Supplementary materials for 'Multi-level meta-analysis of single-case experimental designs using robust variance estimation'

Contents

S1 Additional results from effect size simulations based on the AR(1) Poisson model

S1.1 Sampling variance of effect sizes and correlation between effect sizes and sampling variances

The MLMA model makes the assumption that the model errors are independent of the sampling variance V_{jk}. This assumption will be violated to some degree if the sampling variance estimator V_{jk} is correlated with either the effect size estimator T_{jk} or with the true effect size parameter γ_{jk}, as would be the case if the true sampling variance depends strongly on the magnitude of the effect size parameter. In this section, we examine further sampling properties of the effect size estimators and their variance estimators in order to understand whether such conditions hold.

S1.1.1 LRR

Figures S1 and S2 depict the sampling variance of the LRR estimators and the correlation between the effect size estimators and their variance estimators, for LRR_1 and LRR_2, respectively. The top row of each figure shows that the sampling variance of the effect size estimators depends on the magnitude of the LRR parameter. For negative parameter values, sampling variance increases with the magnitude of the true effect size, while for positive parameter values, sampling variance decreases with the magnitude of the true effect size. The degree of dependence is stronger for short phase lengths and reduced for longer phase lengths. The bottom row of each figure shows that the LRR effect size estimators can be moderately correlated with their variance estimators and that the strength of the correlation becomes larger when the absolute magnitude of the true effect size is larger.

S1.1.2 WC-SMD

Figure S3 depicts the sampling variance of the WC-SMD effect size estimator (top row) and the correlation between this effect size estimator and its sampling variance (bottom row). For short phase lengths ($m = n = 5$), the sampling variance of g is non-monotically related to the true effect size, with larger sampling variance when the effect size is larger in absolute magnitude. The degree of dependence is reduced when phase lengths are longer. Furthermore, for larger true effect sizes, the effect size estimator is strongly correlated with its sampling variance, with sign corresponding to the sign of the true effect size.

S1.1.3 Tau

Figure S4 depicts the sampling variance of the Tau estimator (top row) and the correlation between this effect size estimator and its sampling variance (bottom row). In the top row, it can be seen that the sampling variance of the Tau estimator strongly depends on the magnitude of the true effect size. Sampling variance decreases towards zero as the true parameter value approaches either extreme of the range of Tau–a very different pattern of dependence compared to the patterns for the LRR and WC-SMD effect sizes. In the bottom row of the figure, it can be seen that the Tau estimator can also be strongly correlated with its variance estimator. For negative values of the Tau parameter, the correlation is positive and grows stronger when Tau is closer to the extreme of its

Figure S1: Sampling variance of LRR1 estimator and correlation between LRR1 estimator and its variance estimator under the Poisson AR(1) model.

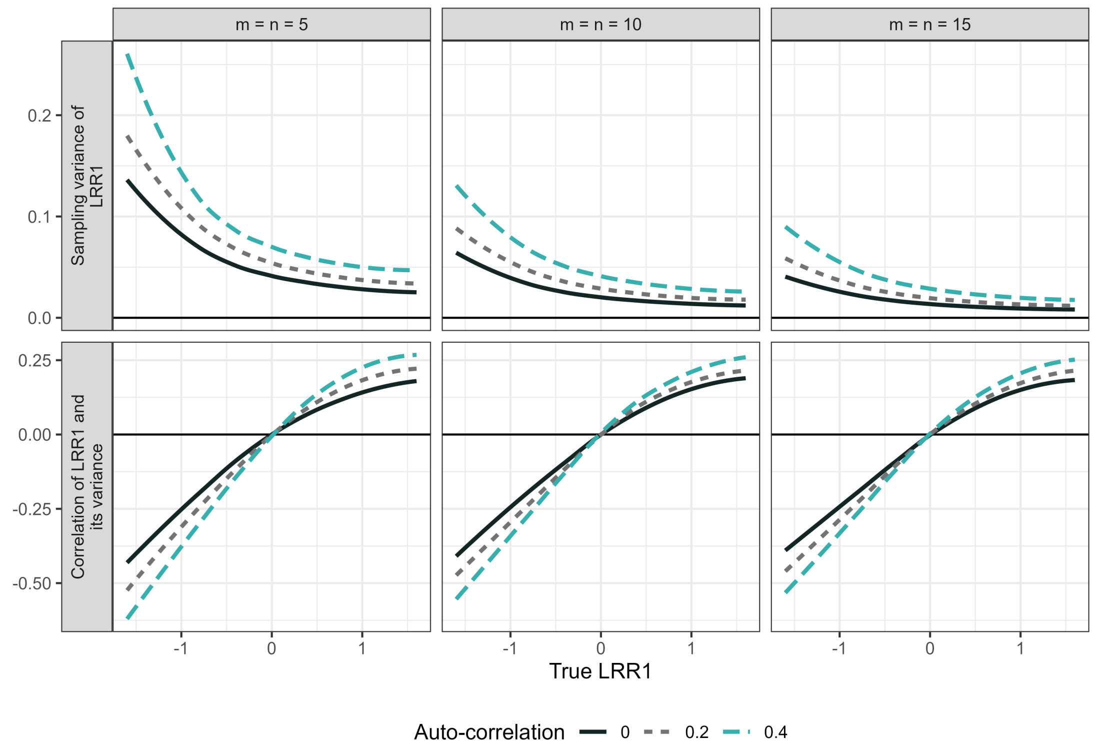


Figure S2: Sampling variance of LRR2 estimator and correlation between LRR2 estimator and its variance estimator under the Poisson AR(1) model.

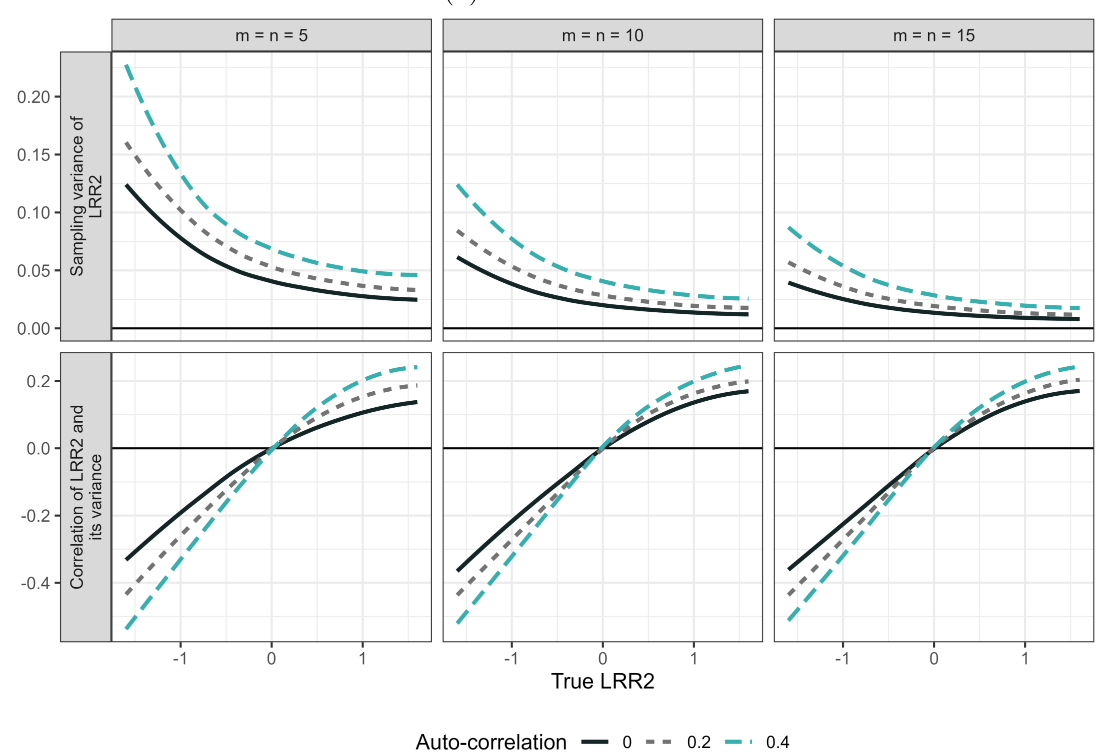


Figure S3: Sampling variance of WC-SMD estimator and correlation between WC-SMD estimator and its variance estimator under the Poisson AR(1) model.

