 10, P_{miss} = .40, N = 100	99.7%	0%
N_{scales} = 6 , N_{items} = 10, P_{miss} = .40, N = 150	100%	96.17%
N_{scales} = 6 , N_{items} = 20, P_{miss} = .10, N = 50	93.06%	0%
N_{scales} = 6 , N_{items} = 20, P_{miss} = .10, N = 100	100%	0%
N_{scales} = 6 , N_{items} = 20, P_{miss} = .10, N = 150	100%	0%
N_{scales} = 6 , N_{items} = 20, P_{miss} = .10, N = 200	100%	0%
N_{scales} = 6 , N_{items} = 20, P_{miss} = .10, N = 250	100%	0%
N_{scales} = 6 , N_{items} = 20, P_{miss} = .10, N = 300	100%	0%
N_{scales} = 6 , N_{items} = 20, P_{miss} = .25, N = 50	0%	0%

$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .25$, $N = 100$	99.79%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .25$, $N = 150$	100%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .25$, $N = 200$	100%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .25$, $N = 250$	100%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .25$, $N = 300$	100%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .40$, $N = 50$	0%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .40$, $N = 100$	0%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .40$, $N = 150$	94.59%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .40$, $N = 200$	100%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .40$, $N = 250$	100%	0%
$N_{scales} = 6$, $N_{items} = 20$, $P_{miss} = .40$, $N = 300$	100%	0%

Figure S1

Percent Bias Values for Conditions Including Two Scales and Five Items for Simulation 1

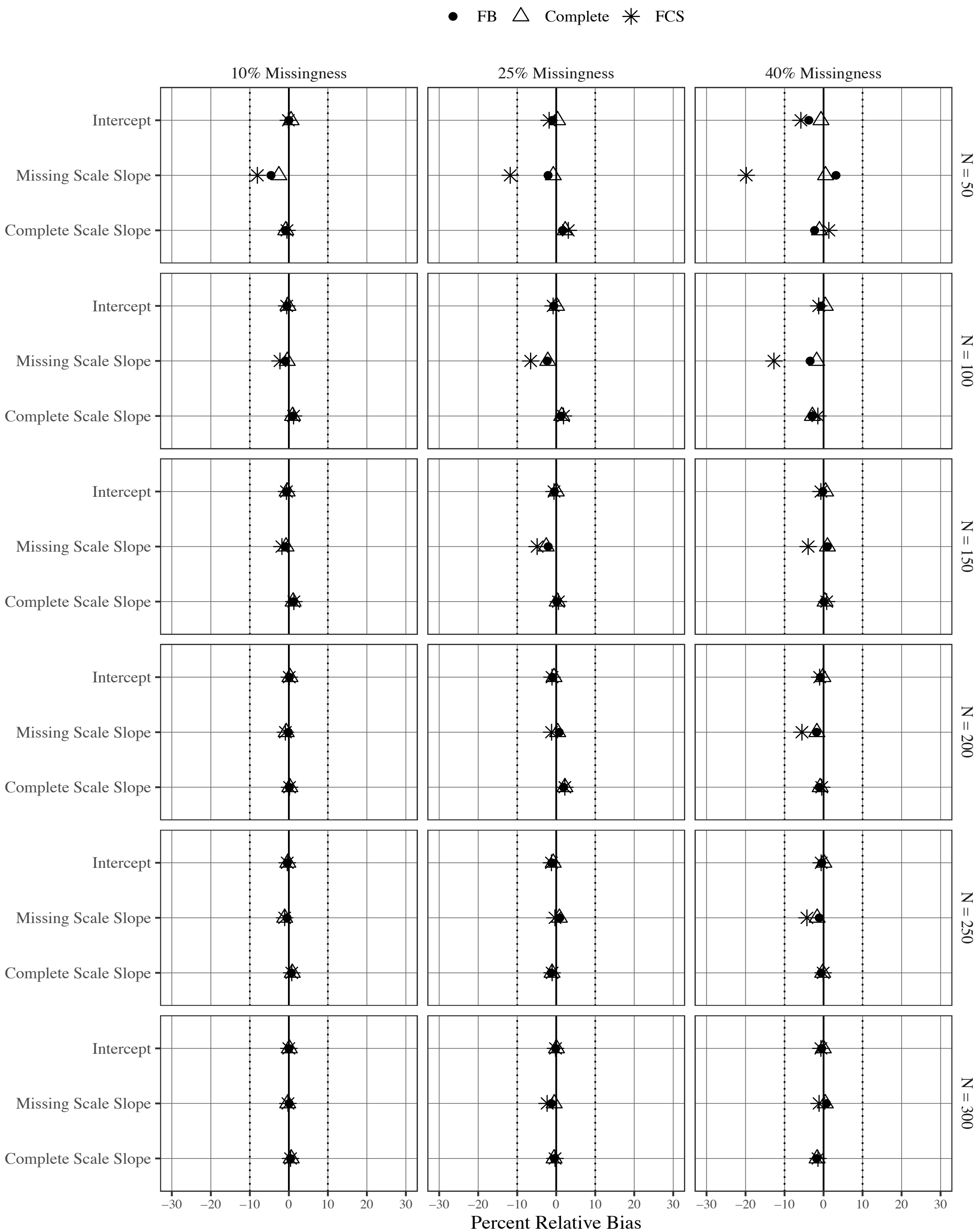


Figure S2

Percent Bias Values for Conditions Including Two Scales and 10 Items for Simulation 1

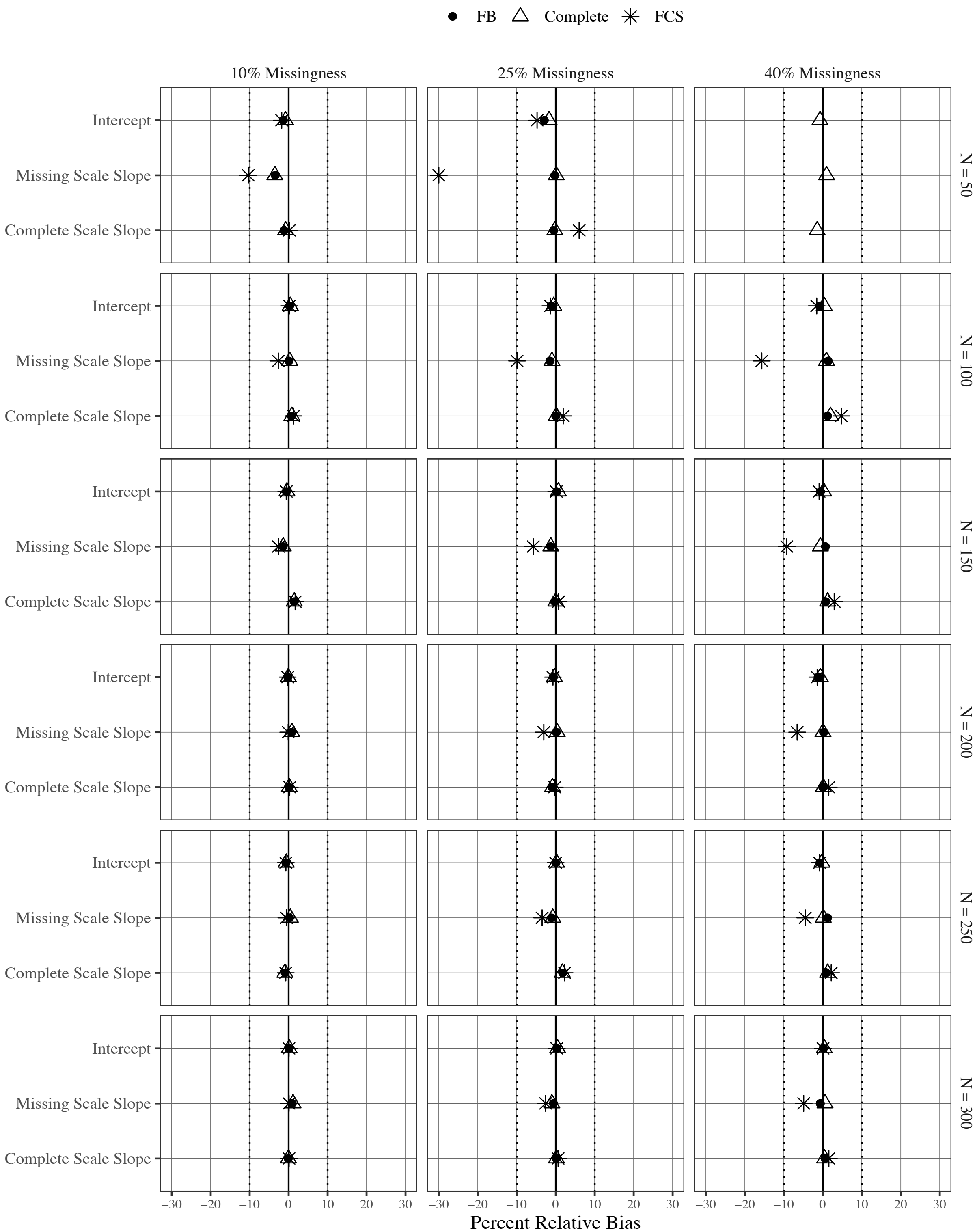


Figure S3

Percent Bias Values for Conditions Including Two Scales and 20 Items for Simulation 1

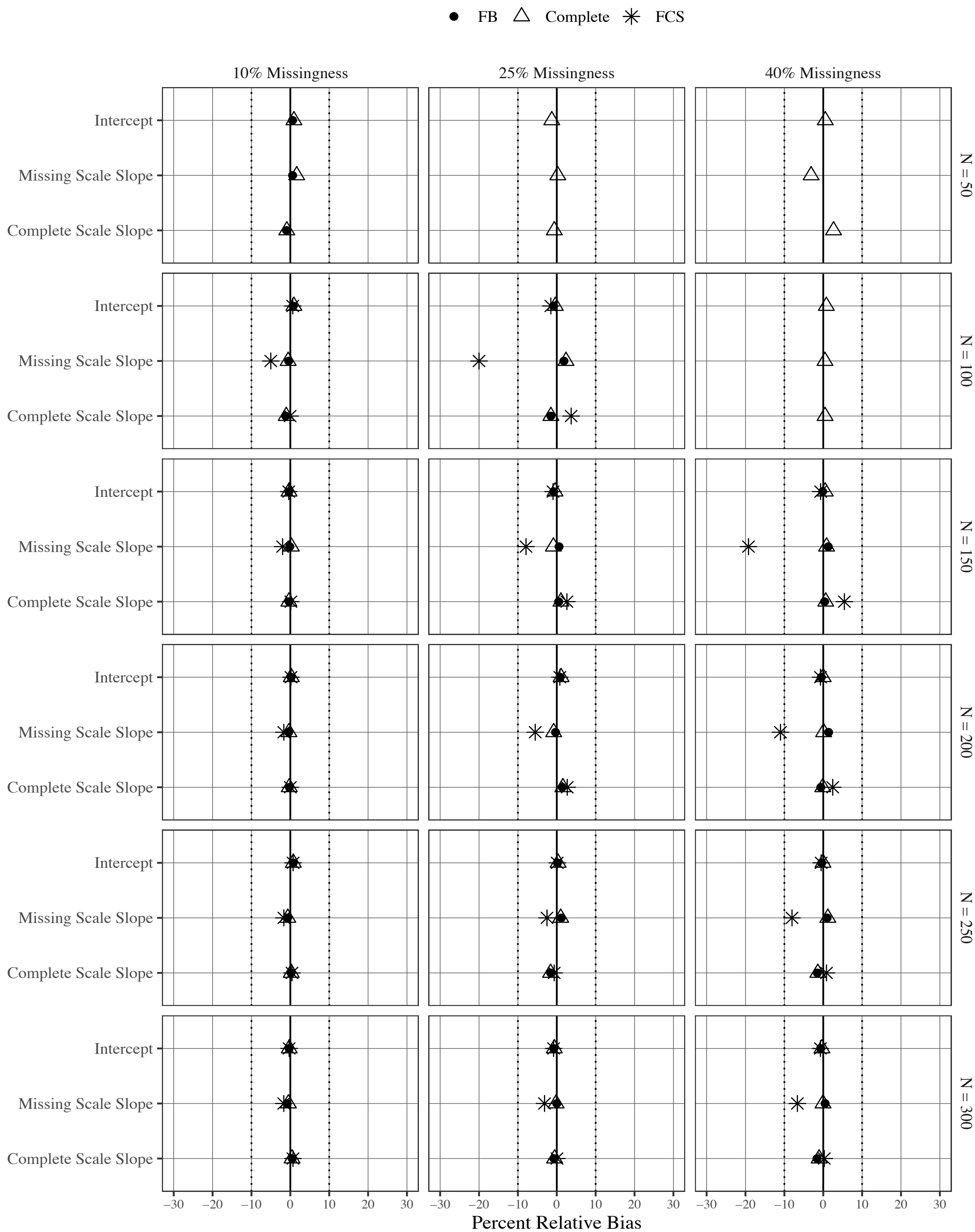


Figure S4

Percent Bias Values for Conditions Including Six Scales and Five Items for Simulation 1

