Measurement-comparable effect sizes for single-case studies of free-operant behavior: Simulation results

James E. Pustejovsky

January 29, 2014

This appendix presents the results of a simulation study examining the finite-sample performance of the effect size estimators proposed in the main text. The following estimators are studied:

- simple and bias-corrected moment estimators for the log-incidence ratio using event counting data;
- simple and bias-corrected moment estimators for the log-prevalence ratio using continuous recording data or momentary time sampling data;
- simple and bias-corrected moment estimators for the log-prevalence odds ratio using continuous recording data or momentary time sampling data; and
- a moment estimator for the log-prevalence odds ratio using partial interval recording data.

For each of these estimators, I compare the bias and root mean-squared error (RMSE) of the bias-corrected moment estimator and the simple moment estimator. I then examine the relative bias of the variance estimators corresponding to each of the proposed point-estimators.[1] Finally, I assess the empirical coverage rates of approximate 95% confidence intervals (CIs) based on each of the estimators. I interpret CIs with approximately nominal coverage rates as indirect evidence that the sampling distribution of the point estimator is approximately normal.

1 Simulation design

I study the properties of each effect size estimator under a common data-generating model, which follows the assumptions of the stable-phase model between sessions and the equilibrium alternating renewal process within each session. I make several assumptions to simplify the design of the simulation and moderate its dimensionality. First, I fix the length of each observation session at 600 s. For the momentary time sampling and partial interval recording

[1] For an effect size estimator R and corresponding variance estimator V_R, the relative bias of the variance estimator is defined as $\mathrm{E}(V_R)/\mathrm{Var}(R)$.

methods, I assume that each interval lasts 20 s. Both of these assumptions are within the range of times observed in practice.[2] Next, I assume that each phase contained an equal number of observation sessions, so that $n_0 = n_1 = n$. Finally, I assume that the average event duration remains unchanged between phases, so that $\mu_0 = \mu_1$; this assumption implies that the log-prevalence ratio is equivalent to the log-incidence ratio and that the log-prevalence odds ratio is equivalent to the log-interim ratio.

Using these assumptions, the simulation varies the within-phase sample size, baseline prevalence, baseline incidence, and prevalence odds ratio, all over wide ranges. The within-phase sample size ranges from $n = 4$ to $n = 20$ in steps of 4. The baseline prevalence ranges from $\phi_0 = 0.1$ to $\phi_0 = 0.9$ in steps of 0.1, covering the bulk of the parameter space for this dimension. The baseline incidence ranges from $\zeta_0 = \frac{1}{100}$ (a moderate frequency event) to $\zeta_0 = \frac{1}{10}$ (a high frequency event), within intermediate values of $\frac{1}{90}, \frac{1}{80}, \frac{1}{70}$, etc. The prevalence odds ratio ranges from $\exp(\psi) = 0.5$ (i.e., a decrease in prevalence-odds of 50%) to $\exp(\psi) = 1.5$ (an increase of 50%), in steps of 0.25. Given that $\mu_0 = \mu_1$, the log-prevalence and log-incidence ratios can be determined from the the baseline prevalence and log-prevalence odd ratio as $\omega^\phi = \omega^\zeta = -\log\left[\phi + (1-\phi)\exp(\psi)\right]$.