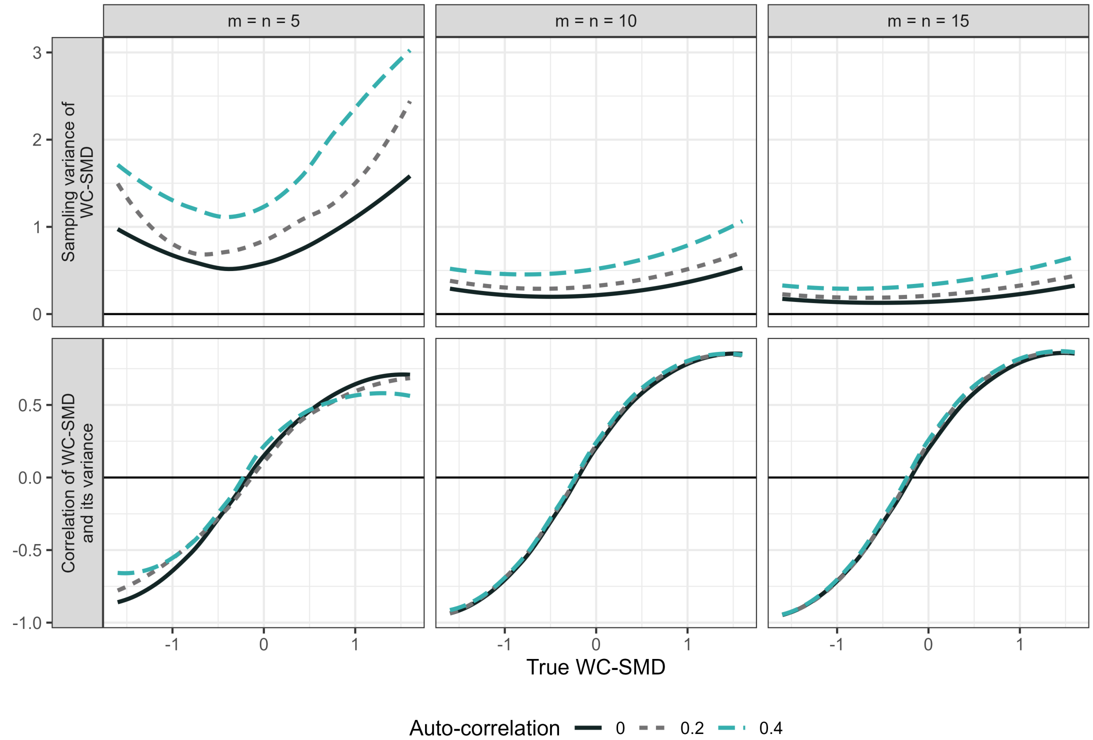
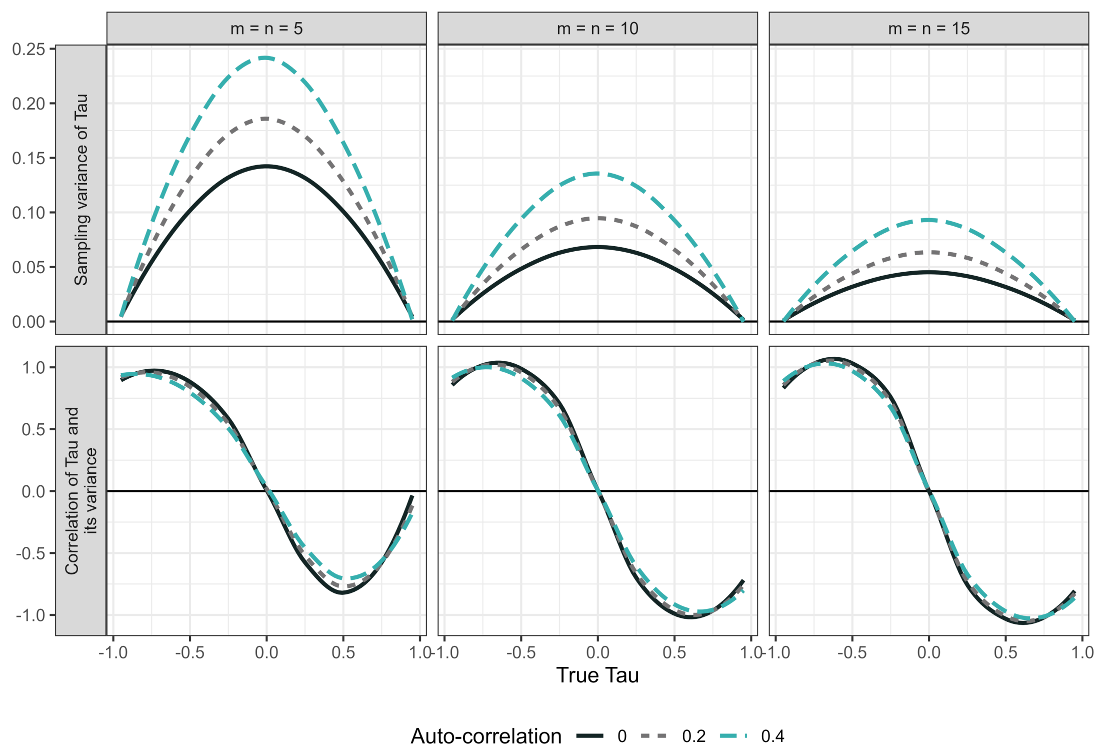


Figure S4: Sampling variance of Tau estimator and correlation between Tau estimator and its variance estimator under the Poisson AR(1) model.

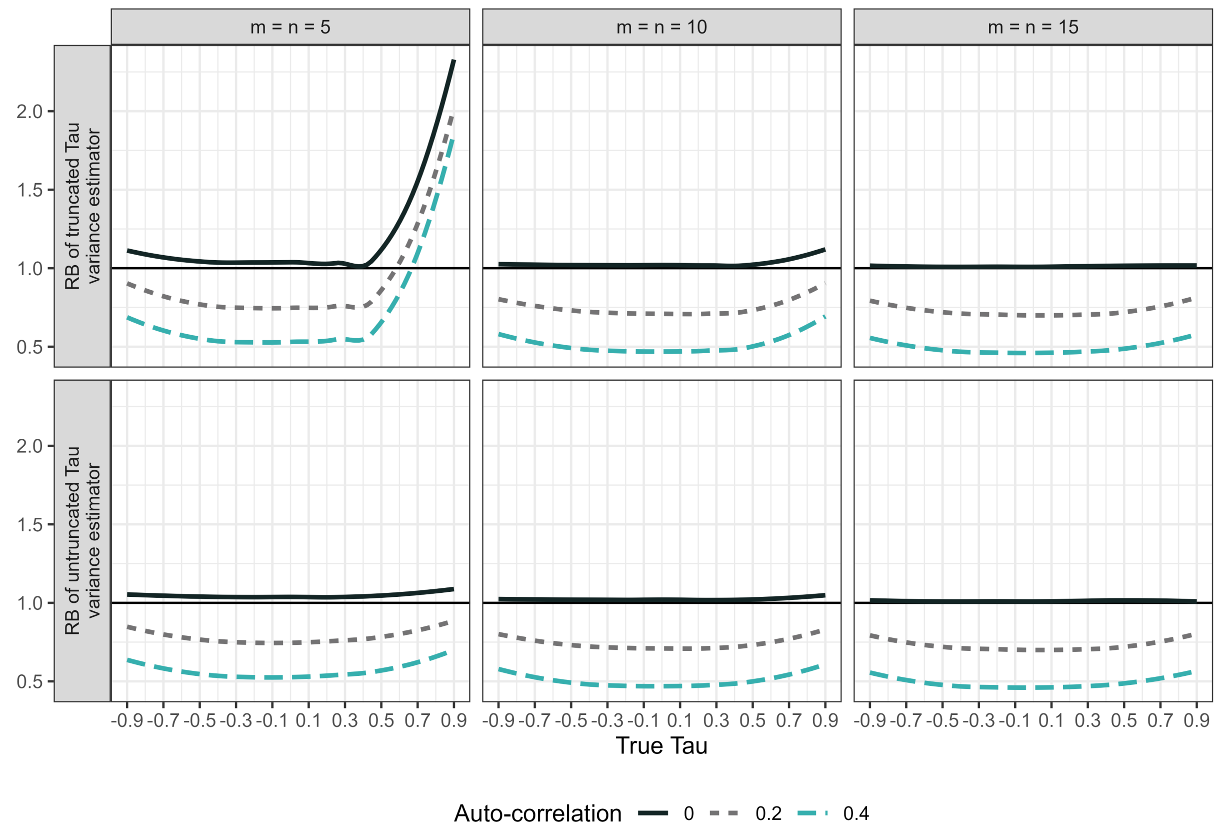


range; for positive values of the Tau parameter, the correlation is negative and attains a maximum near 1 when the Tau parameter is near .6. Again, this is a markedly different pattern compared to what was observed for the LRR and WC-SMD effect sizes.

S1.2 Relative bias for the untruncated variance estimator of Tau

Figure S5 compares the relative bias of the truncated (top row) and untruncated (bottom row) variance estimators for Tau. Overall, the truncated variance estimator has severe bias for short phase lengths ($m = n = 5$) and extreme positive Tau parameters. For short phase lengths ($m = n = 5$), the truncated variance estimator has severe bias for large positive Tau parameter values, in the absence of auto-correlation. However, the untruncated variance estimator has only slight positive bias under the same condition. In the presence of auto-correlation, the discrepancy in the performance of truncated and untruncated variance estimator remains for short phase lengths ($m = n = 5$) and extreme Tau parameters.

Figure S5: Relative bias (RB) of the truncated and untruncated variance estimators for Tau under the Poisson AR(1) model.



S2 Effect size simulations with Gaussian errors

In the effect size simulations presented in the main text, we generated data using an AR(1) Poisson model. Here, we present corresponding simulations based on a model with AR(1) Gaussian errors. Using a model with Gaussian (i.e., normally distributed) outcomes allows us to examine whether any findings from the effect size simulations are specific to the use of a Poisson model with frequency count outcomes. In an outcome series $Y_1, ..., Y_{m+n}$ that follows an AR(1) process, Y_i is marginally normally-distributed with mean μ_i, where $\mu_i = \mu_A$ for $i = 1, ..., m$ and $\mu_i = \mu_B$ for $i = m + 1, ..., m + n$, and unit standard deviation. Deviations from the phase-specific means are correlated at $\text{corr}(Y_h - \mu_h, Y_i - \mu_i) = \phi^{|h-i|}$ for $h, i = 1, ..., m + n$. We simulated single data series using the `arima.sim()` function in R.

Log response ratio effect sizes are only applicable for outcomes that are non-negative. Therefore, for the AR(1) Gaussian simulations, we used truncated normal distributions, in which we first generated a Gaussian AR(1) process $\epsilon_1, ..., \epsilon_{m+n}$ with marginal mean 0 and standard deviation 1, then calculated the outcome as

$$Y_i = \max\{0, \mu_i^* + \epsilon_i\},$$

where μ_i^* is the solution to

$$\mu_i = \mu_i^* + \frac{\phi(\mu_i^*)}{\Phi(\mu_i^*)},$$

so that Y_i has the correct marginal mean, $\text{E}(Y_i) = \mu_i$.