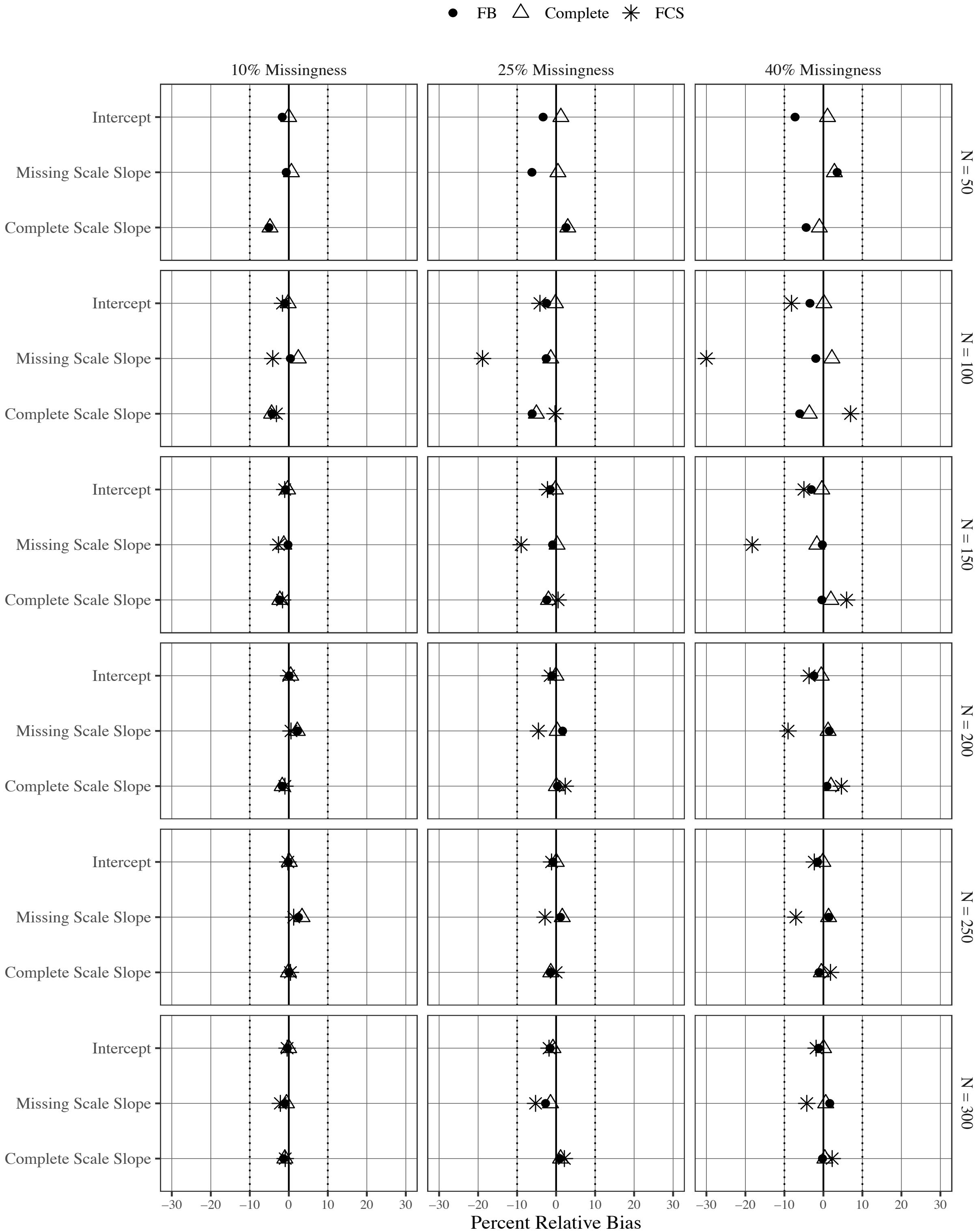


Figure S5

Percent Bias Values for Conditions Including Six Scales and 10 Items for Simulation 1

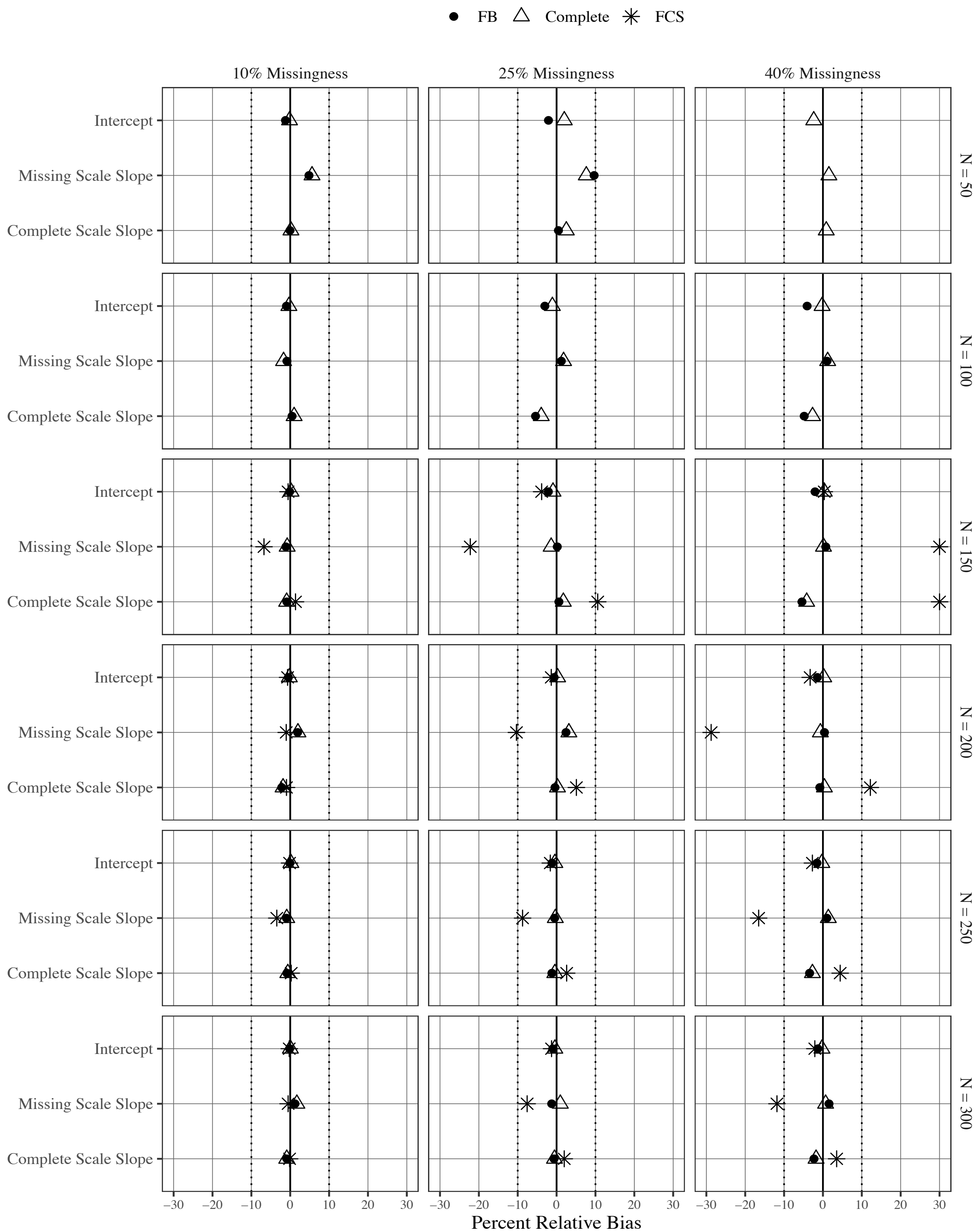


Figure S6

Percent Bias Values for Conditions Including Six Scales and 20 Items for Simulation 1

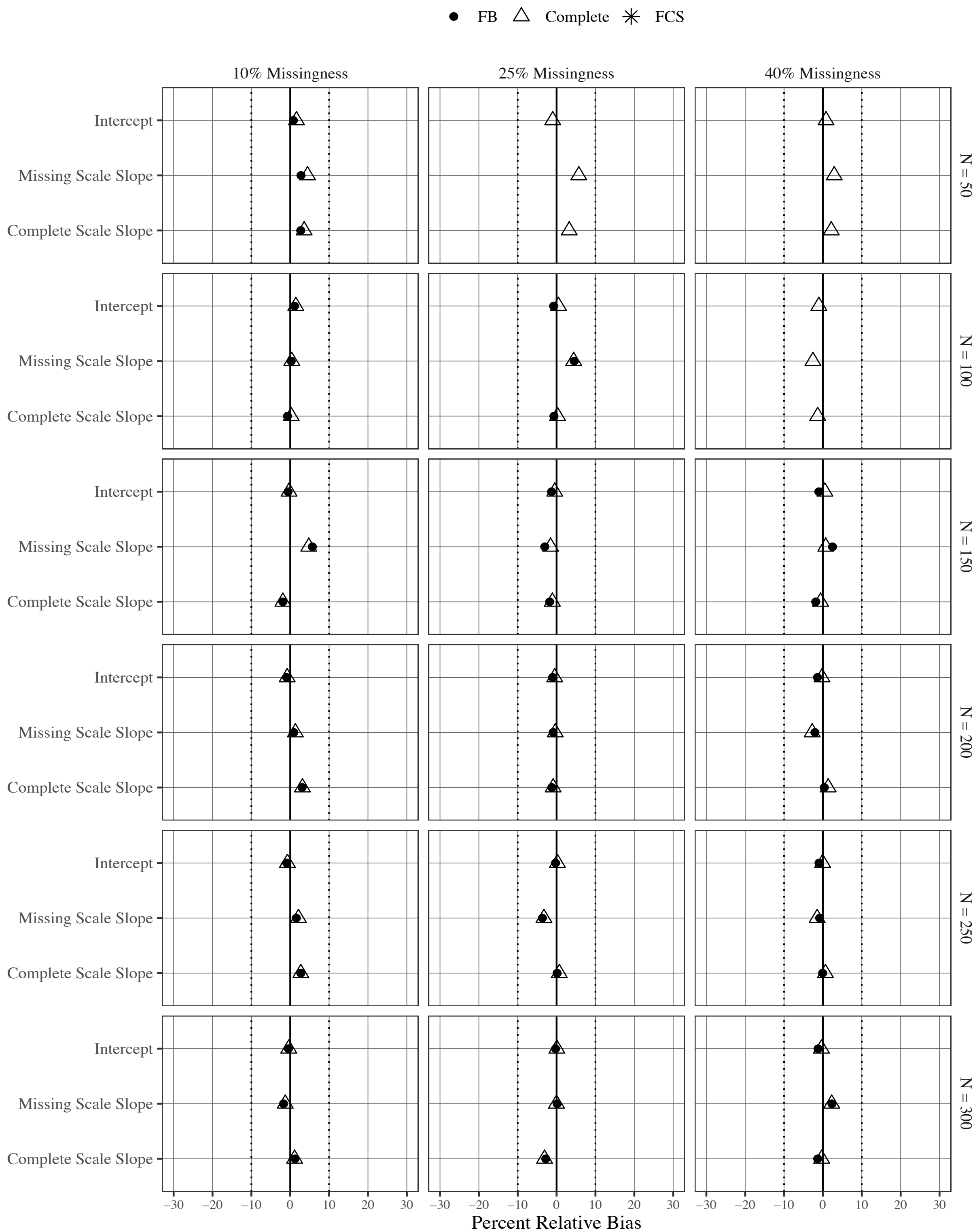


Figure S7

Coverage Rates for Conditions Including Two Scales and Five Items for Simulation 1