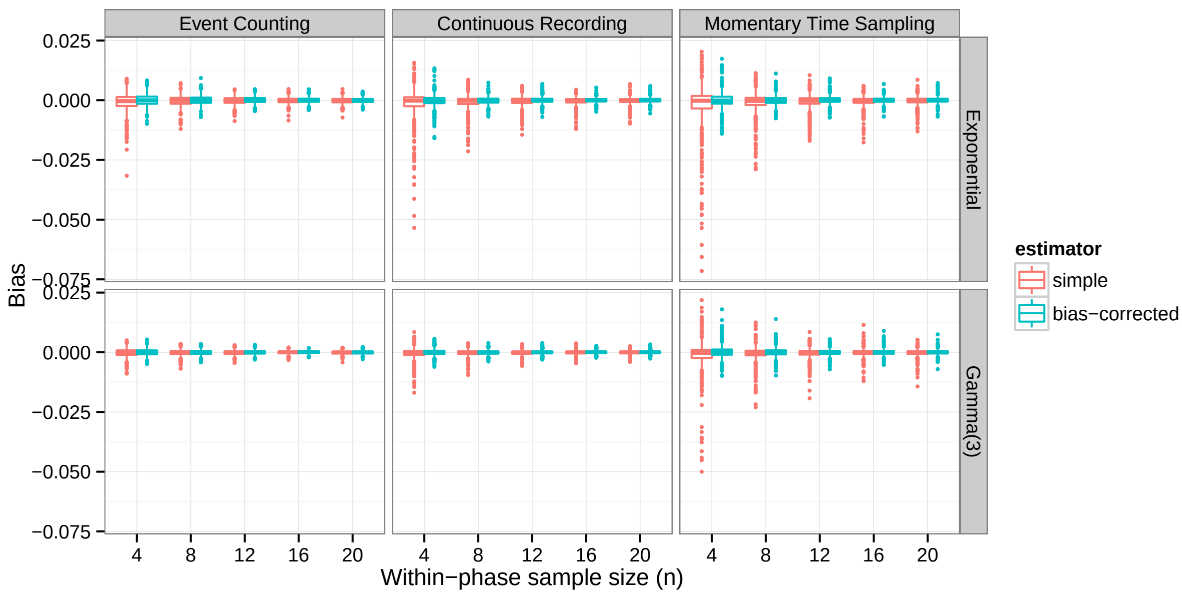
To complete the simulation specification, I select two different parametric forms for the event duration and interim time distributions, both of which are used to generate behavior streams from the alternating renewal process (ARP). In one condition, I use exponential distributions for both event duration and interim time. In the other condition, I use gamma distributions with shape parameters fixed to 3. Both of these choices are arbitrary, but use of more empirically grounded assumptions will not possible until fine-grained continuous recording data on free-operant behavior become available. For each combination of sample sizes, parameter values, and ARP distributions, the distributions of the estimators are simulated over 5,000 replications. R code for reproducing the simulations is available upon request.

2 Log-response ratio estimators

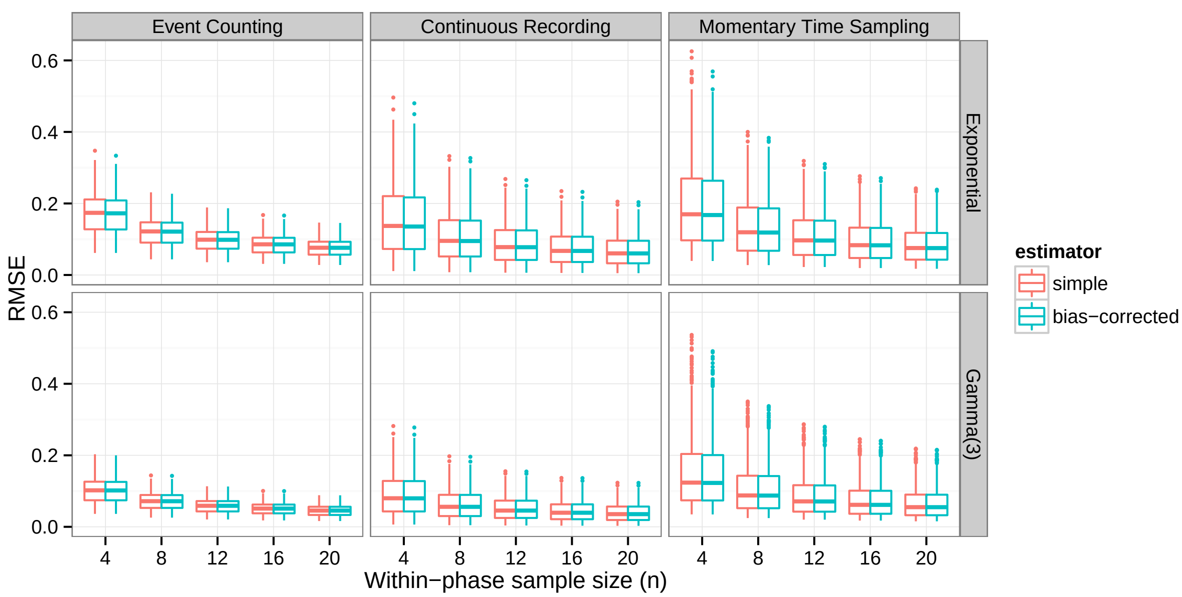
This section presents the results for the log-incidence ratio estimator based on event counting data and the log-prevalence ratio estimators based on continuous recording or momentary time sampling data. Because I have set $\mu_0 = \mu_1$, estimators based on all three types of data share a common estimand, which I refer to as the log-response ratio.