Given a mean level in the baseline phase μ_A and the magnitude of the effect size parameter, we determine the mean level in the treatment phase μ_B as follows:

- The log response ratio effect size parameter is defined as $\lambda = \ln(\mu_B/\mu_A)$. Therefore, just as with the AR(1) Poisson model, μ_B can be calculated as $\mu_B = \mu_A \times \exp(\lambda)$, where exp() denotes the exponential function.
- The WC-SMD effect size parameter is defined as $\delta = (\mu_B - \mu_A)/\sigma_A$. Under the AR(1) Gaussian model, outcomes in the baseline phase are normally distributed with mean μ_A and unit variance, so that $\mu_B = \mu_A + \delta$.
- For an outcome where decrease is desirable, the Tau effect size parameter for a given case is defined as $\theta = 2 \times \Pr(S > 0) + \Pr(S = 0) - 1$, where $S = Y_A - Y_B$. If Y_A and Y_B are independent and normally distributed with unit variances, then S is normally distributed with mean $\mu_A - \mu_B$ and variance of 2. Let Φ denote the cumulative distribution function of the standard normal distribution, with inverse (quantile) function Φ^{-1}. Then $\Pr(S = 0) = 0$, $\Pr(S > 0) = \Phi\left(\frac{\mu_A - \mu_B}{\sqrt{2}}\right)$, and

$$\mu_B = \mu_A - \sqrt{2} \times \Phi^{-1}\left(\frac{\text{Tau} + 1}{2}\right).$$

S2.1 LRR

Figure S6 and S7 show the bias of the LRR_1 and LRR_2 estimator respectively, the relative bias of the variance estimator, the sampling variance of the effect size estimators, and the correlation between effect size and variance estimators, for LRR parameters ranging from $\lambda = -1.6$ to 1.6.

Figure S6: Bias of the LRR1 estimator, relative bias (RB) of its variance estimator, sampling variance of the LRR1 estimator, and correlation between the LRR1 estimator and its variance estimator under the GaussianAR(1) model.

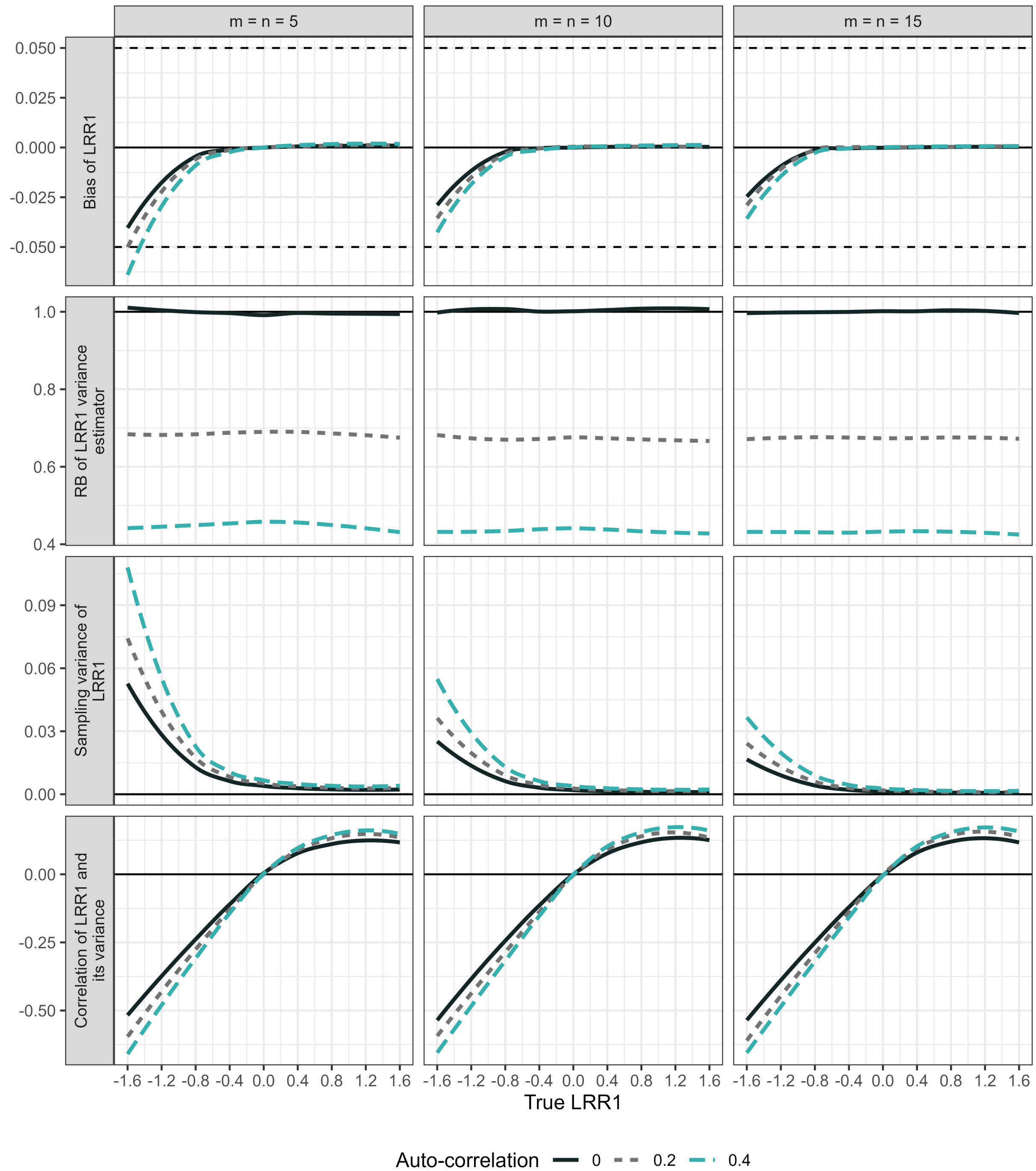
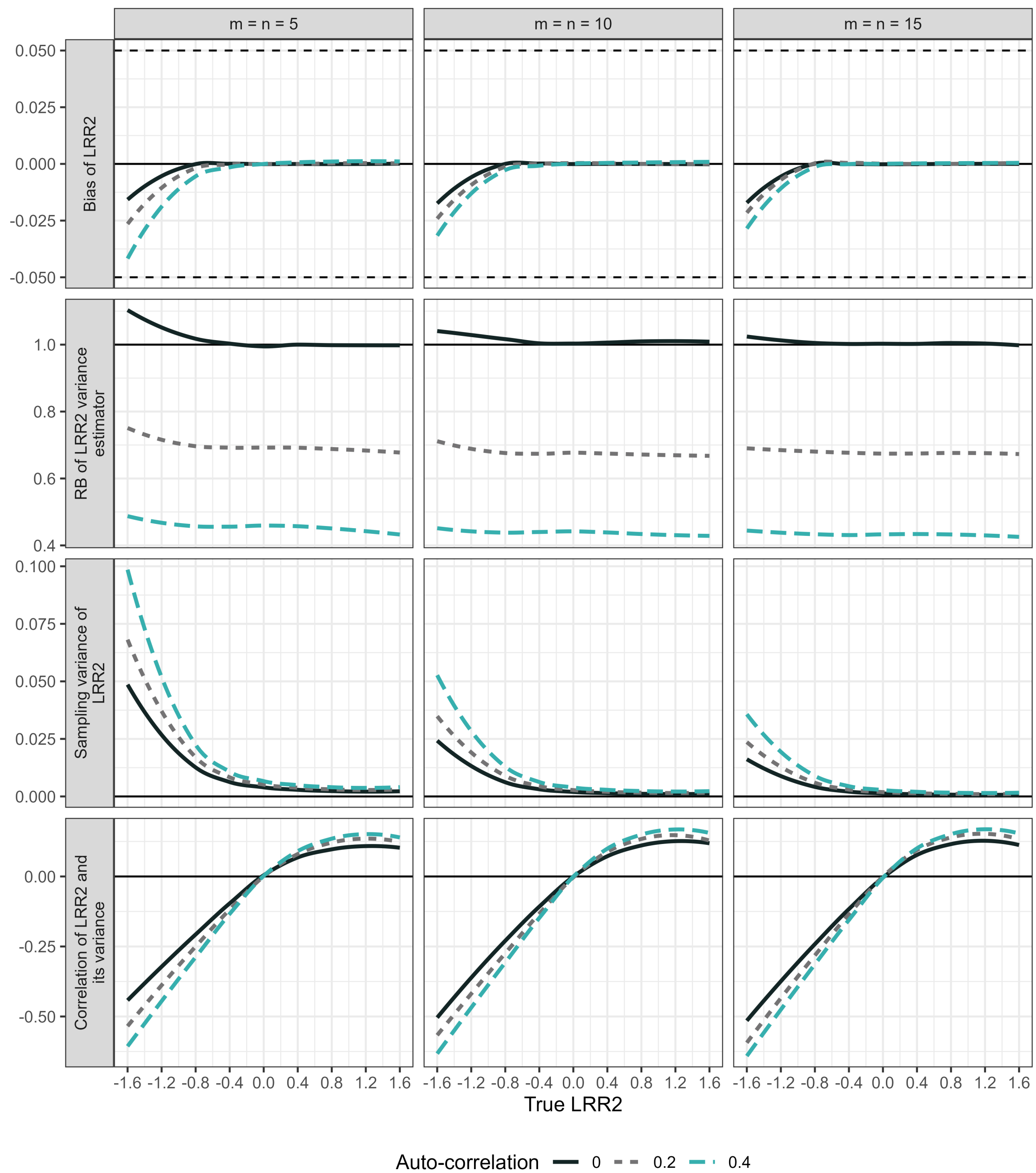


Figure S7: Bias of the LRR2 estimator, relative bias (RB) of its variance estimator, sampling variance of the LRR2 estimator, and correlation between the LRR2 estimator and its variance estimator under the Gaussian AR(1) model.



The LRR_1 and LRR_2 estimator have small negative bias for $\lambda < 0$ and are unbiased for $\lambda > 0$. The bias is slightly smaller for LRR_2 estimator than the LRR_1 estimator. Biases are worse for short phase lengths and large auto-correlation values.

In general, the variance estimators of both LRR_1 and LRR_2 estimators are approximately unbiased when the data are generated independently, except for the small bias in the LRR_2 variance estimator for negative values of λ. The variance estimator tends to under-state the actual sampling variance in the presence of auto-correlation, and the degree of downward bias increases with strength of auto-correlation.