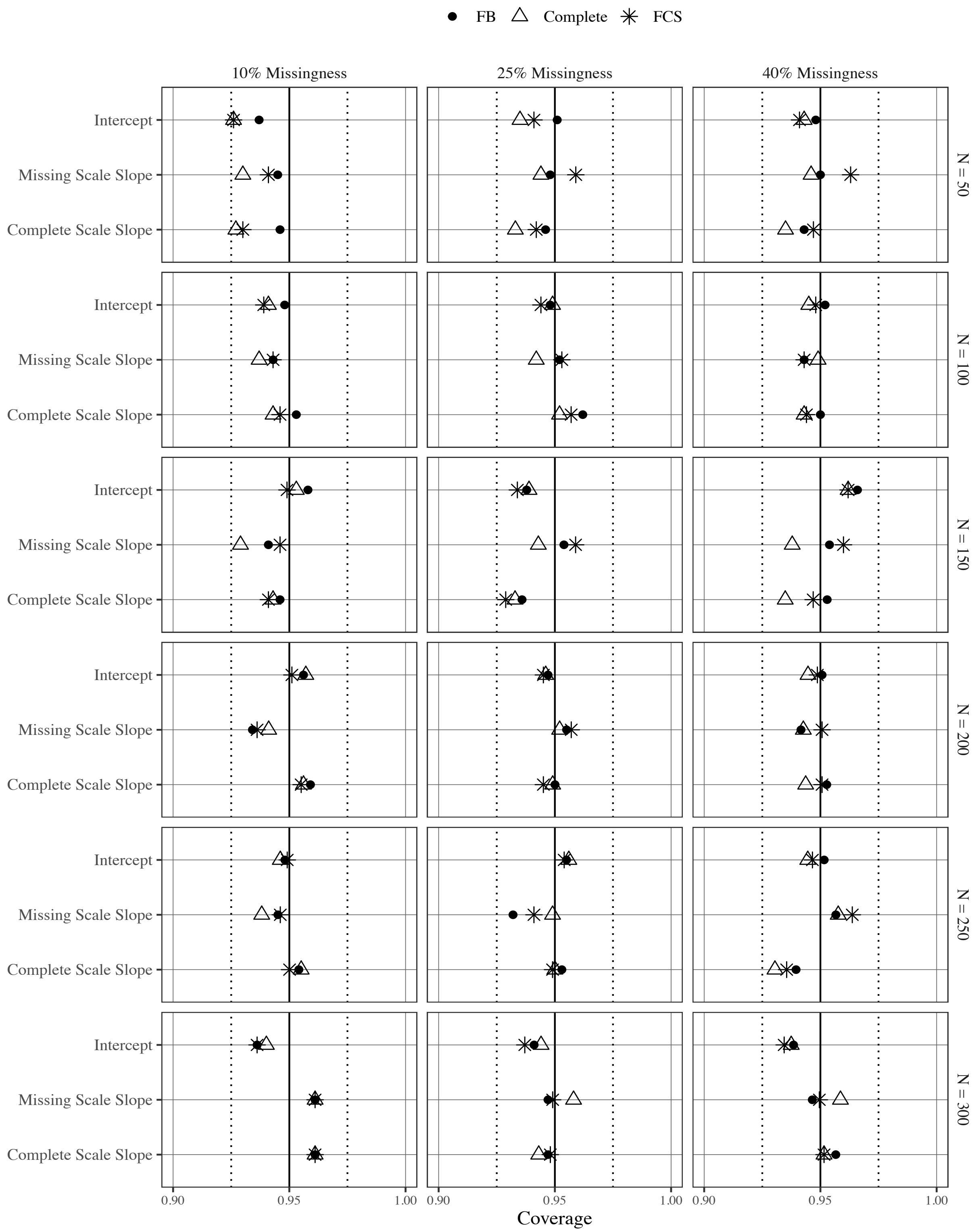


Figure S8

Coverage Rates for Conditions Including Two Scales and 10 Items for Simulation 1

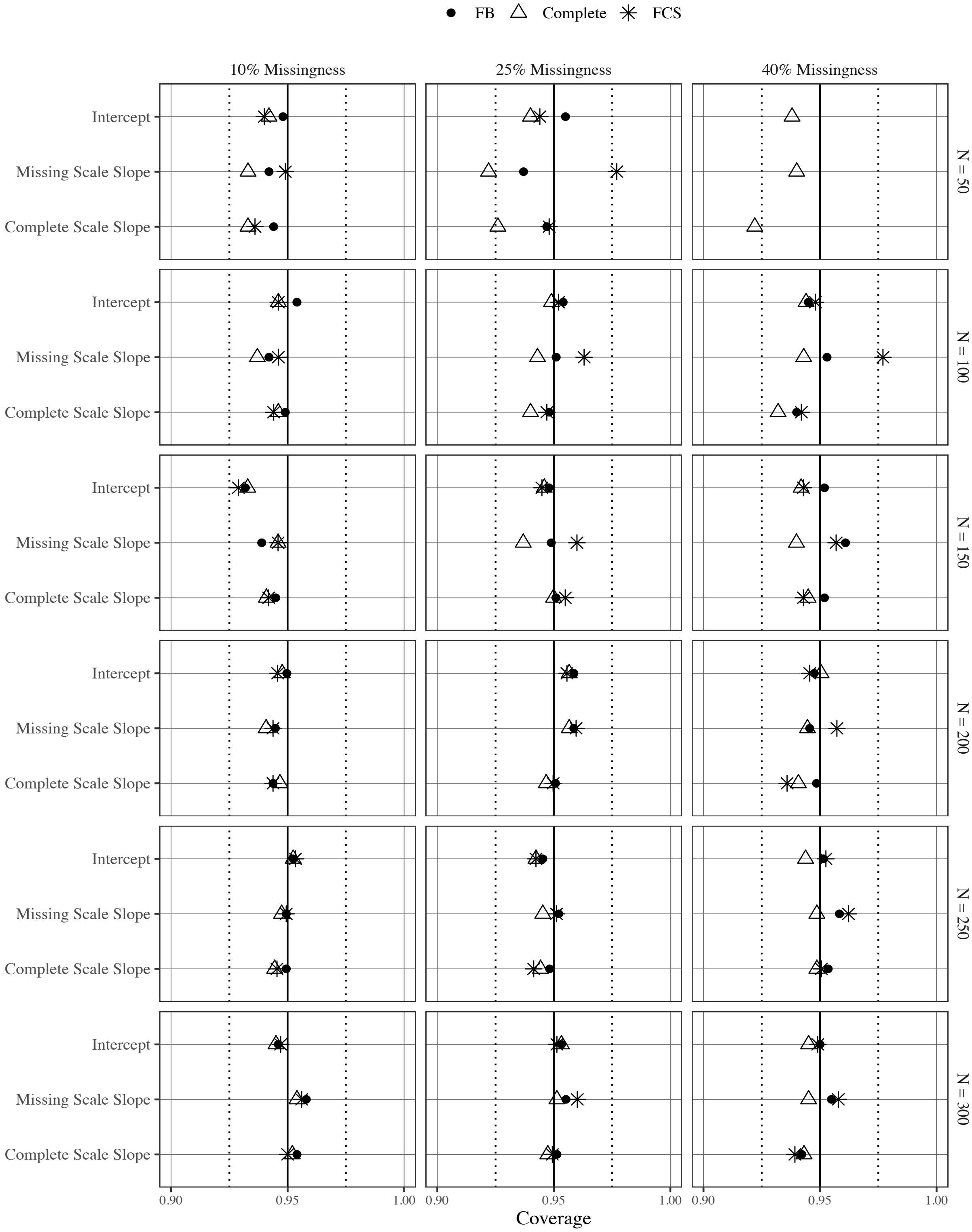


Figure S9

Coverage Rates for Conditions Including Two Scales and 20 Items for Simulation 1

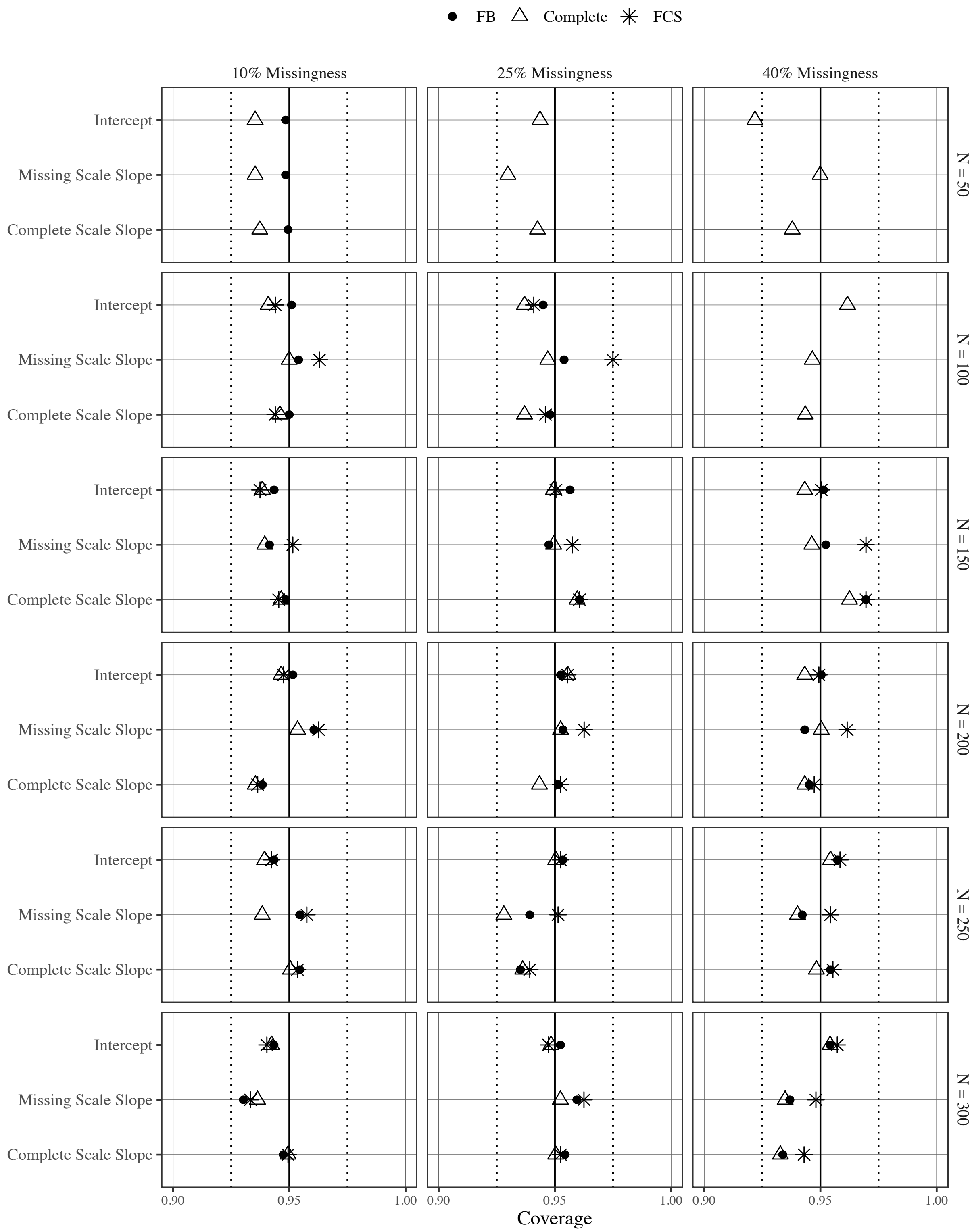


Figure S10

Coverage Rates for Conditions Including Six Scales and Five Items for Simulation 1