Figure 1 presents results for the bias and RMSE of the estimators. For each type of data, the simple (plug-in) moment estimator is compared to the bias-corrected estimator. It can be seen in Figure 1a that the bias-corrected estimators have very small biases, even for very small n. In comparison, the simple estimators have a wider range of biases, particularly when $n = 4$. In addition, it can be seen in Figure 1b that, given the type of data, form of the ARP distribution, and sample size, the estimators display a comparable range of RMSE values. (In fact, the similarity of RMSE holds even for individual levels of ϕ_0, ζ_0, and ψ.) Given that the bias-corrected estimators have comparable RMSE and reduced bias, they are therefore recommended for use in meta-analysis.

[2] Though 600 s may be shorter than average, longer observation lengths will tend to improve the performance of all estimators.

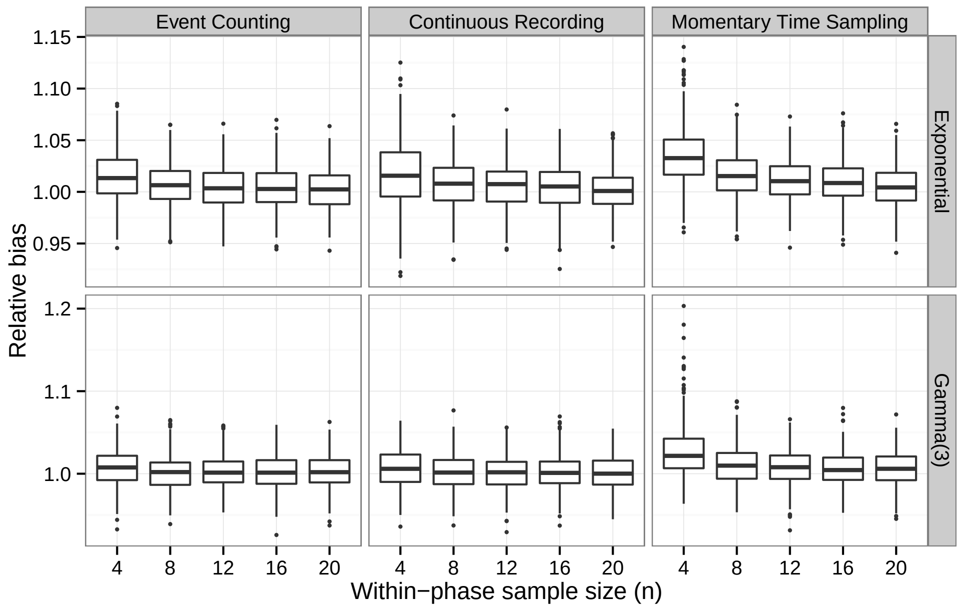


(a) Bias

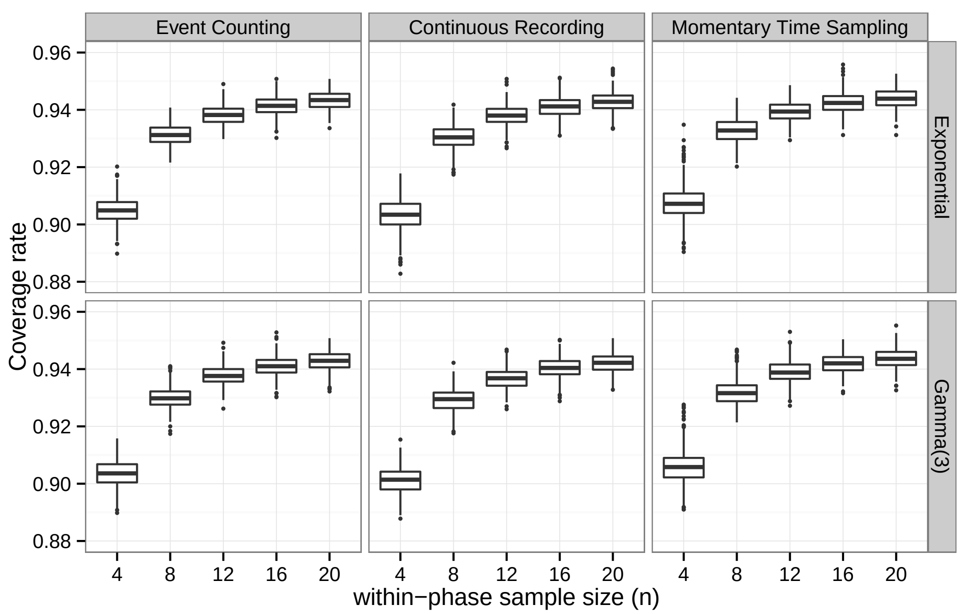


(b) RMSE

Figure 1: Bias and RMSE of simple- and bias-corrected log-response ratio estimators for varying within-phase sample sizes. Each boxplot displays the range across varying levels of ϕ_0, ζ_0, and ψ. Columns of the lattice correspond to data from different observation procedures. Rows of the lattice correspond to different parametric forms for the event duration and interim time distributions.