S2.2 WC-SMD

Figure S8 displays the bias of the WC-SMD estimator (Hedges' g), the relative bias of the variance estimator, the sampling variance of the effect size estimator, and the correlation between effect size and variance estimators for effect size parameters ranging from -1.6 to 1.6. As shown in the top row, the WC-SMD estimator is approximately unbiased when the data are generated without auto-correlation. However, the estimator has a multiplicative bias in the presence of auto-correlation, especially when the number of observations per phase is small ($m = n = 5$). Longer phase lengths mitigate the bias caused by auto-correlation. This is consistent with the pattern of the simulation results based on the AR(1) Poisson model. The second row of Figure S8 demonstrates that the sampling variance of g is under-estimated—even in the absence of auto-correlation. As expected, the variance estimator has a strong downward bias for positive values of auto-correlation.

S2.3 Tau

Figure S9 illustrates the bias of Tau estimator, the relative bias of its (truncated) variance estimator, the sampling variance of the effect size estimators, and the correlation between effect size and variance estimators. The Tau estimator is approximately unbiased across the full range of parameter values. The truncation is for variance estimator not for Tau estimator and thus it would not cause bias in the effect size estimator.

Figure S10 compares the relative bias of the truncated (top row) and untruncated (bottom row) variance estimators for Tau. For short phase lengths ($m = n = 5$), the truncated variance estimator has severe bias for large positive Tau parameter values, while the untruncated variance estimator remains nearly unbiased. This is consistent with the finding in the AR(1) Poisson model. In the presence of auto-correlation, the untruncated variance estimator tends to under-estimate the actual sampling variances and higher auto-correlation would lead to stronger downward bias. This is also true for the truncated variance estimator for most conditions with the exception that the truncation leads to increasing positive bias for large Tau parameter values for short phase lengths ($m = n = 5$).

Figure S8: Bias of the WC-SMD estimator, relative bias (RB) of its variance estimator, sampling variance of the WC-SMD estimator, and correlation between the WC-SMD estimator and its variance estimator under the Gaussian AR(1) model.

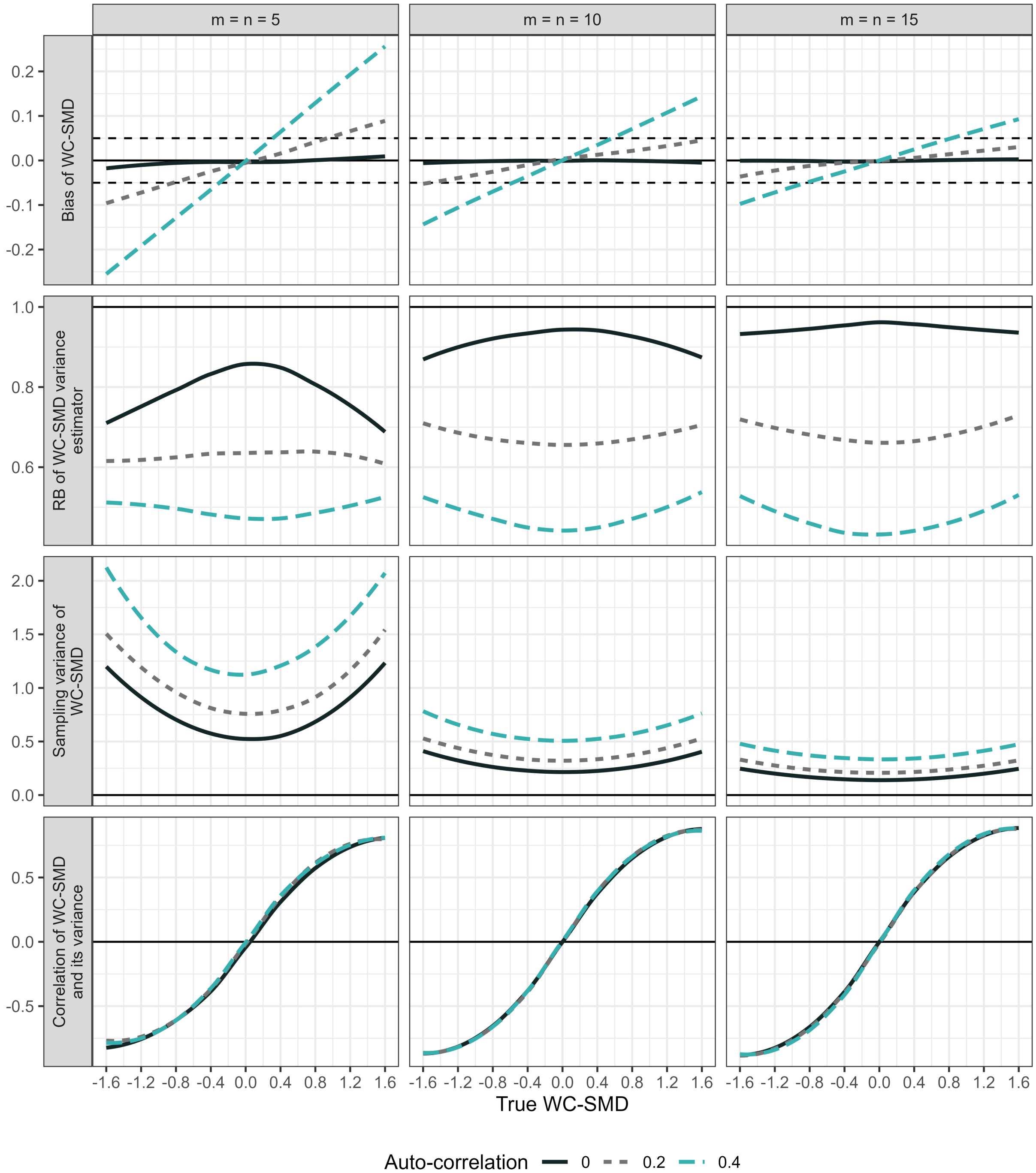


Figure S9: Bias of the Tau estimator, relative bias (RB) of its truncated variance estimator, sampling variance of the Tau estimator, and correlation between the Tau estimator and its truncated variance estimator under the Gaussian AR(1) model.

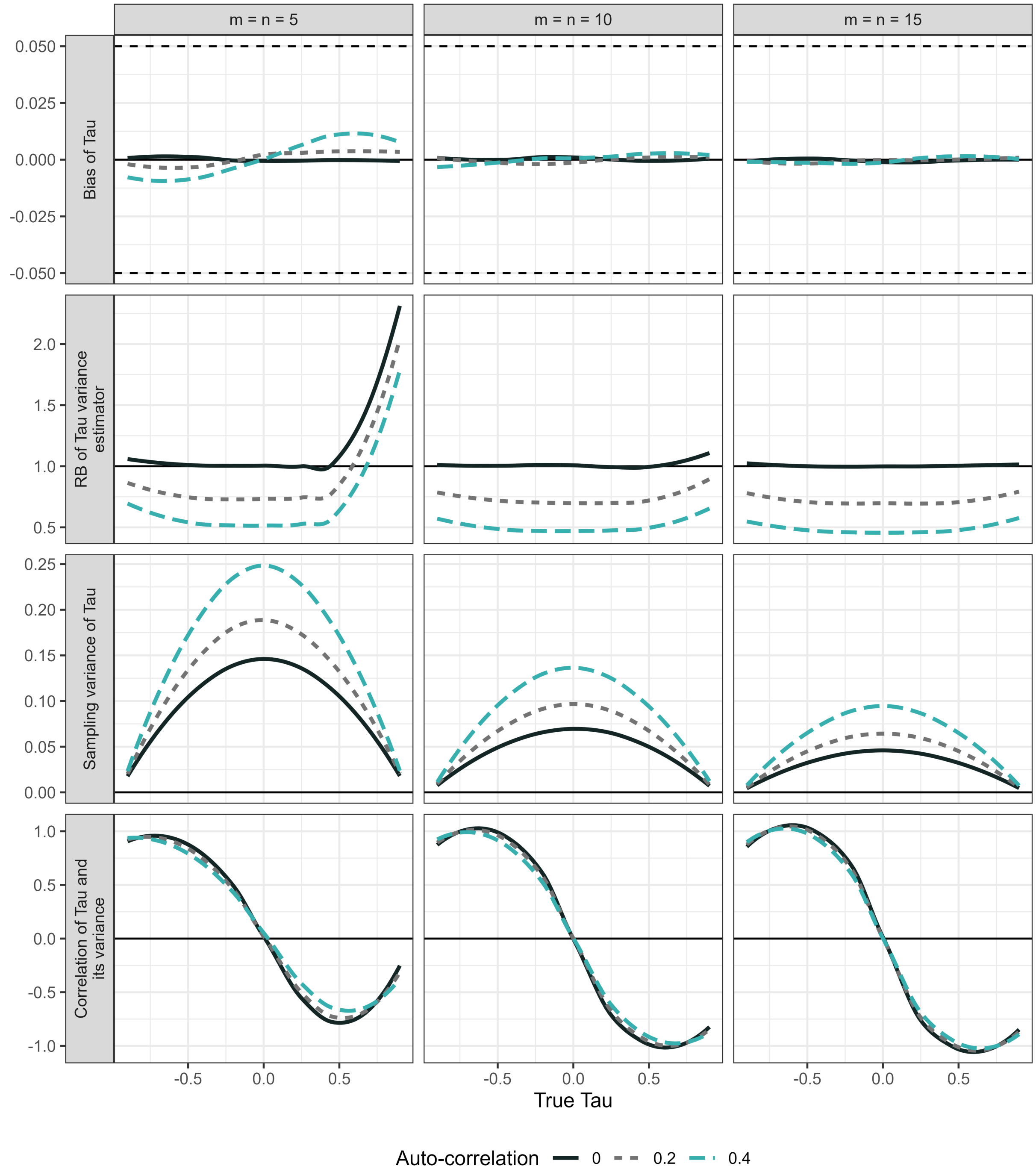
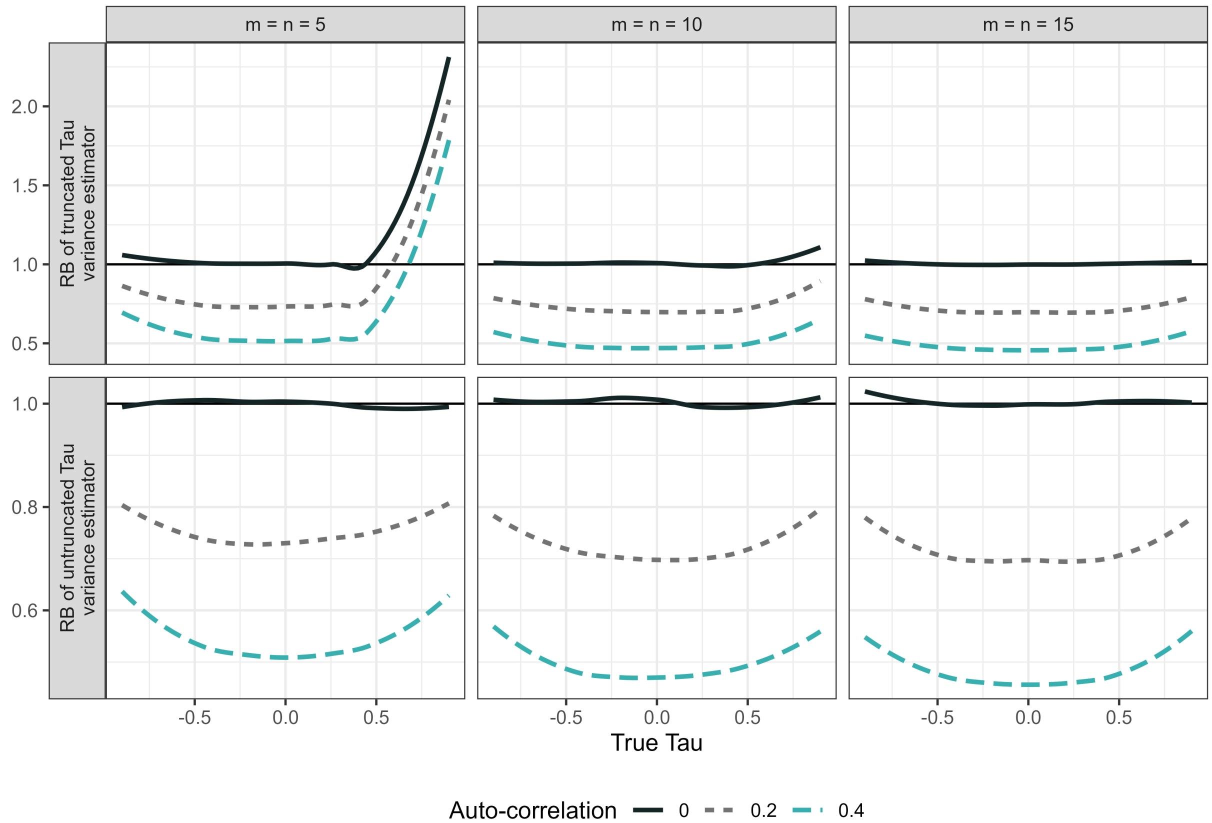