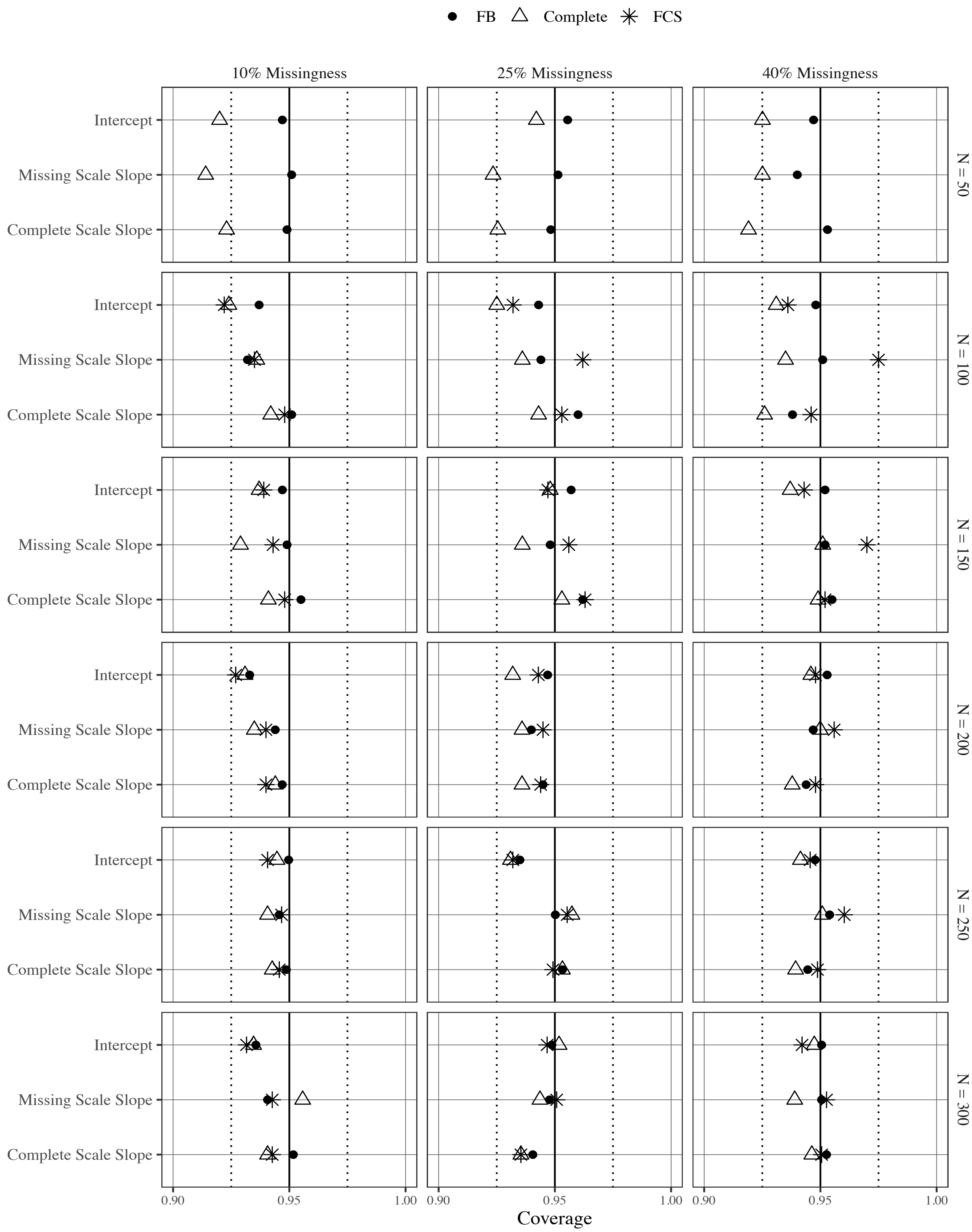


Figure S11

Coverage Rates for Conditions Including Six Scales and 10 Items for Simulation 1

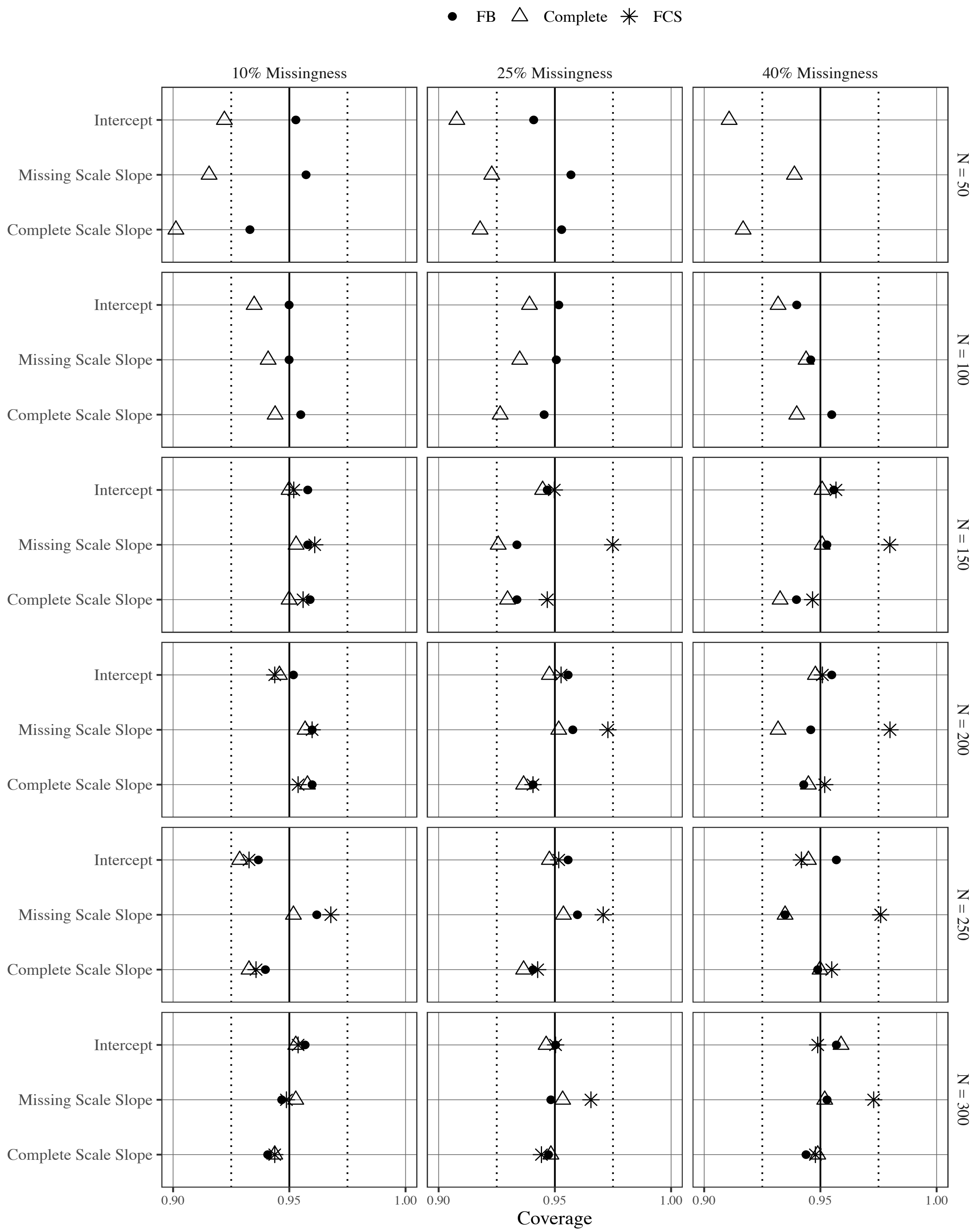


Figure S12

Coverage Rates for Conditions Including Six Scales and 20 Items for Simulation 1

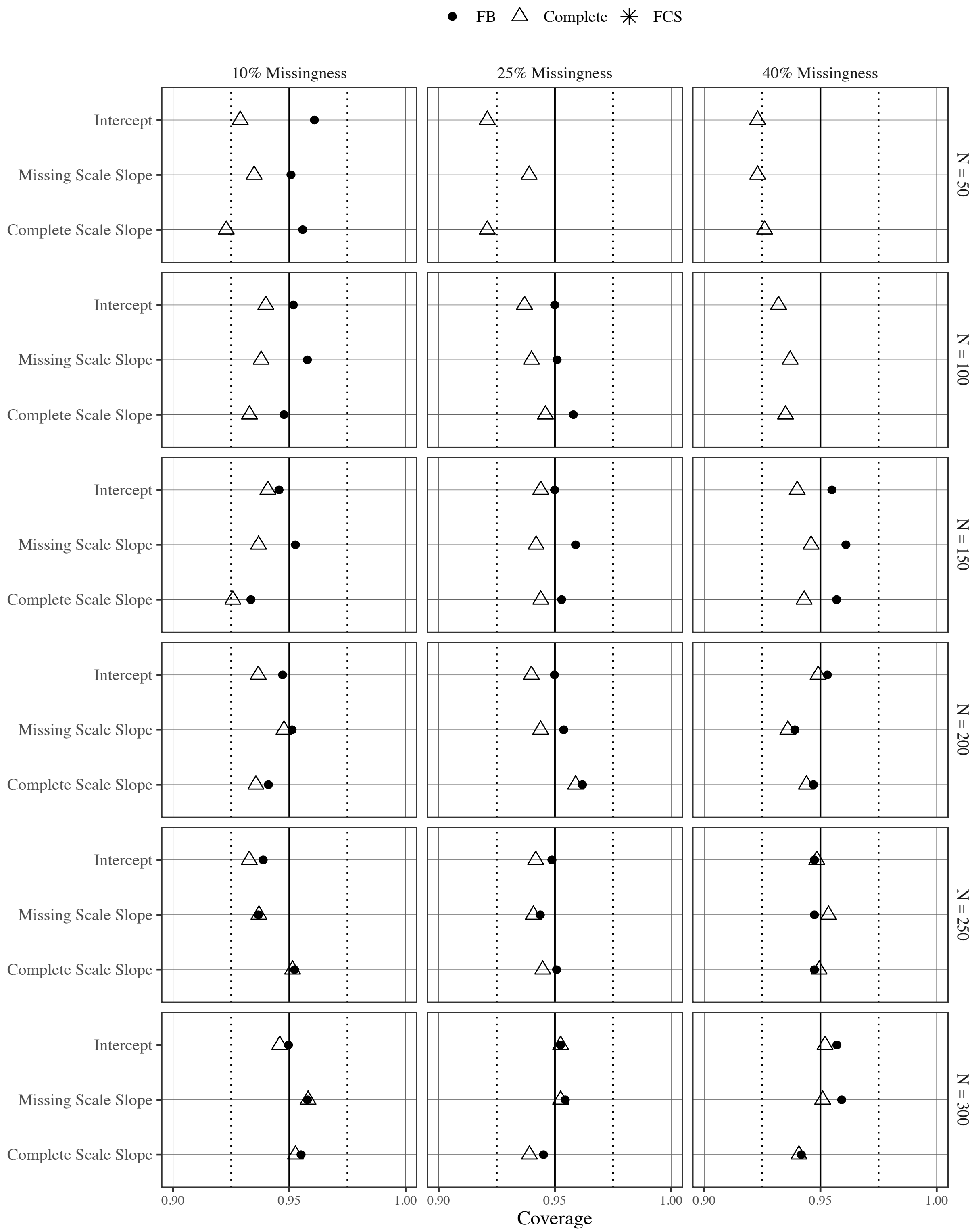


Figure S13

MSE Ratios for Conditions Including Two Scales and Five Items for Simulation 1

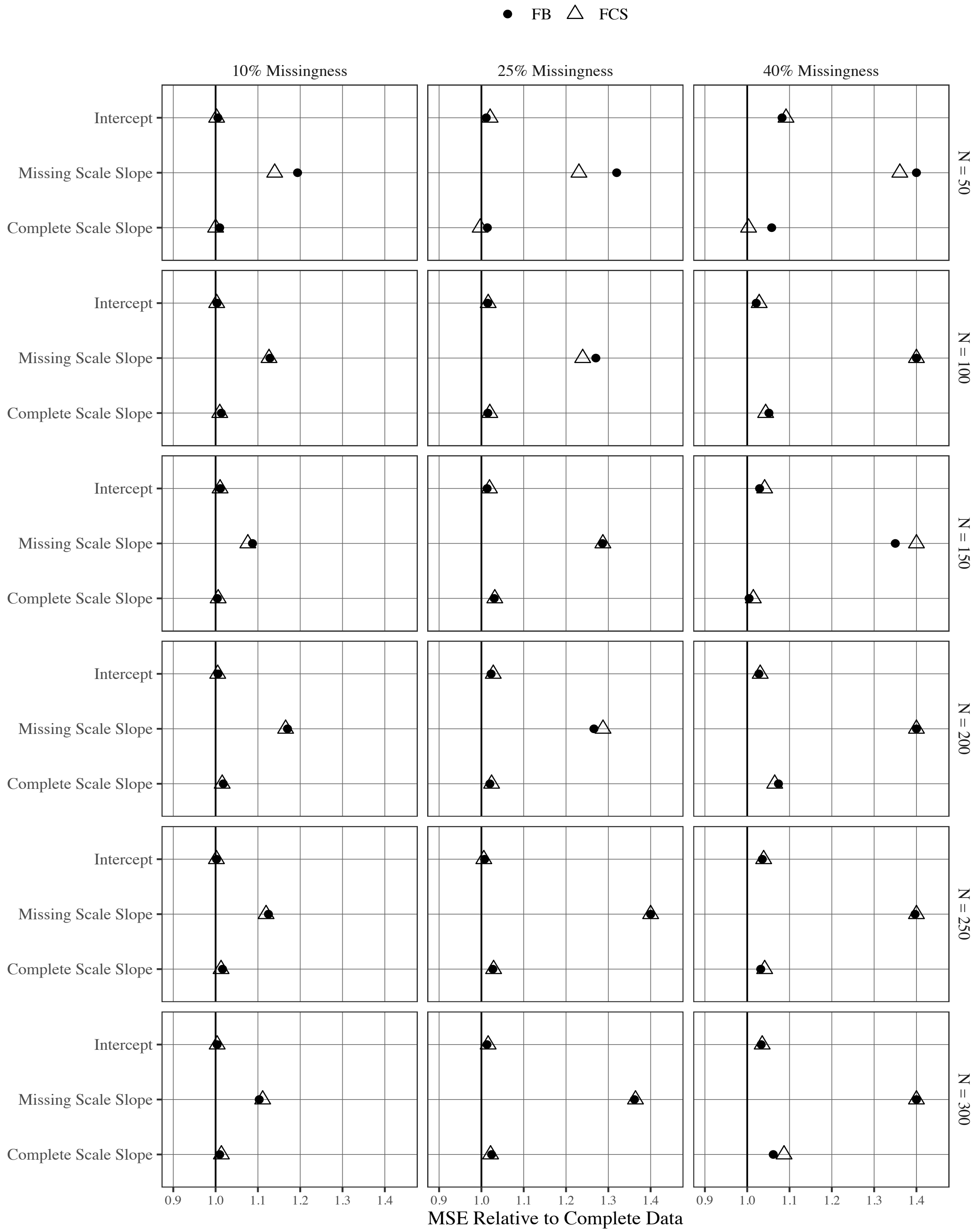


Figure S14

MSE Ratios for Conditions Including Two Scales and 10 Items for Simulation 1