(a) Relative bias of variance estimators.



(b) 95% CI coverage rate.

Figure 2: Relative bias of variance estimators and CI coverage for log-response ratio. Each boxplot displays the range across varying levels of ϕ_0, ζ_0, and ψ.

Figure 2a displays the relative bias of the variance estimators corresponding to the bias-corrected log-response ratio estimators. Across data types and ARP distributions, the variance estimators are close to unbiased, having expectations that are largely within 5% of the true variance when $n \geq 8$. At the smallest within-phase sample size ($n = 4$), the variance estimators tend to slightly overstate the actual variance.

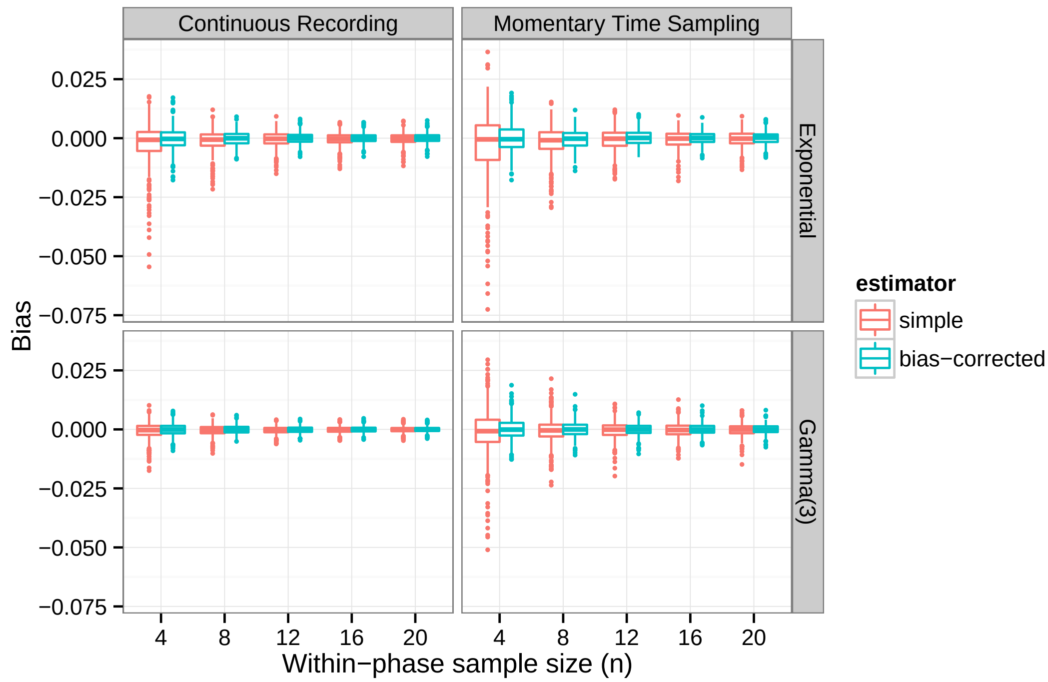
Figure 2b displays the range of actual coverage rates of 95% CIs for the log-response ratio. The CIs have less than nominal coverage, as might be expected given the use of standard normal critical values. At the smallest sample size considered, the coverage rates are sometimes lower than 90%. For $n \geq 8$, the coverage is at least 92% across all combinations of parameter values; the median coverage rate is 93% for $n = 8$ and 94% for $n = 12$. A simple small-sample correction (such as using critical values based on a t_{2n-2} distribution rather than a standard normal distribution) would be expected to improve the coverage rates of the proposed CIs.