Figure S10: Relative bias (RB) of the truncated and untruncated variance estimators for Tau under the Gaussian AR(1) model.



S3 Additional results from meta-analytic simulations

S3.1 MLMA Convergence

The restricted maximum likelihood estimation process converged very well under most conditions for all three effect size measures. The convergence rates were lower for Tau effect size metric which were above 96% except for one case. The lowest convergence rate for Tau is 92% under the conditions that $\gamma = 0.8,\ K = 30,\ \tau = 0.1,\ \omega = 0.2,\ \phi = 0.4$.

Table S1: Convergence rates for restricted maximum likelihood estimation process

metric	mean	median	min
LRR	1.000	1.000	0.997
Tau	0.998	1.000	0.918
WC-SMD	0.997	0.999	0.974

S3.2 Type I error rates

Figure S11 displays the Type I error rates of OLS-RVE, MLMA, and MLMA-RVE estimators for the four LRR_1, LRR_2, WC-SMD (g), and Tau effect size measures (rows) across different degrees of auto-correlation (columns). We present Type I error rates for $K = 10$ and $K = 30$ (the horizontal axis). Type I error rates of the average effect sizes are examined when the true effect is zero. The conventional nominal level of $\alpha = .05$ is used.

In the top two rows, when used with the LRR_1 or LRR_2 effect sizes, the OLS-RVE and MLMA-RVE estimators have similar Type I error rates that are close to the nominal level, especially when the number of studies is large. In comparison, the MLMA estimator has higher Type I error rates that are above the nominal level. This indicates that applying RVE to MLMA model results in better control of Type I error rates. The results are consistent for LRR_1 and LRR_2 estimators.

In the third row, when used with the WC-SMD effect size, the three meta-analytic methods have similar Type I error rates that are near the nominal level for $K = 10$ but above the nominal level for $K = 30$. The error rates are slightly better controlled when applying RVE to the MLMA model.

In the bottom row, when used with the Tau effect size, OLS-RVE has Type I error rates that are at or close to the nominal level and the estimator performs the best compared to MLMA or MLMA-RVE estimator under most conditions.

Figure S11: Type I error rates of the average effect size estimates.

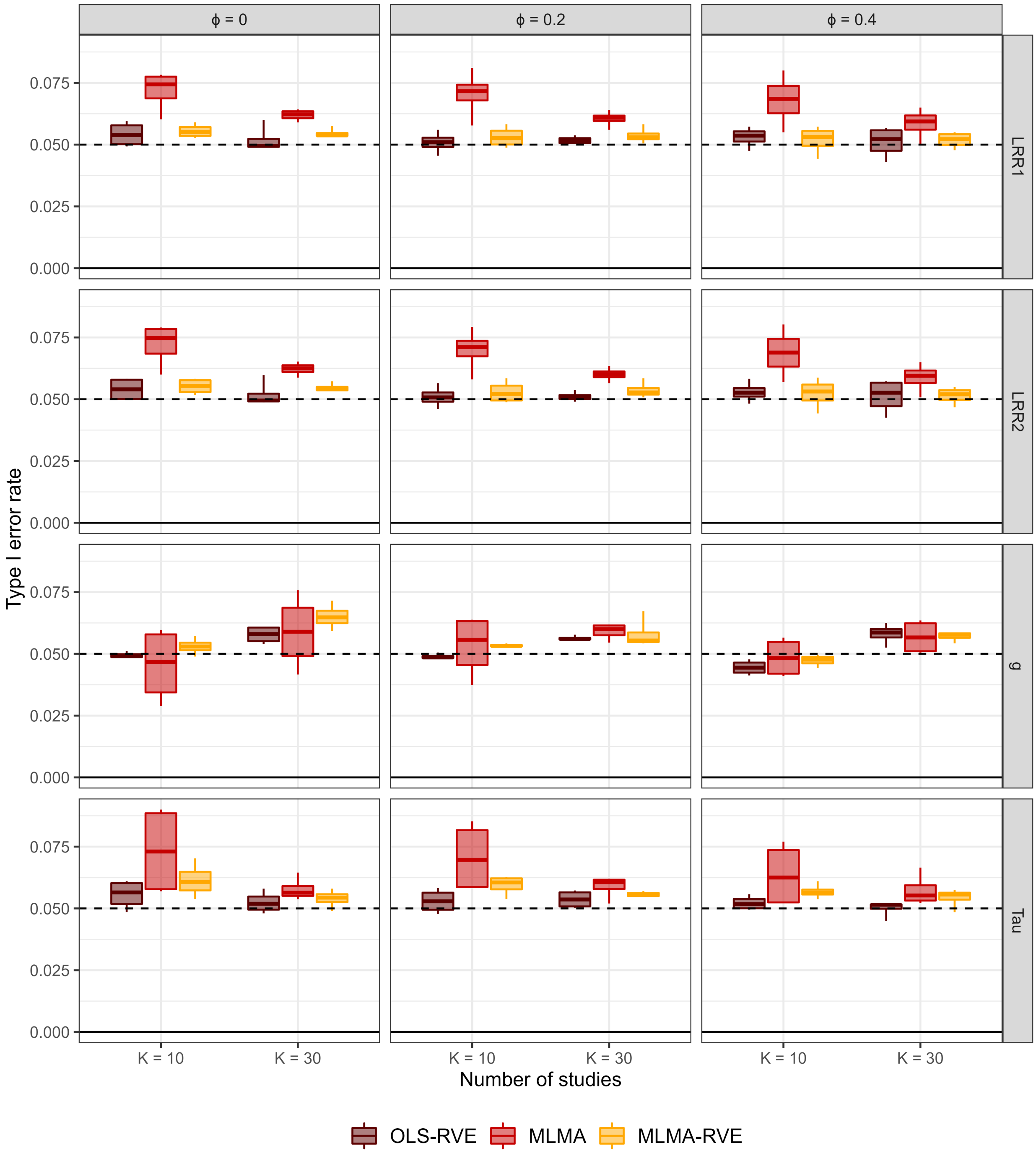
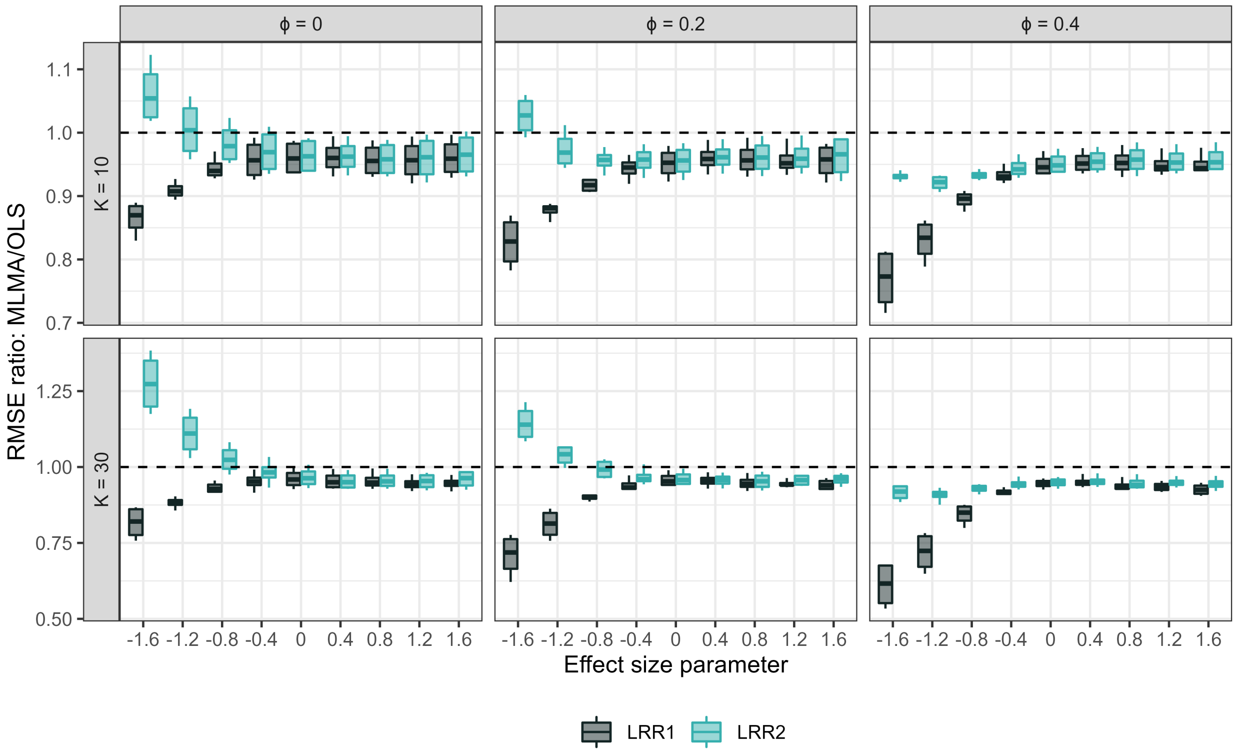