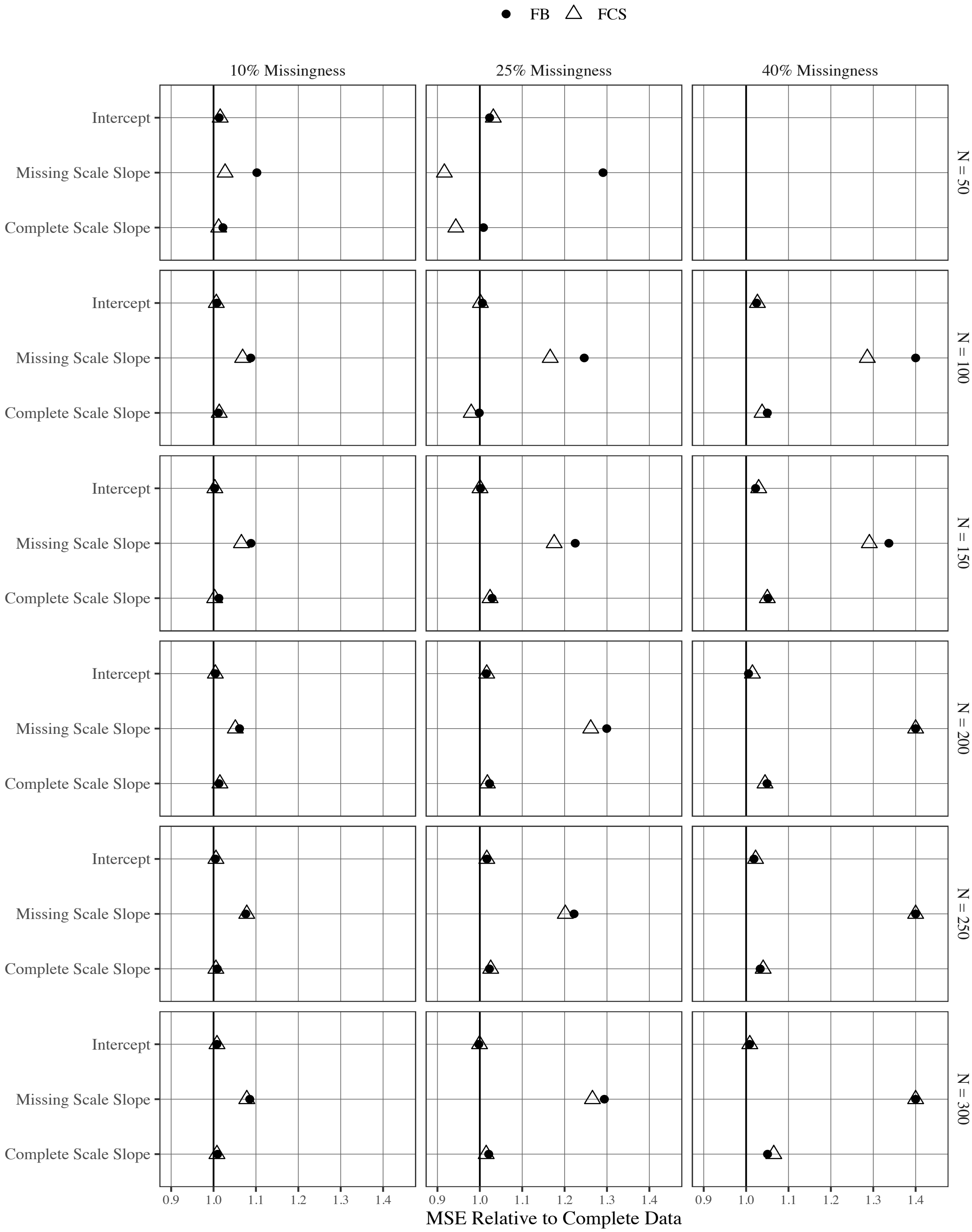


Figure S15

MSE Ratios for Conditions Including Two Scales and 20 Items for Simulation 1

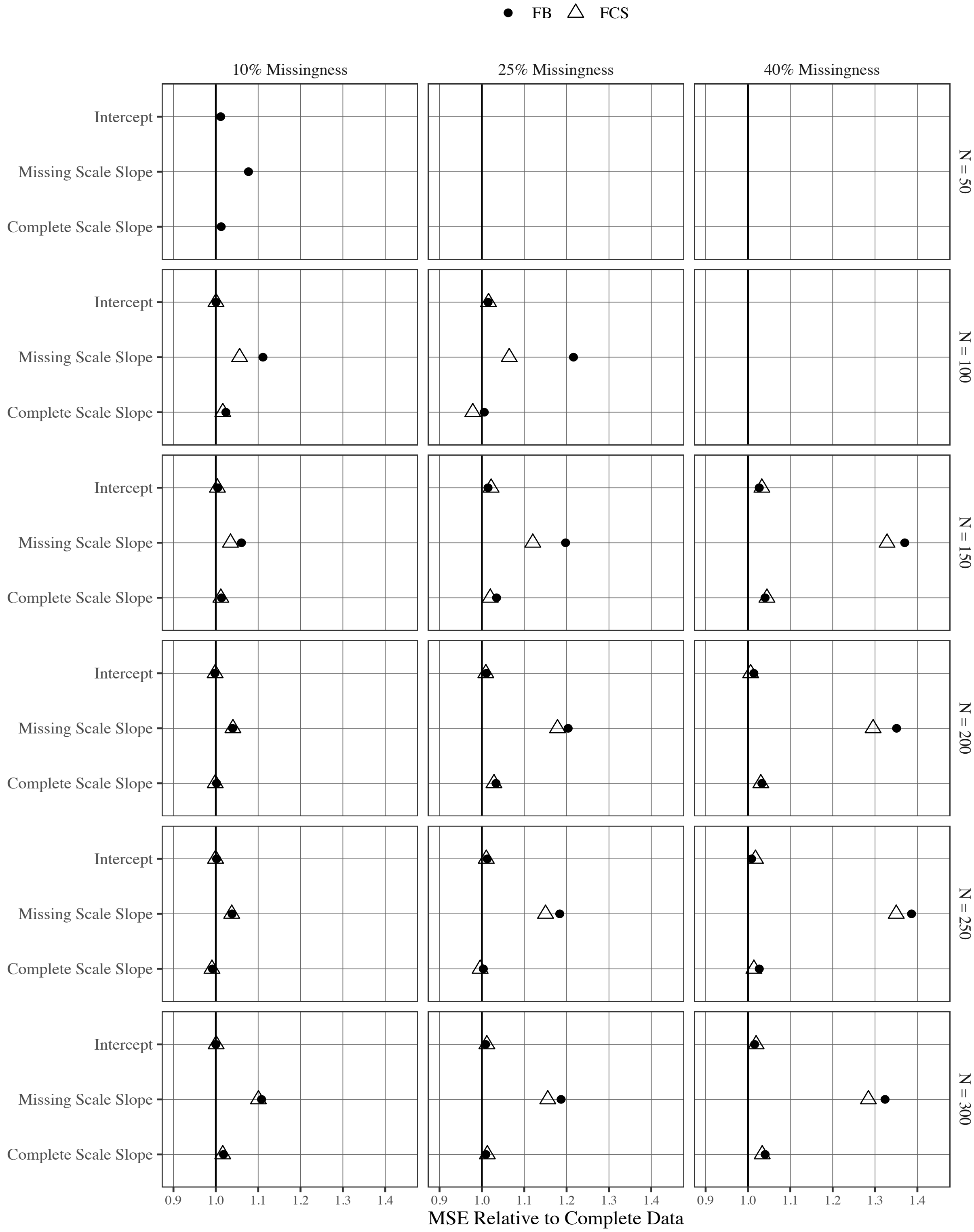


Figure S16

MSE Ratios for Conditions Including Six Scales and Five Items for Simulation 1

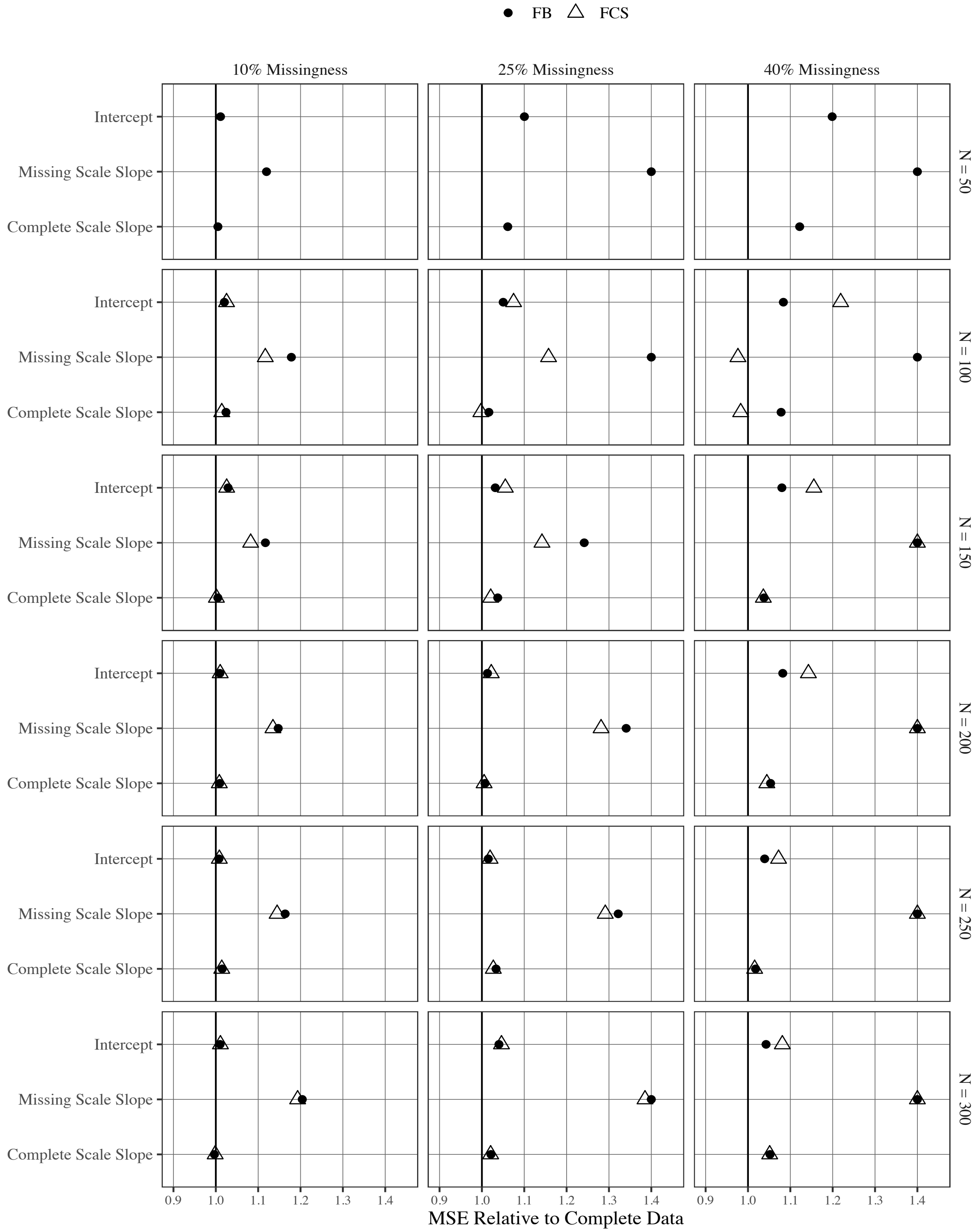


Figure S17

MSE Ratios for Conditions Including Six Scales and 10 Items for Simulation 1

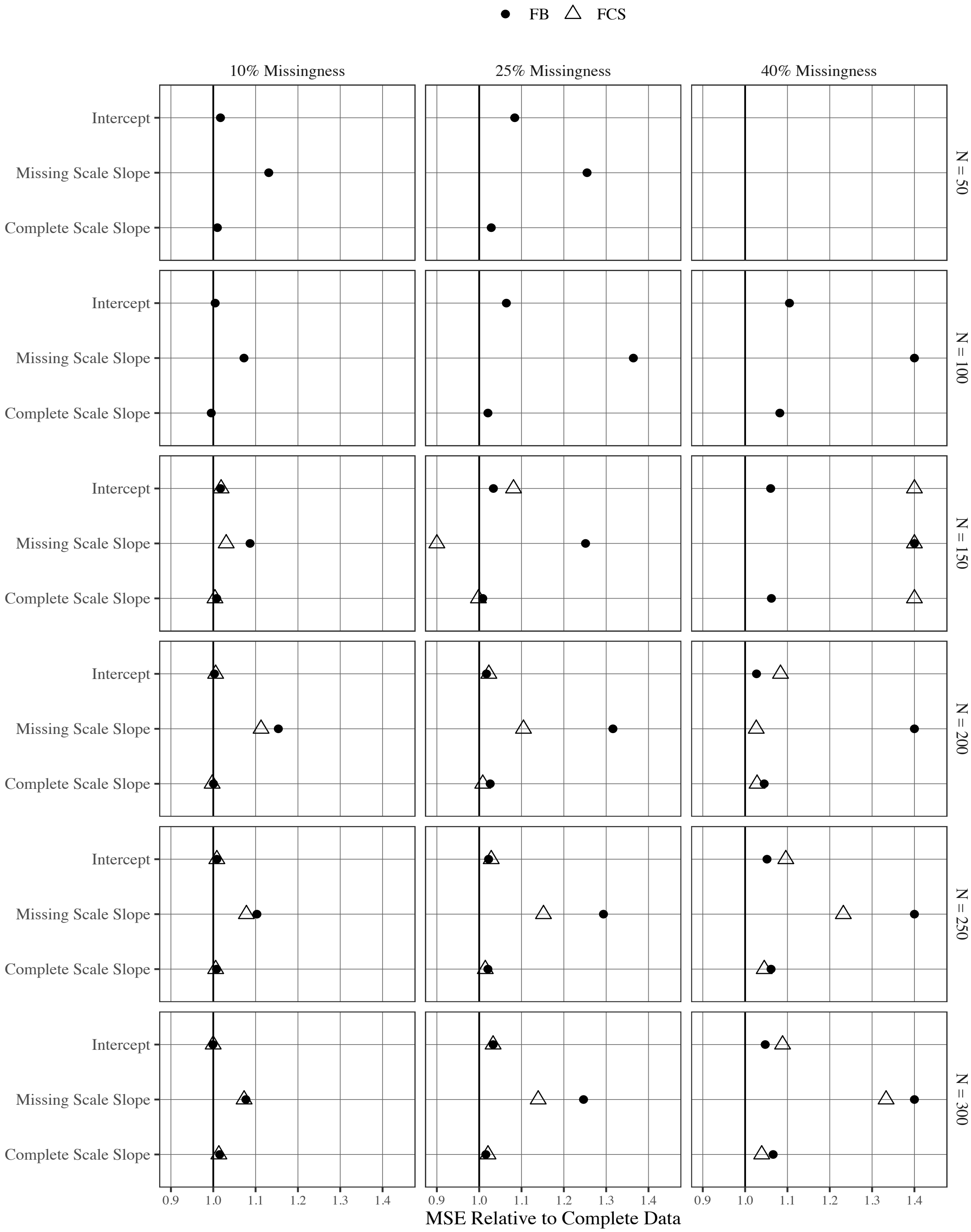
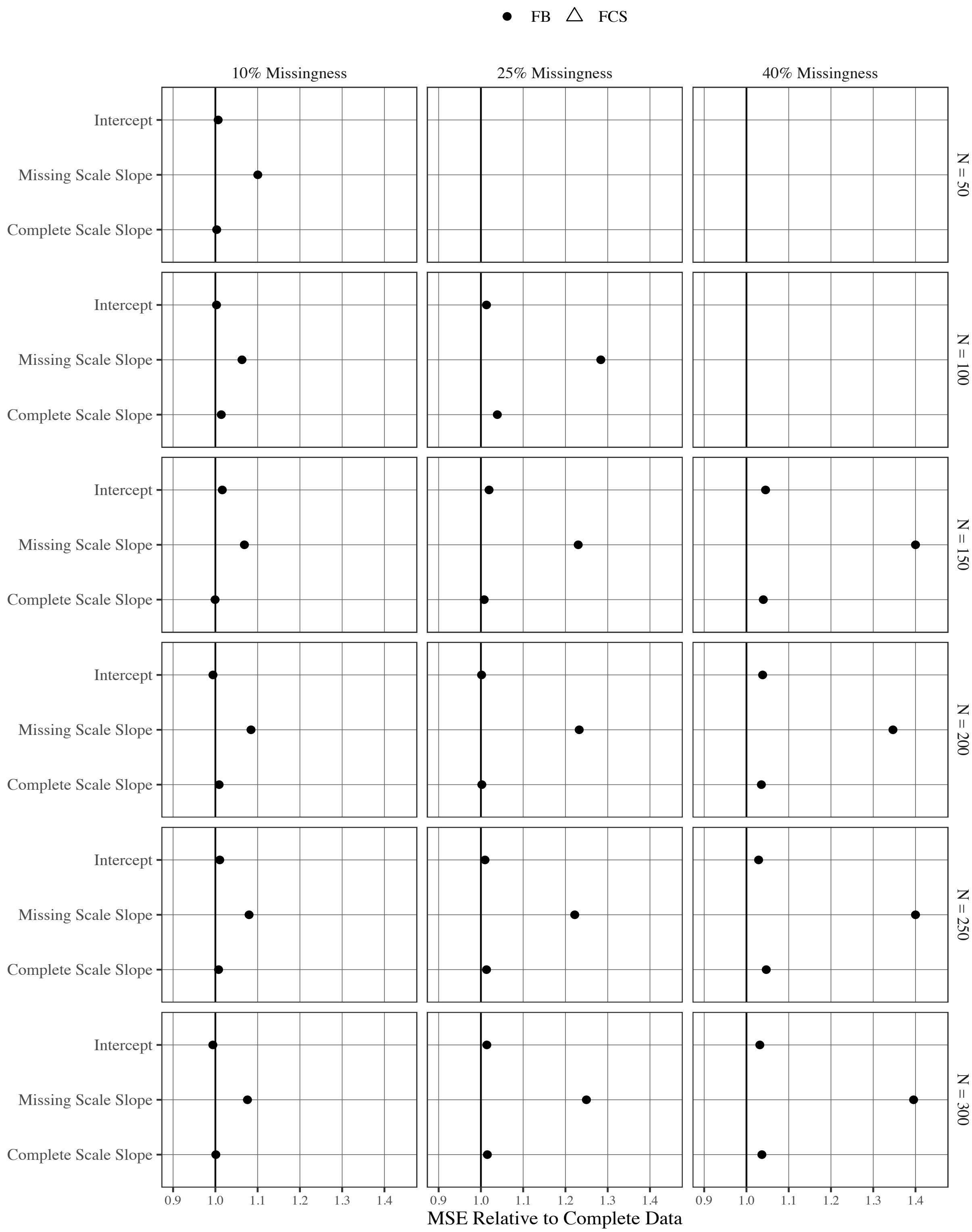