3 Log-prevalence odds ratio estimators based on continuous recording or momentary time sampling

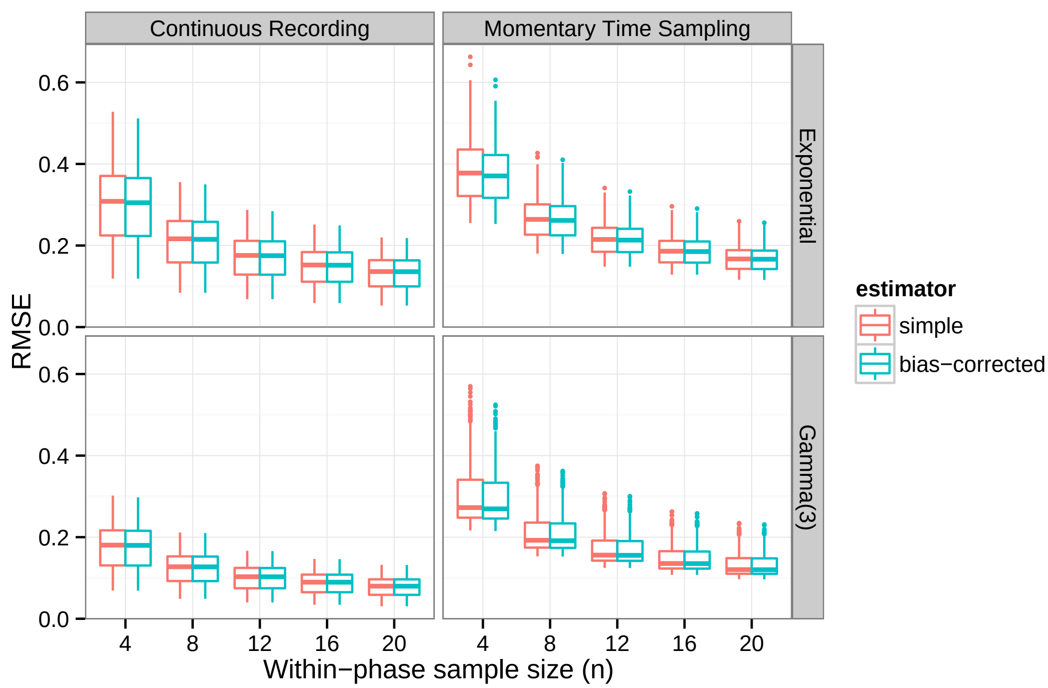
This section presents the results for the log-prevalence odds ratio estimators based on continuous recording or momentary time sampling data. Figure 3 displays the bias and RMSE of the estimators, and is constructed in the same fashion as Figure 1. Just as with the log-response ratio estimators, the bias-corrected log-prevalence odds ratio estimators have less bias than the simple moment estimators, while maintaining comparable RMSE. Even at the smallest sample size considered, the bias is less than 0.02 in absolute magnitude for both continuous recording and momentary time sampling data. By comparison, the simple moment estimator has maximum bias of 0.05 when based on continuous recording with $n = 4$ and 0.07 when based on momentary time sampling with $n = 4$.