Figure S12: RMSE ratio: MLMA/OLS.



S3.3 RMSE ratio

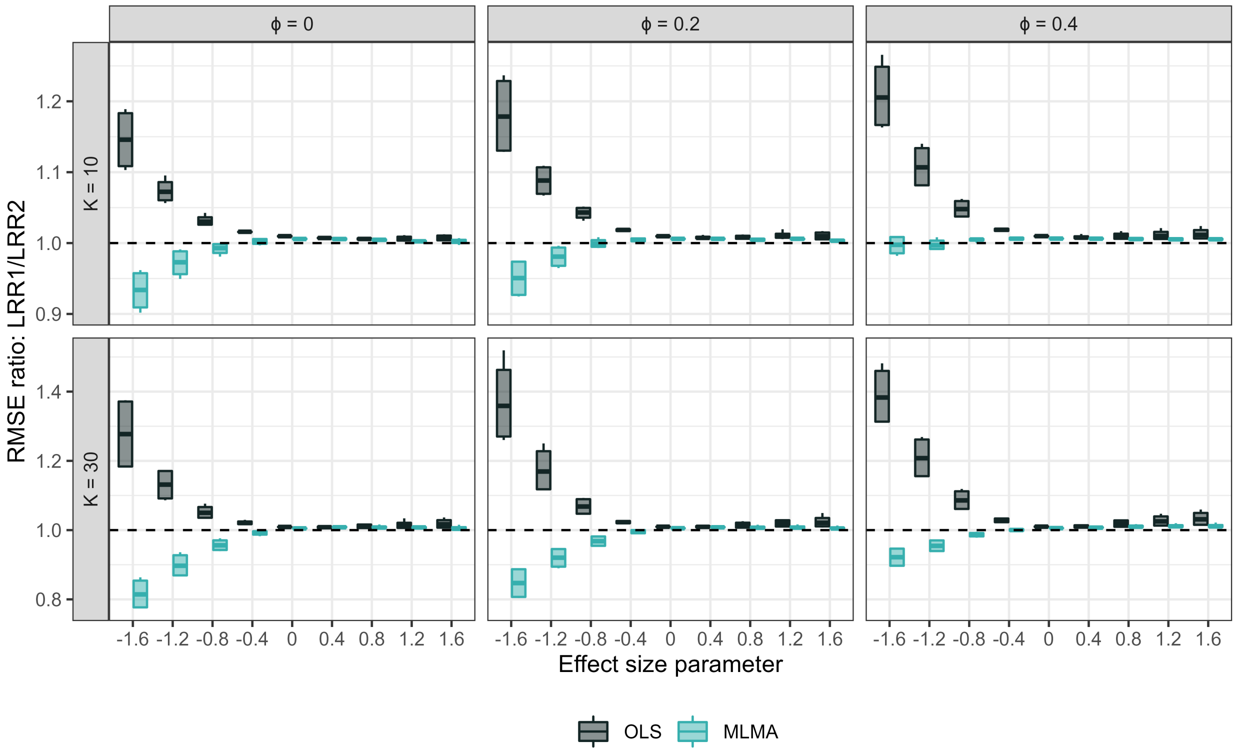
S3.3.1 MLMA/OLS

Figure S12 illustrates that the MLMA model is more accurate than OLS across all conditions for LRR1 estimator in meta-analysis. For LRR1, RMSE of MLMA ranges from 0.534 to 1.003 times the RMSE of OLS, with a median of 0.934. The MLMA model is more accurate than OLS for LRR2 estimator for most conditions. For LRR2, the relative RMSE of MLMA over OLS ranges from 0.876 to 1.384, with a median of 0.958.

S3.3.2 LRR1/LRR2

Figure S13 illustrates that LRR_2 estimator is more accurate than LRR_1 under OLS estimation. However, under MLMA estimation, LRR_1 estimator is more accurate than LRR_2 for larger negative parameter values and has similar accuracy for other conditions.

Figure S13: RMSE ratio: LRR1/LRR2.



S3.4 Variance components

The between-study heterogeneity and within-study heterogeneity in effect sizes were estimated for the MLMA model. We calculate and present the relative bias of the between-study heterogeneity variance estimates $\hat{\tau}^2$ and within-study heterogeneity variance estimates $\hat{\omega}^2$.

S3.4.1 LRR1

Figure S14 displays the relative bias of between-study heterogeneity variance estimates for average LRR_1 when $K = 10$ and $K = 30$. The between-study heterogeneity variance estimates are biased when the true heterogeneity is small ($\tau = 0.1$) and the bias is larger for stronger auto-correlation ($\phi = 0.4$) and for smaller number of studies ($K = 10$). When the true heterogeneity is larger ($\tau = 0.3$), the heterogeneity variance estimates are near unbiased for most conditions.

Figure S15 shows that the relative bias of $\hat{\omega}^2$ is similar for $K = 10$ and $K = 30$. The estimated within-study heterogeneity variances for the overall average LRR_1 are more biased when the true heterogeneity is small and when the auto-correlation value is large. For instance, when $\omega = 0.1$ and $\phi = 0.4$, the estimated within-study heterogeneity variances are severely biased.

Figure S14: Relative bias of the between-study heterogeneity variance estimates for LRR1.

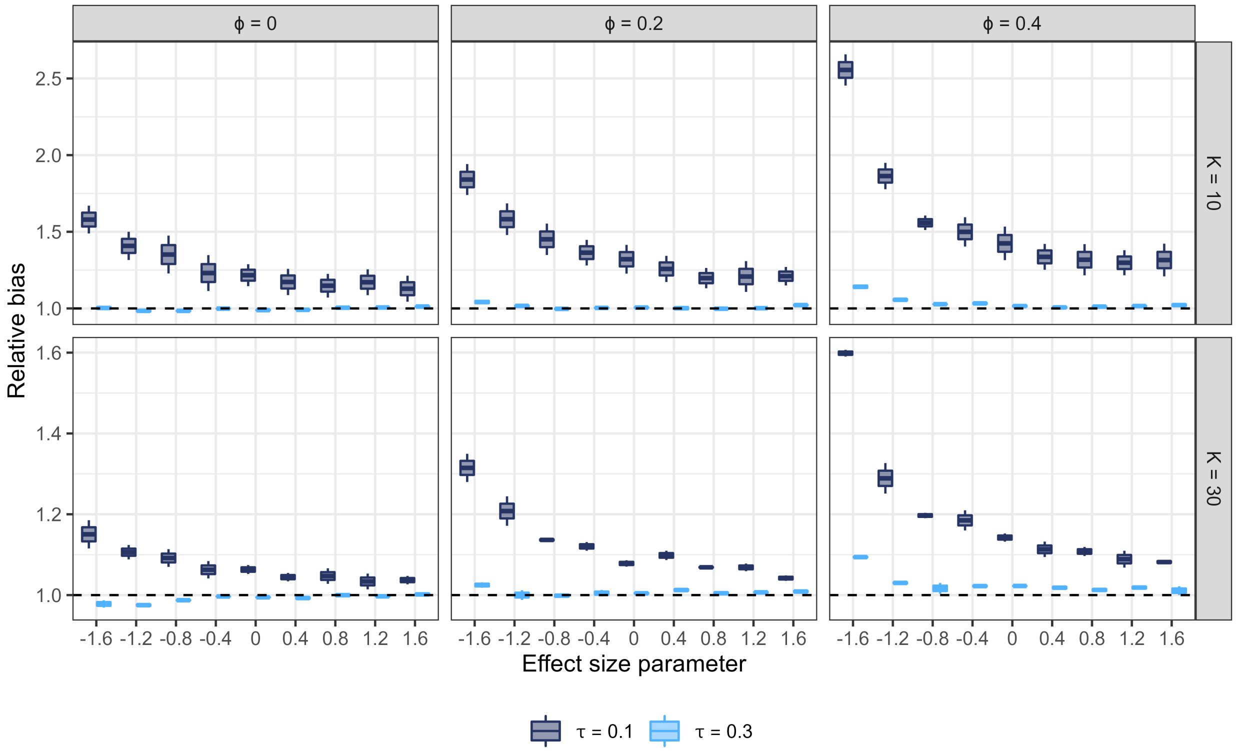
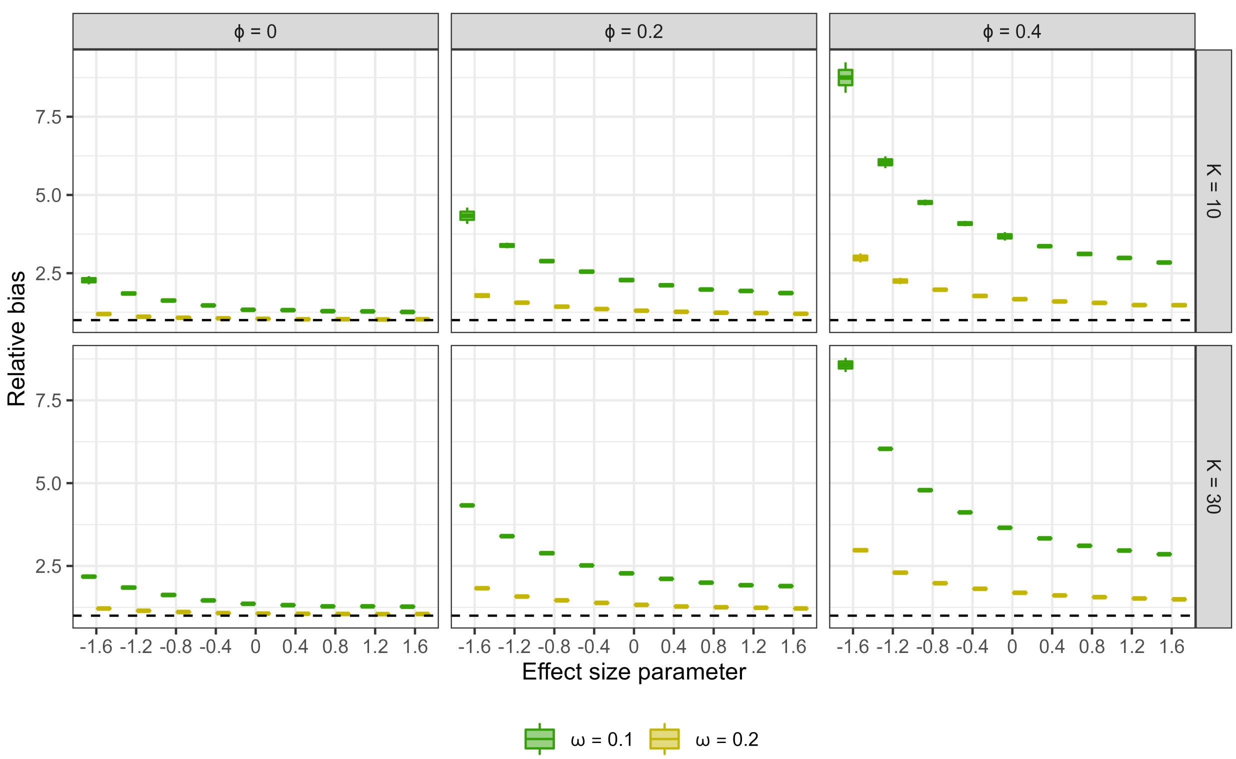