Figure S18

MSE Ratios for Conditions Including Six Scales and 20 Items for Simulation 1



Supplemental Simulation Study 2: Within-Scale Constraints

Even though the between-scale constraints we imposed in the first simulation study improved convergence relative to FCS, there were still conditions where Bayesian estimation failed or had a poor convergence rate (i.e., conditions with very small sample sizes such as $N = 50$). As discussed previously, the number of parameters can further be reduced by imposing within-scale constraints on coefficients linking items from the same scale. Recall that within-scale constraints effectively assume that all items within a scale exert the same influence on a particular item (i.e., items x_2 through x_5 have the same influence on x_1).

As our goal is to explore the lower sample size limit for Bayesian estimation, we examined a subset of conditions from the first simulation study in which convergence was unacceptably low or at zero. For illustration we focus on conditions that included sample size of $N = 50$, two or six scales with 10 or 20 items each, and a 40% missing data rate. Recall that within-scale constraints effectively assume that, for any given scale item, all other items within a scale exert the same influence (e.g., items x_2 through x_5 have the same influence on x_1). To explore these constraints under optimal conditions, we generated data where all intra-scale correlations were held constant at .30 (see the Population Model and Data Generation section for the first simulation).

We can describe the results from Simulation 2 very succinctly. A full graphical summary that contains all eight conditions for this simulation in Figures S19 through S21. Consistent with the previous simulation study, factored regression Bayesian estimates generally tracked closely to the complete data and were largely unbiased, even with 40% missing data. These results are especially encouraging in light of the fact that the number of items exceeded the total sample size in half of the design cells (conditions with 60 and 120 items). The coverage results were somewhat unremarkable and were all within Bradley's bounds. Finally, MSE ratios were not relevant for this simulation because we focused on conditions where FCS multiple imputation was impossible.

Table S3
Convergence Values for all Conditions in Simulation 2

Condition	Convergence FB	Convergence FCS
N_{scales}= 2, N_{items} = 10, P_{miss}= .40, N = 50	100%	66.7%
N_{scales} = 2, N_{items} = 20, P_{miss} = .25, N = 50	100%	0%
N_{scales} = 2, N_{items} = 20, P_{miss} = .40, N = 50	100%	0%
N_{scales} = 2, N_{items} = 20, P_{miss} = .40, N = 100	100%	83.65%
N_{scales} = 6, N_{items} = 10, P_{miss} = .40, N = 50	98.3%	0%
N_{scales} = 6, N_{items} = 20, P_{miss} = .25, N = 50	100%	0%
N_{scales} = 6, N_{items} = 20, P_{miss} = .40, N = 50	98.3%	0%
N_{scales} = 6, N_{items} = 20, P_{miss} = .40, N = 100	99.4%	0%

Figure S19

Percent Bias Values for All Conditions in Simulation 2

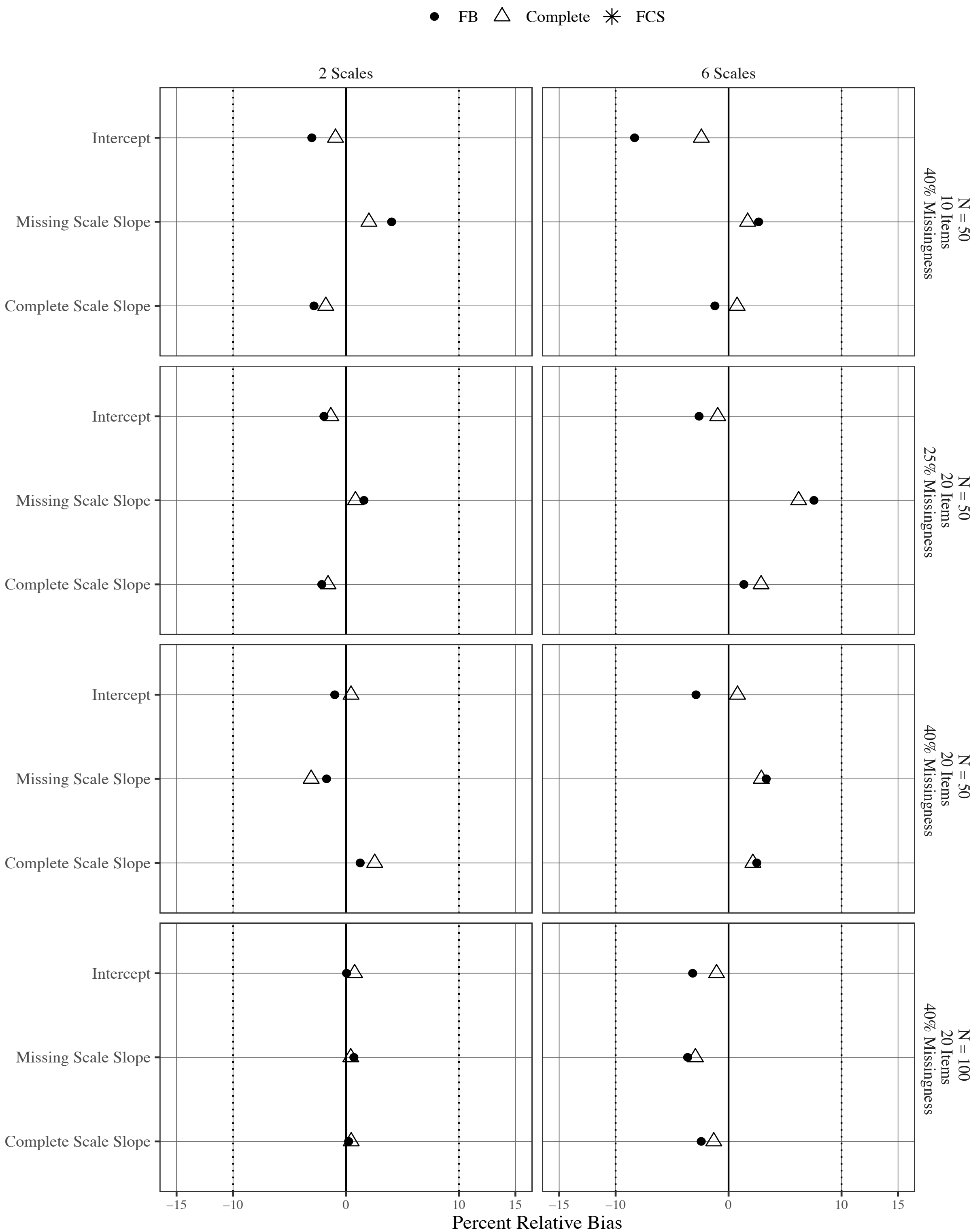


Figure S20
Coverage Rates for All Conditions in Simulation 2

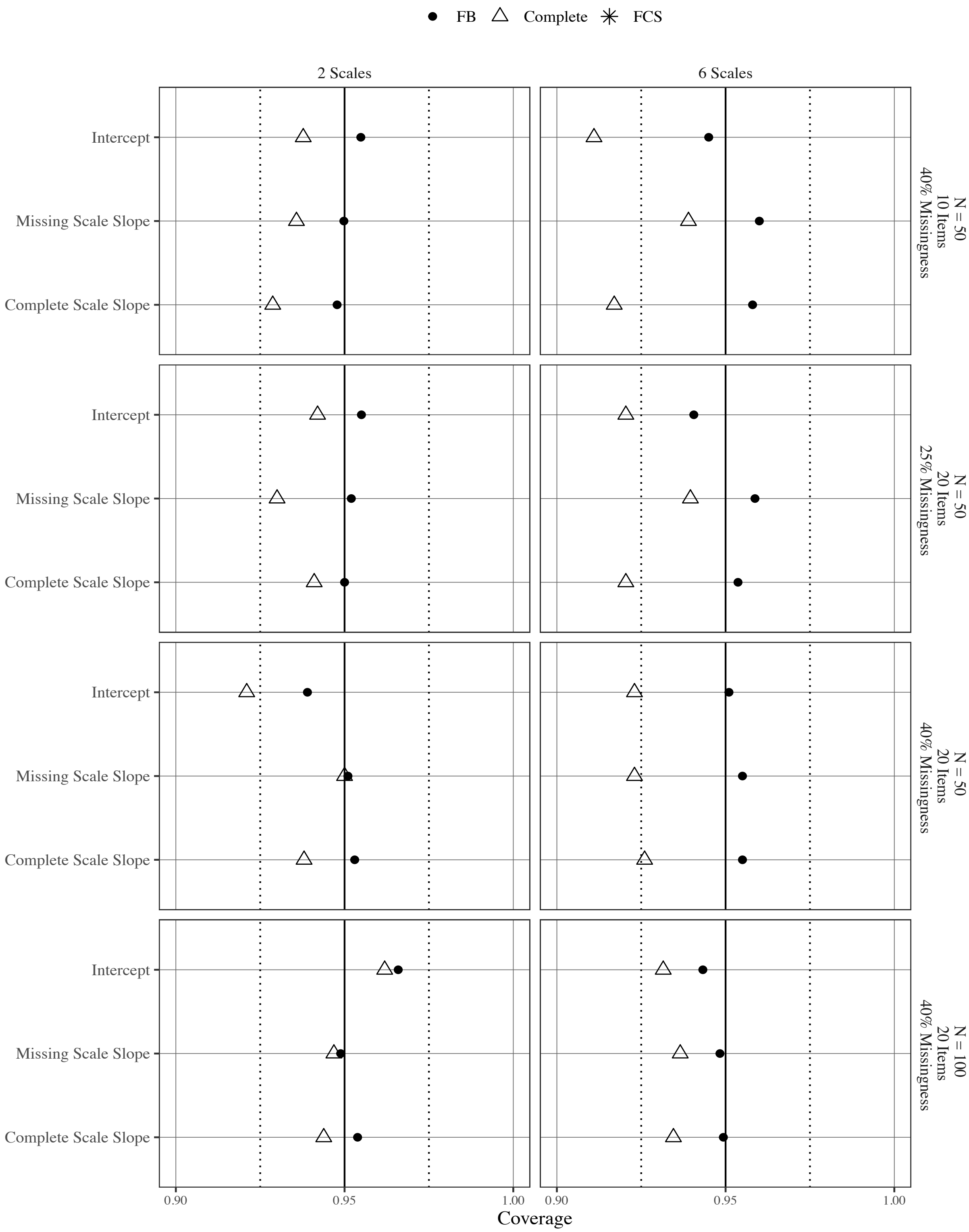
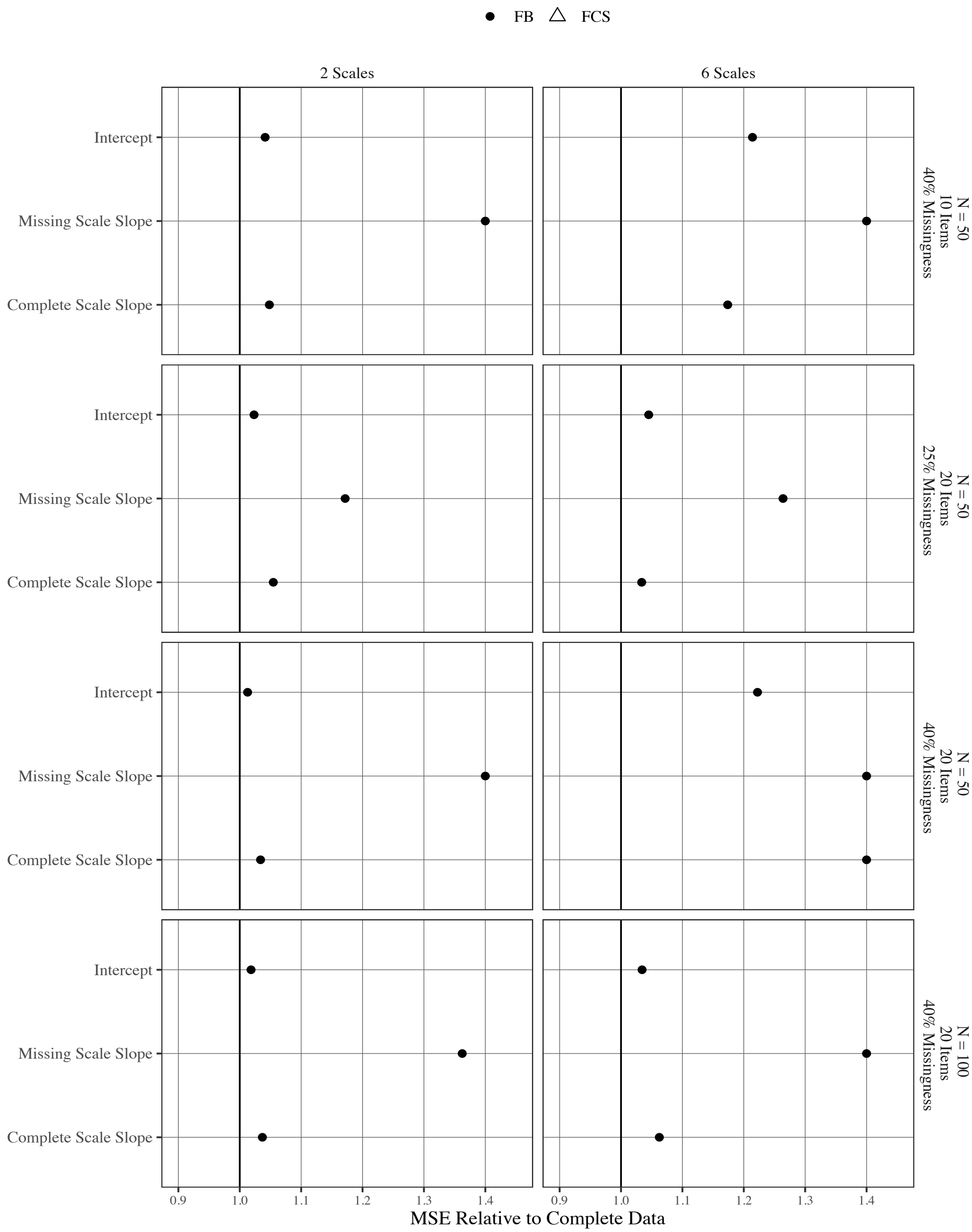