Figure 4a displays the relative bias of the variance estimators corresponding to the bias-corrected log-prevalence odds ratio estimators. Across data types and ARP distributions, the variance estimators have small biases. For sample sizes of $n \geq 8$, the relative bias is always within 10% of the true variance, and is usually much closer. For small sample sizes, the estimators tend to slightly overstate the variance, particularly when based on momentary time sampling data.

Figure 4b displays the range of actual coverage rates of 95% CIs for the log-prevalence odds ratio. Justs with the log-response ratio, these CIs have less than nominal coverage for most combinations of parameters. At the smallest sample size considered, the coverage rates are sometimes lower than 90%, but improve to near-nominal levels as sample size increases. For $n \geq 8$, the coverage is at least 92% across all combinations of parameter values.

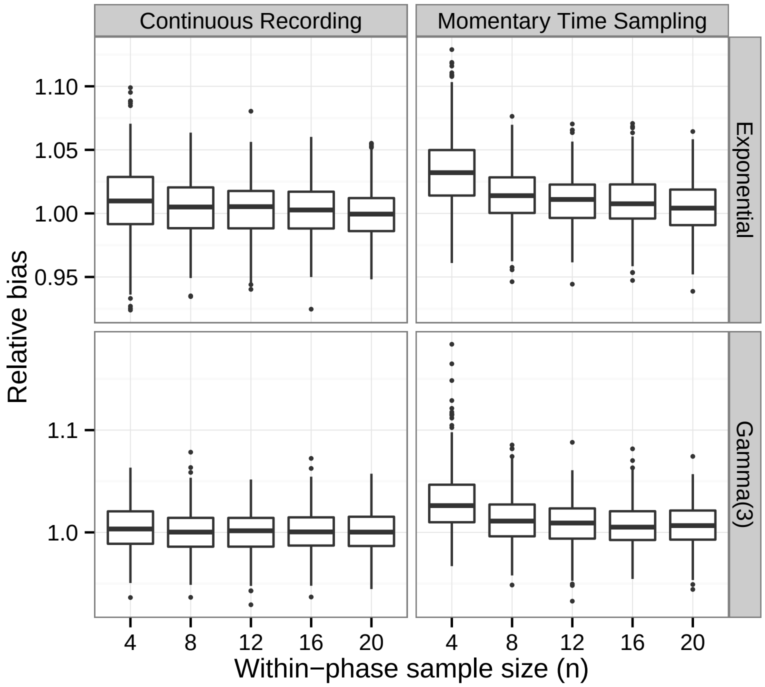


(a) Bias

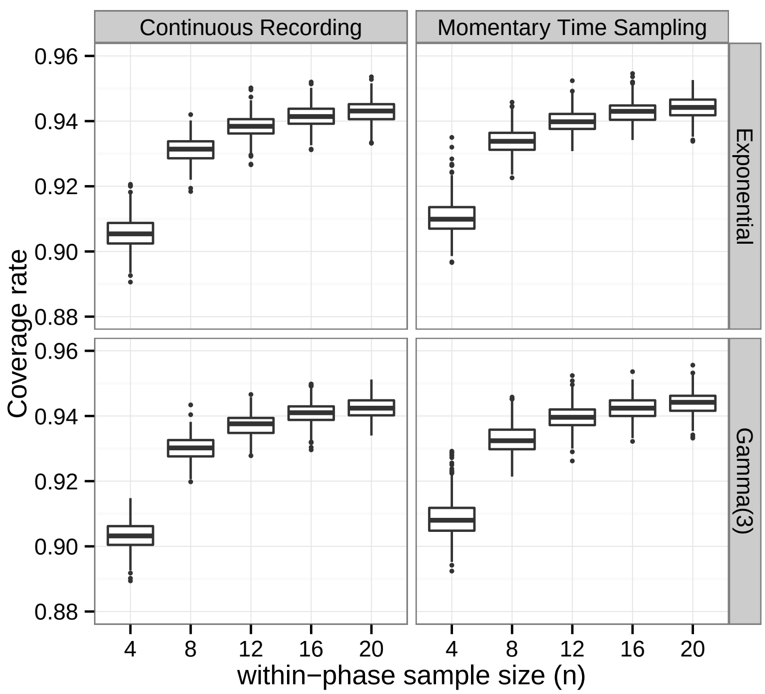


(b) RMSE

Figure 3: Bias and RMSE of simple- and bias-corrected log-prevalence odds ratio estimators based on continuous recording or momentary time sampling data, for varying within-phase sample sizes.