Figure S15: Relative bias of the within-study heterogeneity variance estimates for LRR1.



S3.4.2 LRR2

For LRR_2, despite the number of studies involved, there is positive relative bias in the estimated between-study heterogeneity variance $\hat{\tau}^2$ when $\tau = 0.1$, while there is small bias for negative effect size parameters when $\tau = 0.3$. Importantly, Figure S16 shows that the between-study heterogeneity variance estimates tend to be more biased for large auto-correlation values and for small number of studies.

Figure S17 displays the same pattern of within-study heterogeneity variance estimates for LRR_2 as for LRR_1 shown in the Figure S15. The estimated within-study heterogeneity variances are more biased when the true heterogeneity is small ($\omega = 0.1$) and the auto-correlation value is large ($\phi = 0.4$).

Figure S16: Relative bias of the between-study heterogeneity variance estimates for LRR2.

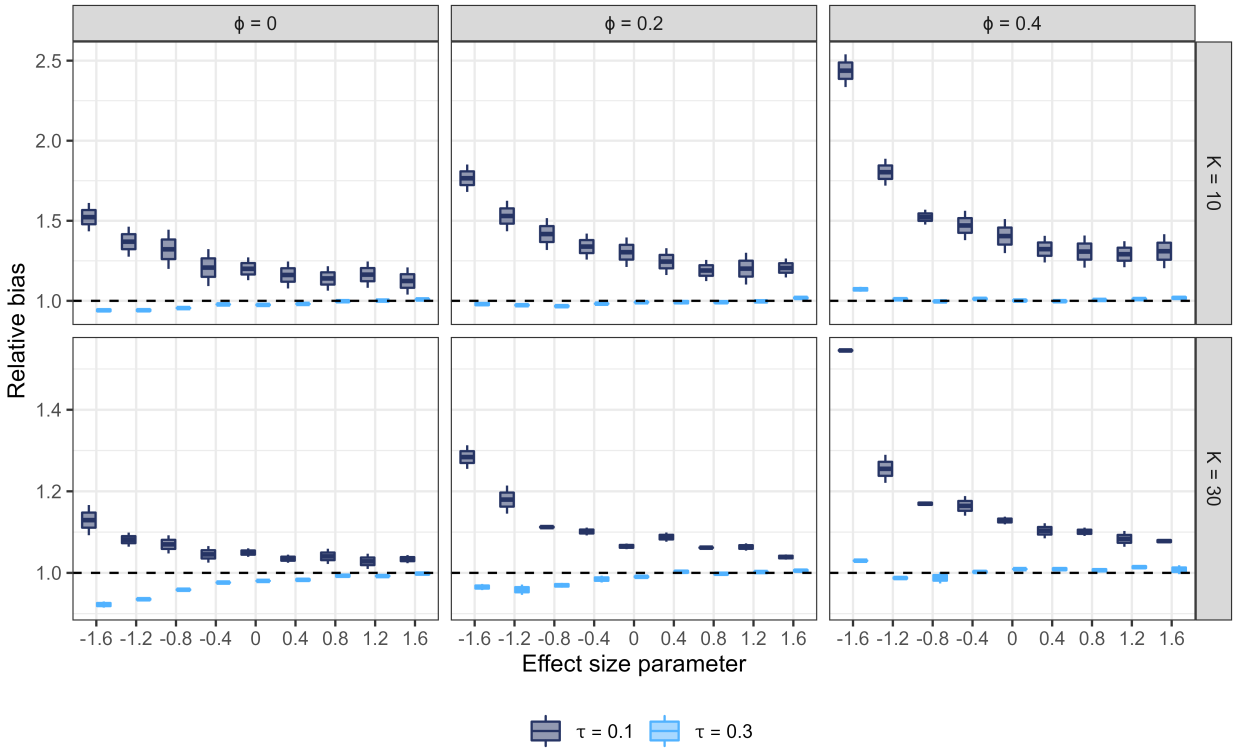
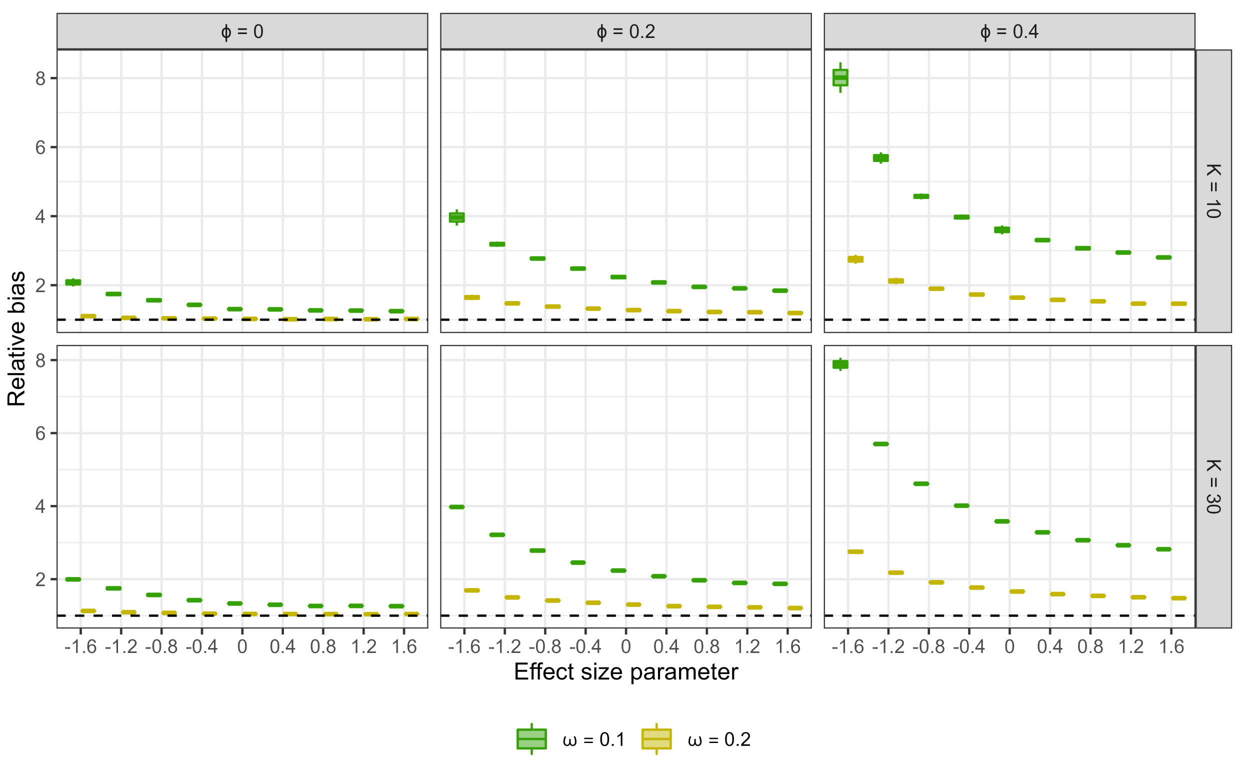


Figure S17: Relative bias of the within-study heterogeneity variance estimates for LRR2.



S3.4.3 WC-SMD

Figure S18 displays that there is bias in the estimated between-study heterogeneity variance and it is worse for small true heterogeneity $\tau = 0.1$, for large auto-correlation values ($\phi = 0.4$), and for small number of studies involved ($K = 10$).

Figure S19 illustrates that the relative bias of $\hat{\omega}^2$ is similar despite the number of studies involved. The within-study heterogeneity variance estimates are more biased for small true heterogeneity value ($\omega = 0.1$) and large auto-correlation value ($\phi = 0.4$).