Figure S21

MSE Ratios for All Conditions in Simulation 2



Supplemental Simulation Study 3: Within-Scale Constraints With Heterogeneous Intra-Scale Correlations

In the previous section we demonstrated how implementing within-scale constraints can be especially useful when sample sizes are very small and can even accommodate scenarios where the number of items exceeds the total sample size—a situation where FCS multiple imputation is generally impossible. As explained previously, these constraints do make an assumption about the structure of the variance–covariance matrix that could be at odds with the true data-generating model. In order to test whether this assumption substantially impacts the quality of the estimates, we generated data that varied the item correlations within each scale. We generated the data in the same way using a latent factor model, but we used five different factor loadings to induce correlation differences. The 10-item condition repeated this set of five loadings twice, and the 20-item condition used this block of loadings four times. These loadings produced intra-scale item correlations ranging from roughly .15 (i.e., .39 × .39) to approximately .55 (i.e., .74 × .74). In order to facilitate comparison, we used the same combination of eight conditions from the previous simulation.

We can describe the results from Simulation 3 very succinctly, and interested readers can find a full graphical summary in Figures S22 through S24. First, all percent bias values were less than 10%, and the trellis plot for these quantities can be seen in the figures. We believe that the range of within-scale correlation values likely represents an upper bound for many unidimensional scales in widespread use. The absence of meaningful bias is encouraging and suggests that these arbitrary constraints on the predictors may not impact the focal model parameters, even when the model-implied covariance structure is at odds with the data. It is again noteworthy that the number of items exceeded the total sample size in half of the design cells (conditions with 60 and 120 items). The coverage results were again somewhat unremarkable and were all within Bradley's bound. Finally, MSE ratios were not relevant for this simulation because we focused on conditions where FCS multiple imputation was impossible.

Table S4

Convergence Values for all Conditions in Simulation 3

Condition	Convergence FB	Convergence FCS
N_{scales}= 2, N_{items} = 10, P_{miss}= .40, N = 50	100%	68.9%
N_{scales} = 2, N_{items} = 20, P_{miss} = .25, N = 50	100%	0%
N_{scales} = 2, N_{items} = 20, P_{miss} = .40, N = 50	100%	0%
N_{scales} = 2, N_{items} = 20, P_{miss} = .40, N = 100	100%	83.5%
N_{scales} = 6, N_{items} = 10, P_{miss} = .40, N = 50	98.8%	0%
N_{scales} = 6, N_{items} = 20, P_{miss} = .25, N = 50	100%	0%
N_{scales} = 6, N_{items} = 20, P_{miss} = .40, N = 50	99.1%	0%
N_{scales} = 6, N_{items} = 20, P_{miss} = .40, N = 100	100%	0%

Figure S22

Percent Bias Values for All Conditions in Simulation 3

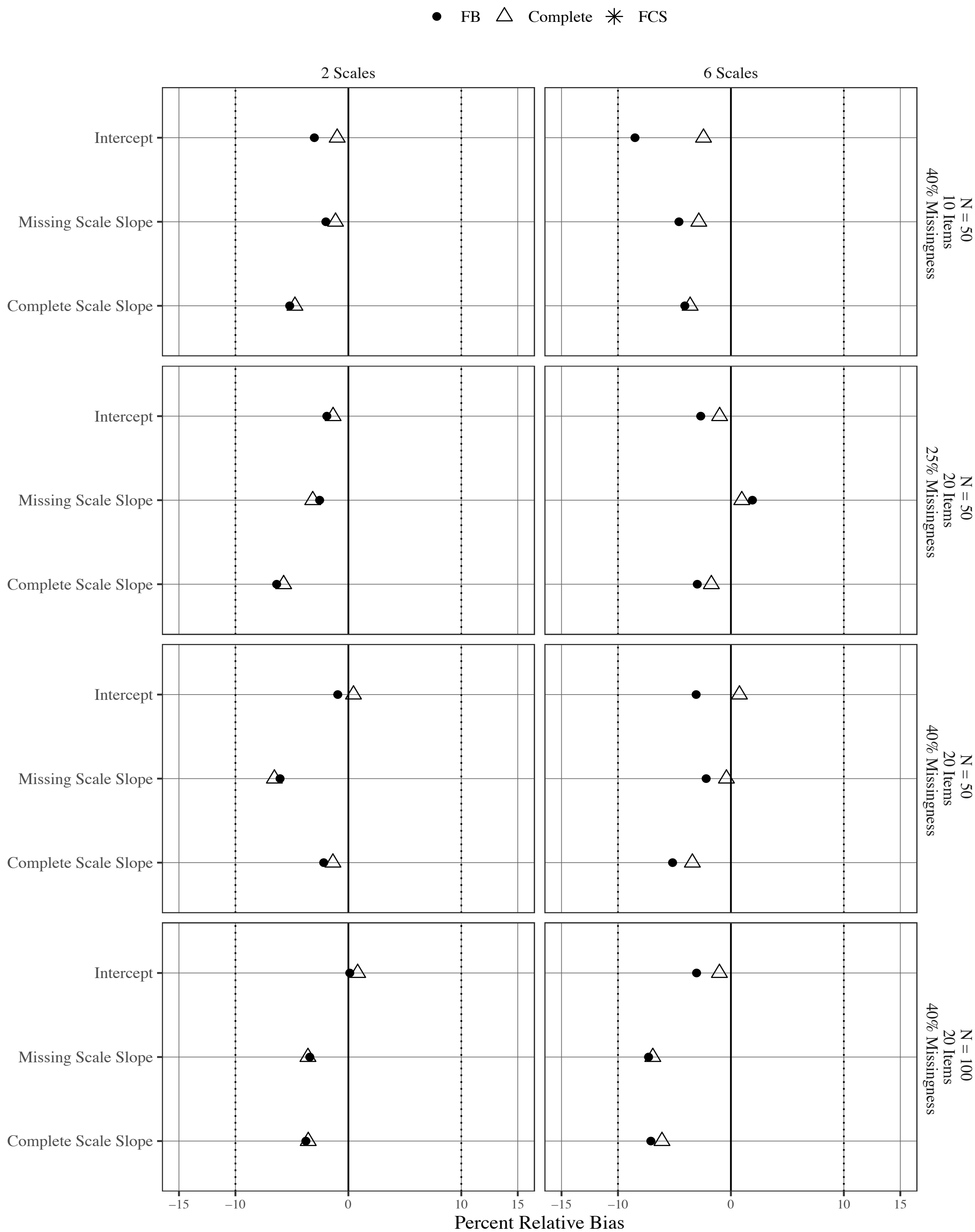


Figure S23

Coverage Rates for All Conditions in Simulation 3

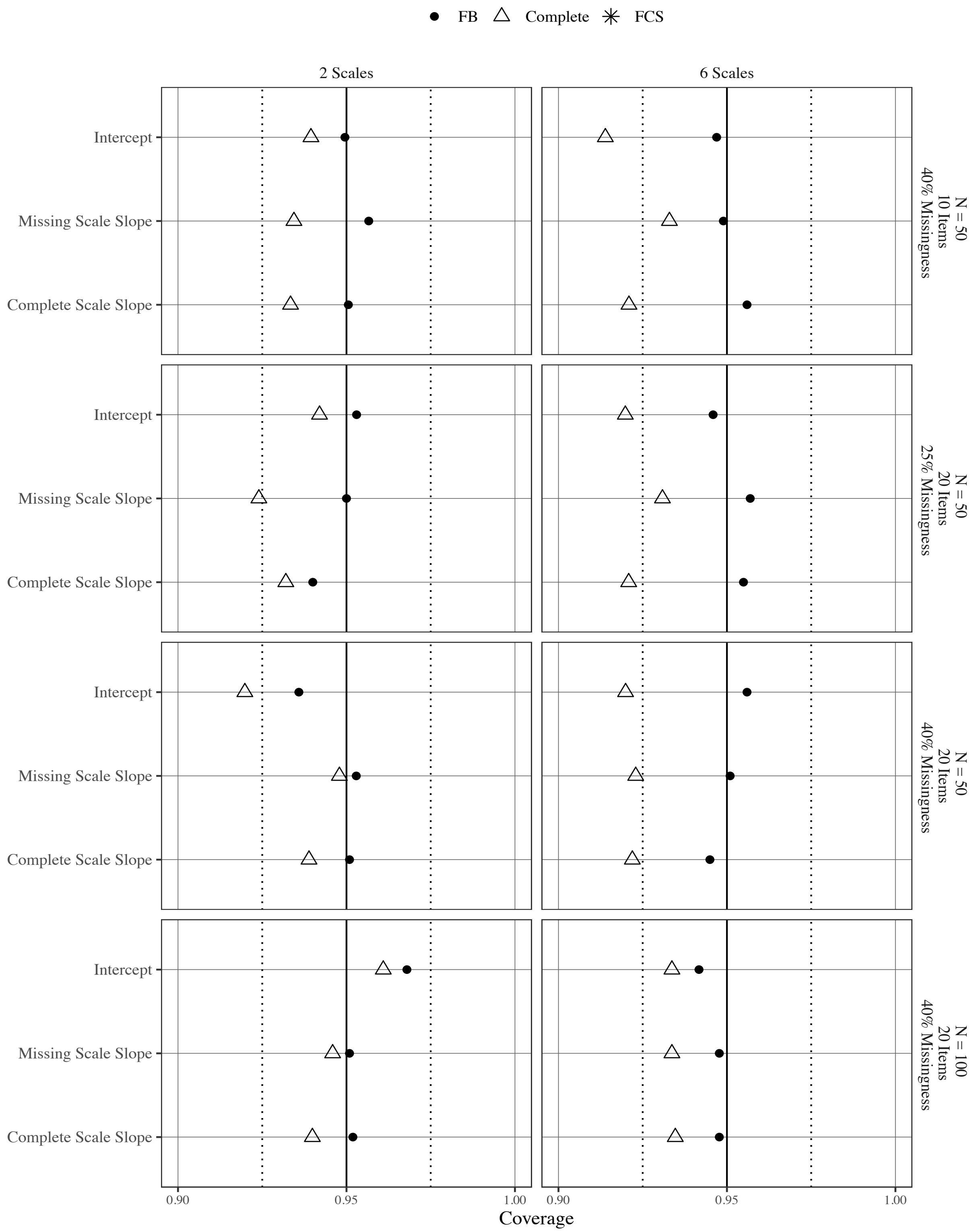
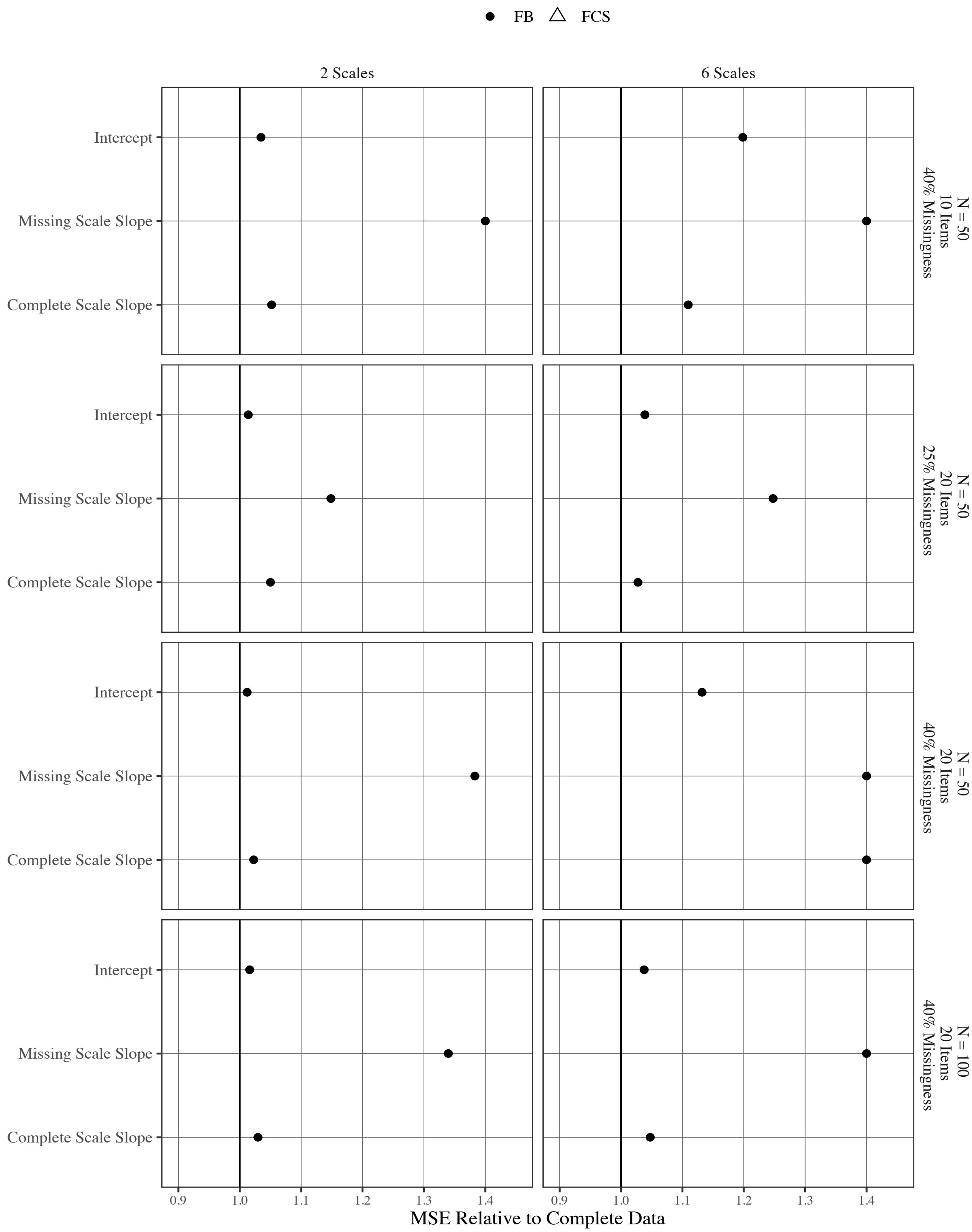