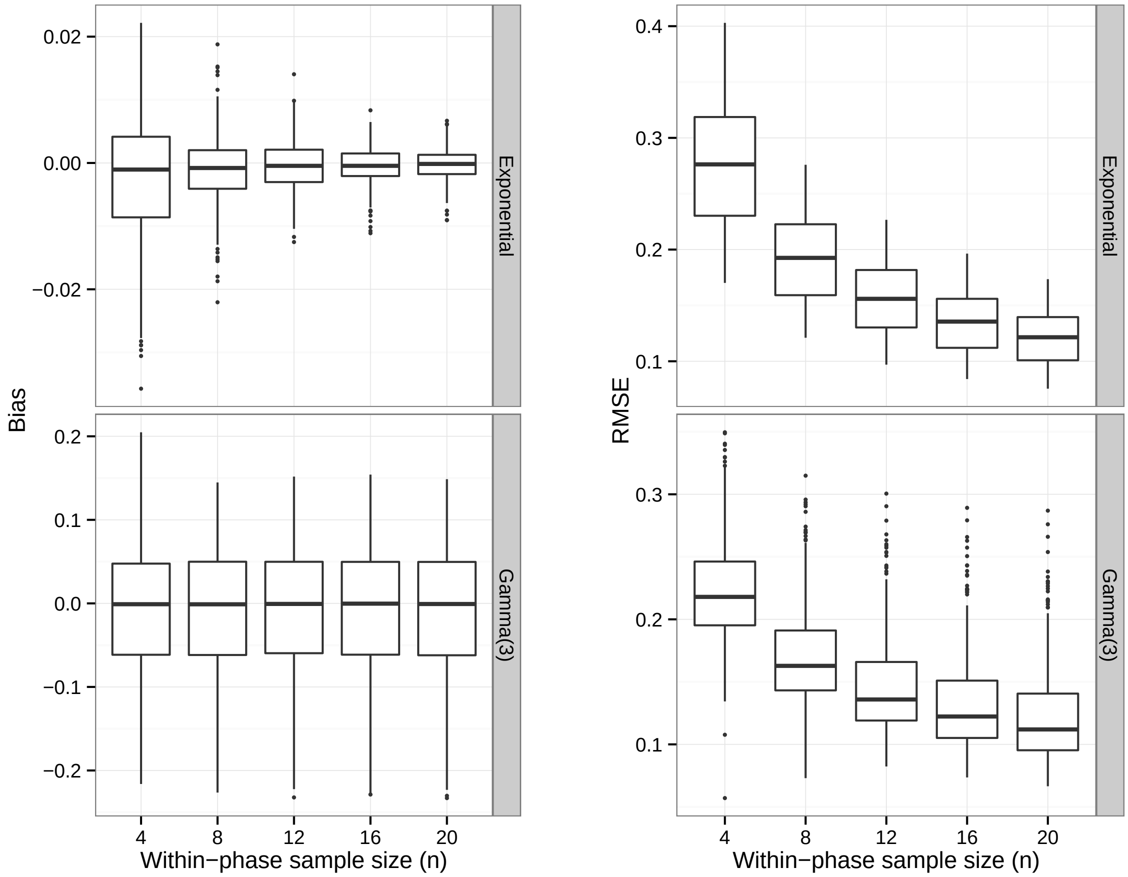


(a) Relative bias of variance estimators.



(b) 95% CI coverage rate.

Figure 4: Relative bias of variance estimators and CI coverage for the log-prevalence odds ratio, when based on continuous recording or momentary time sampling data.