Figure S18: Relative bias of the between-study heterogeneity variance estimates for WC-SMD.

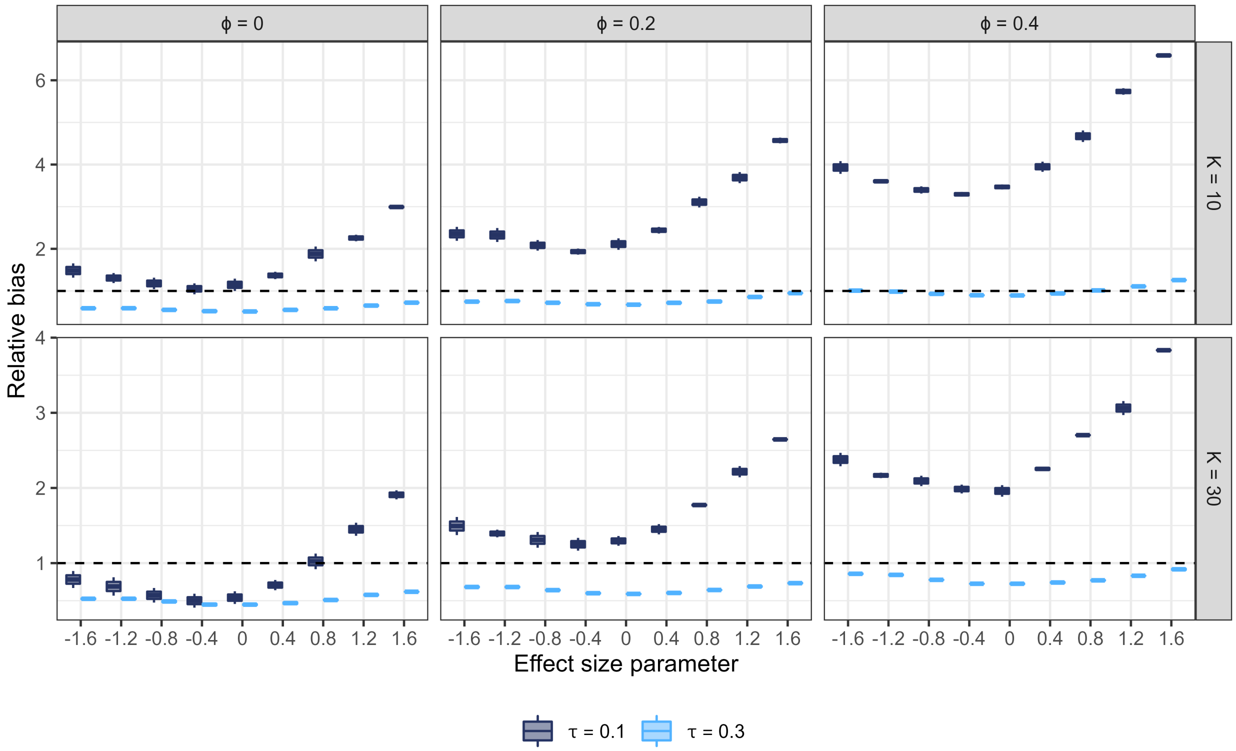
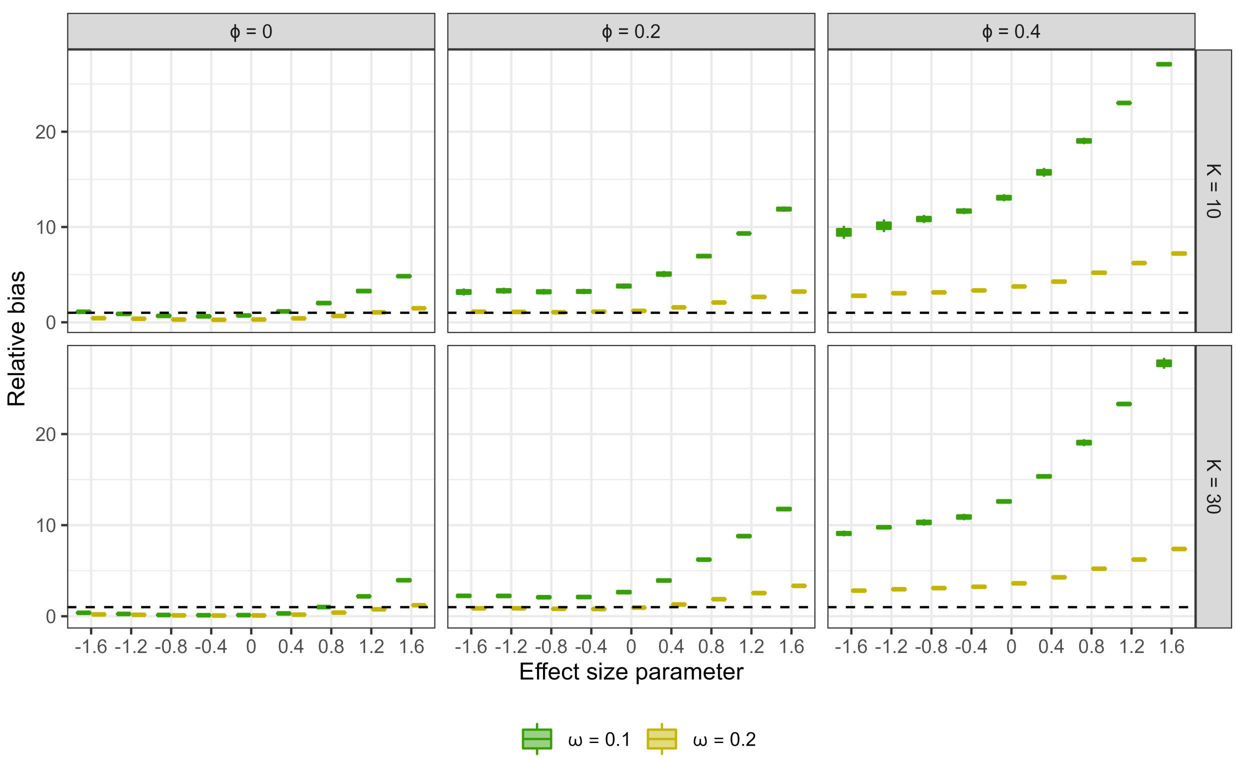


Figure S19: Relative bias of the within-study heterogeneity variance estimates for WC-SMD.



S3.4.4 Tau

Figure S20 shows that the estimated between-study heterogeneity variances are more biased when the true between-study heterogeneity is small ($\tau = 0.1$) and the number of studies is small ($K = 10$). The between-study heterogeneity variance is over-estimated for small true effects and under-estimated for large true effects. The bias does not seem to vary much depending on the levels of auto-correlation.

Figure S21 displays that the relative bias of $\hat{\omega}^2$ is similar for $K = 10$ and $K = 30$. The within-study heterogeneity variance estimates tend to be more biased for small true heterogeneity value ($\omega = 0.1$) and large auto-correlation value ($\phi = 0.4$).

Figure S20: Relative bias of the between-study heterogeneity variance estimates for Tau.

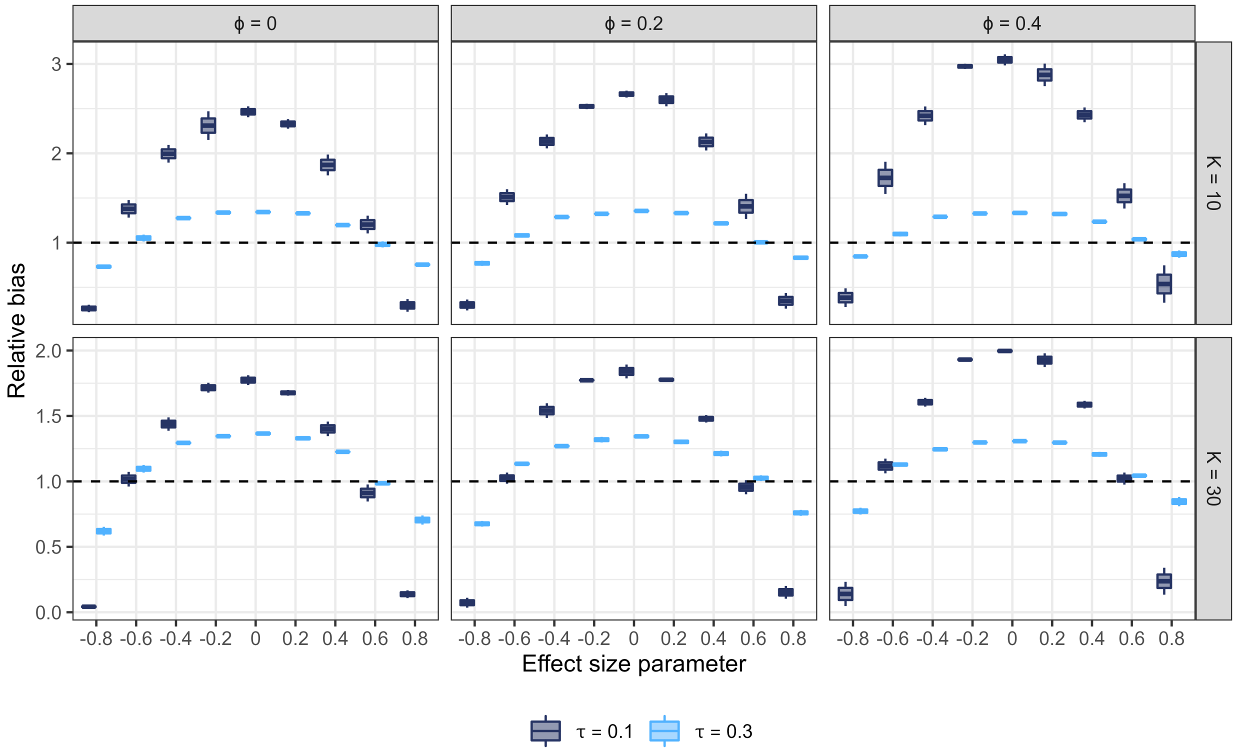


Figure S21: Relative bias of the within-study heterogeneity variance estimates for Tau.