Figure S24

MSE Ratios for All Conditions in Simulation 3



Supplemental Simulation Study 4: Categorical Items

The first three supplemental simulation studies considered composite score predictors that were functions of continuous items, but categorical responses (e.g., questionnaire or test items) are an important and predominant application for our proposed method. Accordingly, the fourth simulation study extended the procedure to ordered categorical items. We generated the data in the same way as the first simulation study, but we instead created five categories per item (e.g., a 5-point Likert scale) by using four threshold parameters to carve the underlying normal distribution into discrete segments. The categorical responses followed a symmetric distribution with category percentages equal to 15%, 20%, 30%, 20%, and 15%. Because the indicators were standardized, the threshold parameters are z-score cutoffs from a standard normal distribution (i.e., the lowest threshold, $z = -1.04$, separates the lowest 15% of the distribution from the rest).

Our goal for this simulation was twofold: examine parameter recovery and explore the lower sample size limit for Bayesian estimation. Perhaps not surprisingly, the $N = 50$ conditions did not reliably converge because the lowest and highest categories generally had insufficient number of observations with which to estimate threshold parameters. Estimation might be possible with fewer response options (e.g., binary items have a single, fixed threshold), but we felt that five categories were more broadly representative. As such, we considered sample sizes of $N = 100$ and 150, two scales with 20 items each, six scales with 10 items each, and a 25% missing data rate. Following Simulation 1, we considered a fairly complex model with only between-scale constraints, and all of the supporting item-level models were probit rather than linear regressions. We again used version 3 of the Blimp application to implement factored regression Bayesian estimation, but additional threshold parameters for each item-level regression required a longer burn-in period of 40,000 iterations. We chose these burn-in periods after examining potential scale reduction factors (Gelman & Rubin, 1992) from several artificial datasets.

We can again describe the results from Simulation 4 very succinctly, and Figures S25 through S27 present the graphical displays. Consistent with the previous simulation studies, the factored regression Bayesian estimates generally tracked closely to those of the complete data, and the focal model parameters were approximately unbiased (we did not consider the bias in the item-level thresholds, as these supporting parameters are not of substantive interest). We view these results as encouraging in light of the fact that the procedure can accommodate up to 60 categorical items with a sample size of only $N = 100$. Finally, coverage values were comparable to those in Figure 4, and MSE ratios were not relevant for this simulation because FCS multiple imputation failed completely.

Table S4
Convergence Values for all Conditions in Simulation 4

Condition	Convergence FB
$N_{scales} = 2$, $N_{items} = 20$, $P_{miss} = .25$, $N = 100$	98.2%
$N_{scales} = 2$, $N_{items} = 20$, $P_{miss} = .25$, $N = 150$	100%
$N_{scales} = 6$, $N_{items} = 10$, $P_{miss} = .25$, $N = 100$	99.89%
$N_{scales} = 6$, $N_{items} = 10$, $P_{miss} = .25$, $N = 150$	100%

Figure S25

Percent Bias Values for All Conditions in Simulation 4

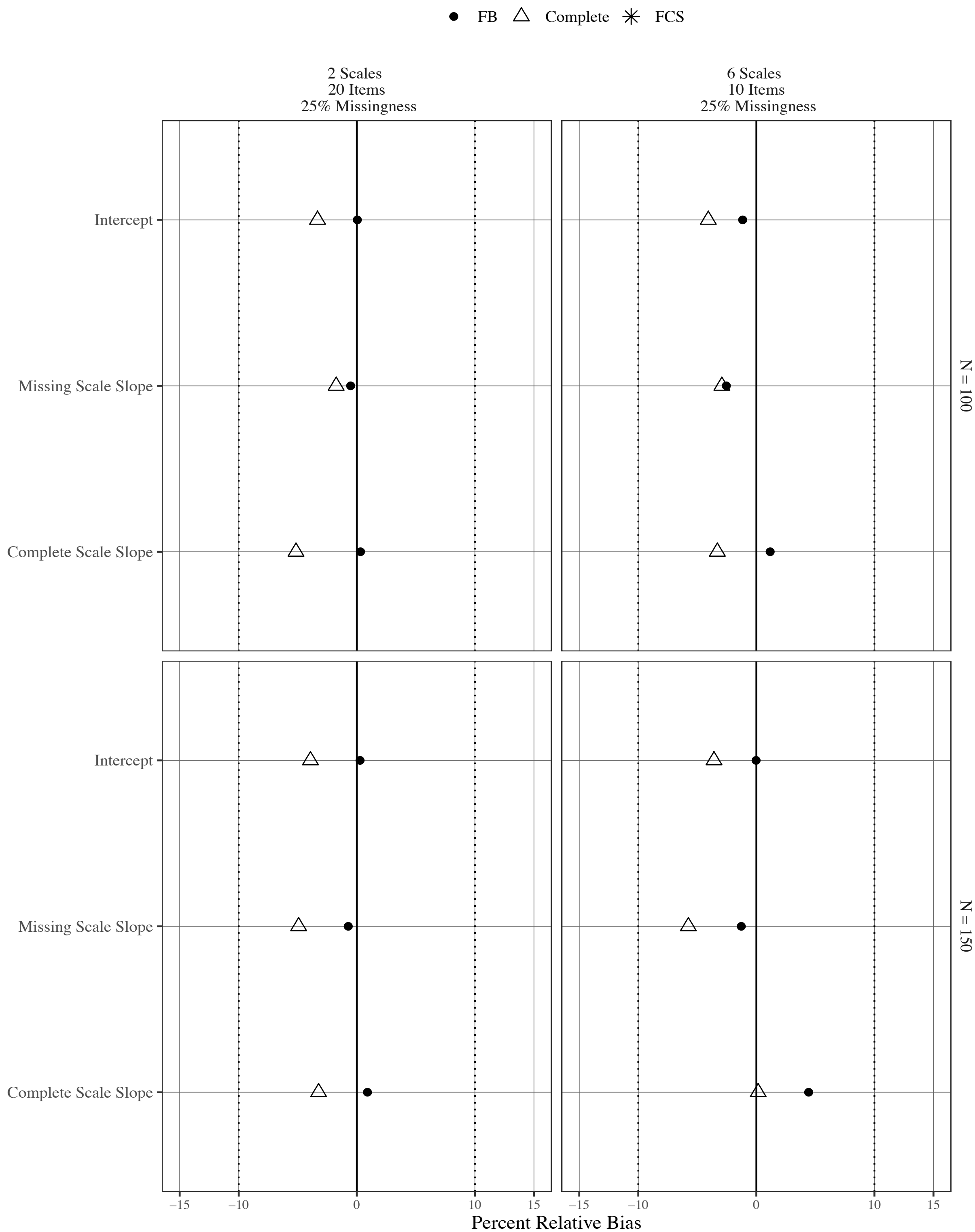


Figure S26

Coverage Rates for All Conditions in Simulation 4

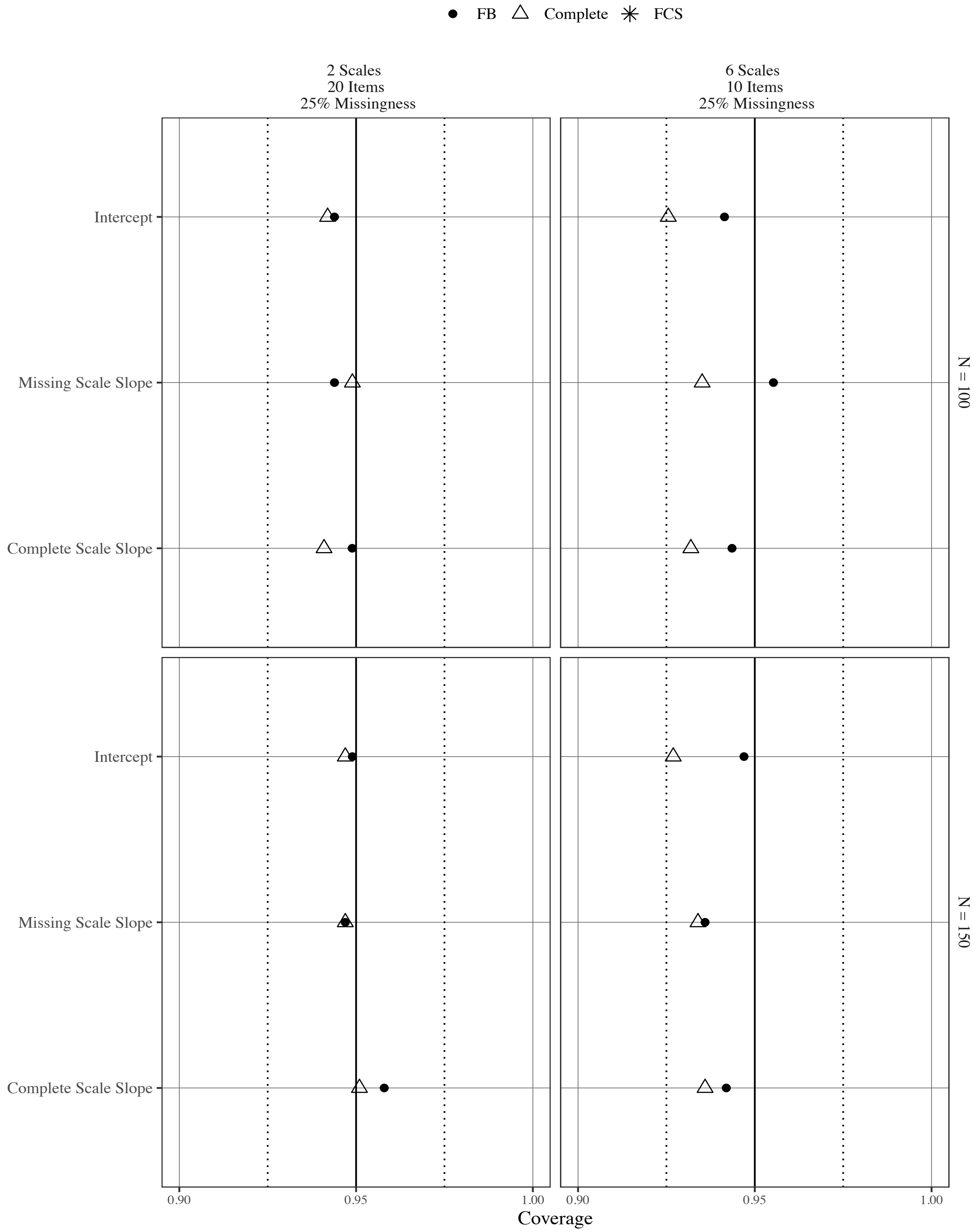


Figure S27

MSE Ratios for All Conditions in Simulation 4