(a) Bias

(b) RMSE

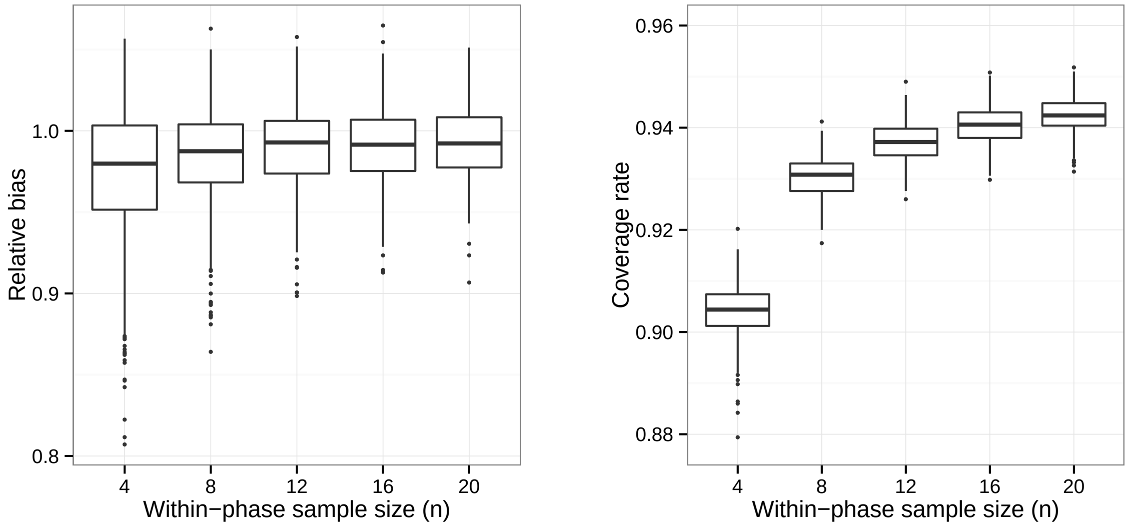
Figure 5: Bias and RMSE of log-prevalence odds ratio estimators based on partial interval recording data, for varying within-phase sample sizes.

4 Log-prevalence odds ratio estimator based on partial interval recording

This section presents the results for the log-prevalence odds ratio estimator based on partial interval recording data. The proposed estimator is premised on two assumptions: that the average event duration does not change across treatment conditions ($\mu_0 = \mu_1$) and that the interim times follow exponential distributions. The first assumption holds given the simulation design; the latter assumption does not.

Partial interval recording data is upwardly biased relative to the prevalence of a behavior. At high levels of prevalence or incidence, it therefore displays ceiling effects that limit its sensitivity to change. To address this feature, I limit the simulation parameter combinations such that the expectation of the sample mean is less than 0.98 under each treatment condition. This reduces the number of unique parameter combinations of ϕ_0, ζ_0, and ψ from 450 to 345.

Figure 5 displays the bias and RMSE of the proposed moment estimator when the ARP



(a) Relative bias of variance estimator.

(b) 95% CI coverage rate.

Figure 6: Relative bias of variance estimator and CI coverage for the log-prevalence odds ratio, when based on partial interval recording data.

uses exponential distributions or gamma distributions. When the interim time distribution is exponential, the estimator has only small biases. Even at the smallest sample size considered, the bias is less than 0.04 in absolute magnitude; for $n \geq 8$, the absolute bias is always less than 0.02. However, the estimator can be badly biased if the interim time distributions are not exponential, as in the lower panel of Figure 5a. Although the estimator actually displays lower RMSE when the interim times are gamma-distributed than when they are exponentially distributed, this is merely an artifact of the simulation design.

Figure 6a displays the relative bias of the variance estimator based on partial interval recording data, under the condition that interim times are exponentially distributed. The variance estimator tends to somewhat understate the true variance of the effect size estimator, particularly when the within-phase sample size is very small. For $n \geq 12$, the variance estimator has bias of less than 10%.

Figure 6b displays the range of actual coverage rates of 95% CIs for the log-prevalence odds ratio. The coverage rates are very similar to those of CIs based on other types of data. They are generally less than nominal, but improve to near-nominal levels as sample size increases. For $n \geq 8$, the coverage is at least 92% across all combinations of parameter values under consideration.