A moment estimator for ϕ and ζ: Simulation results

June 19, 2014

This appendix describes the results of a simulation study examining the performance of the moment estimators for ϕ and ζ proposed in the main text. We examined the following properties of the estimators:

- the bias of the moment estimators, in both absolute and relative terms;
- the relative efficiency of the moment estimator for ϕ compared to the "naïve" estimator of ϕ; and
- the coverage levels of parametric bootstrap confidence intervals for ϕ and ζ.

The following sections describe the design of the simulations and present the results. The final section provides and summary and guidance on the use of the moment estimators.

1 Simulation design

We studied the performance of the moment estimators for a single sample of measurements, where each measurement is generated by applying PIR to a behavior stream that follows an equilibrium Alternating Poisson Process. The Alternating Poisson Process is a specific instance of the Alternating Renewal Process in which the interim times and event durations are both exponentially distributed. We conducted two simulations, one to study the bias and efficiency of the moment estimators and a second to study the coverage of the bootstrap confidence intervals. Both simulations used fully crossed factorial designs with four factors: the prevalence (ϕ) and incidence (ζ) of the behavior stream, the length of the observation session as indicated by the number of intervals (K) in the session, and the number of observation sessions in the sample (n). Throughout, we assumed that there was no rest period in between active intervals and we fixed the length of the active interval (c) to 1. The behavior's incidence is therefore measured in terms of behaviors per active interval (i.e., $\zeta = 1/5$ corresponds to an average of one behavior per five intervals).

1.1 Initial simulation

The third column of Table 1 summarizes the factor levels used in the initial simulation. We varied ϕ over nearly its entire range, using 50 values from 0.01 to 0.99 that were equally spaced in the logit of ϕ. We also included the value of 0, in order to cover the full range of

Table 1: Simulation Design

Parameter	Definition	Initial Simulation	Bootstrap simulation
ϕ	Prevalence	0; 0.01 to 0.99 in 50 steps	0.01 to 0.50 in 20 steps
ζ	Incidence	0.00001; 0.005 to 0.90 in 50 steps; 1, 2, 3	0.005 to 0.50 in 20 steps
K	Intervals	20, 40, 60, 80	20, 40, 60, 80
n	Measurements	8, 12, 16, 20, 24, 48	8, 12, 16, 20, 24, 48

possible estimates when simulating the bootstrap confidence intervals. We varied ζ over a range wider than is likely to be encountered in practice, taking 50 values from 0.005 to 0.90 that were equally spaced in the natural logarithm of ζ. We also included values of 0.00001, 1, 2, and 3 in order to cover outlying estimates of zeta when simulating the bootstrap confidence intervals. The number of intervals ranges from $K = 20$ to $K = 80$ in steps of 20. Because the active interval length was fixed, larger K corresponded to longer observation sessions. We varied the number of observation sessions per sample from 8 to 24 in steps of 4 and also included 48 observations per sample. We included $n = 48$ in order to examine the performance of the moment estimator in very large samples, but note this sample size is much larger than would typically be found in a single-case design.

We simulated the PIR measurements using the ARPobservation package for the R statistical computing environment. For each combination of parameters, we generated 5000 samples, each consisting of n PIR summary measurements. Each PIR measurement was generated by simulating a behavior stream of length K from an Alternating Poisson Process with mean event duration $\mu = \phi/\zeta$ and mean interim time $\lambda = (1-\phi)/\zeta$, then applying PIR with K intervals. For each simulated sample, we then calculated the mean and variance of the measurements and solved equations (17) and (18) in the main text to obtain moment estimators $\hat{\phi}$ and $\hat{\zeta}$.

1.2 Bootstrap simulation

We also studied the performance of parametric percentile bootstrap confidence intervals for prevalence and incidence. To moderate the dimension of this simulation, we excluded extreme levels of ϕ and ζ, where the moment estimators would not be expected to perform well. The fourth column of Table 1 summarizes the design of the bootstrap simulation. We varied ϕ from 0.01 to 0.50, taking 20 equal steps in the logit of ϕ. We varied ζ from 0.05 to 0.50, taking 20 equal steps in the natural logarithm of ζ. The levels of K and n matched those from the initial simulation.

Given the large number of conditions in the design (more than 60,000), the amount of computing time necessary to simulate bootstrap samples for each replicate of the simulation would have been prohibitive. Instead, we estimated the bootstrap confidence intervals based on the initial simulation results. In the initial simulations, we calculated the 2.5% and 97.5% percentiles across the 5,000 simulated values of the moment estimators for each combination of parameter levels. In the bootstrap simulations, we began by simulating 5,000 values of the moment estimators for each combination of factor levels, just as in the previous simulation.

We then used bilinear interpolation of the percentiles from the initial simulation to predict the 2.5% and 97.5% percentiles of the bootstrap distribution of the moment estimators, simulated from a model with parameter values set to the estimates $\hat{\phi}$ and $\hat{\zeta}$.

R code for reproducing the simulations is available upon request.

2 Bias of the moment estimators

We assessed the bias of the moment estimators in both absolute and relative terms. In absolute terms, the bias of the prevalence estimator is simply $\mathrm{E}\left(\hat{\phi} - \phi\right)$ and that of the incidence estimator is $\mathrm{E}\left(\hat{\zeta} - \zeta\right)$. We calculated the relative bias of the prevalence estimator in terms of odds, as $\mathrm{E}\left[\hat{\phi}/(1-\hat{\phi})\right] / [\phi/(1-\phi)] - 1$, and the relative bias of the incidence estimator as $\mathrm{E}\left(\hat{\zeta}\right)/\zeta - 1$. Lacking alternative methods of estimation, an analyst may consider estimators with moderate biases to be tolerable. We therefore adopted a liberal criteria for relative bias, taking relative bias of between $\pm 5\%$ to be "approximately unbiased."

Figure 1 depicts the raw bias of the moment estimator for prevalence. The portion of the parameter space with low bias (-.02 to .02) increases slightly n increases, but the value of K has relatively little impact on bias. At $n = 16$, the moment estimator produce estimates with low bias when $\phi < .15$ and $\zeta < 0.35$, or when $\phi < .75$ and $\zeta < .10$. As the sample size increases, these ranges expand slightly.

Figure 2 presents results for the relative bias of the estimate of prevalence odds. For any substantial portion of the parameter space to be unbiased a sample size of at least $n = 16$ is required. When $n = 16$ and $K = 60$, the moment estimator produces approximately unbiased estimates of prevalence only when $.25 < \phi < .75$ and incidence is relatively low, $\zeta < 0.10$. This range narrows slightly when the number of intervals is smaller and expands slightly for $K = 80$. As n increases, the range where approximately unbiased estimates are obtained expands to include smaller and larger values of ϕ as well as larger values of ζ.

Figure 3 depicts the raw bias of the moment estimator for incidence. The portion of the parameter space with low bias increases as n increases but the value of K has no systematic impact. At $n = 16$, the moment estimator produces estimates very low bias when $\zeta < 0.25$ and $\phi < .12$ or when $\zeta < 0.10$ and $\phi < .50$. As the value of n increases, these areas widen to includes higher values of ϕ and ζ.

Figure 4 presents results for the relative bias of the estimate of incidence. Compared to the results for the relative bias of $\hat{\phi}$, even larger sample sizes are required to obtain approximately unbiased estimates of ζ. A minimum sample size of $n = 20$ with $K = 40$ intervals or more is necessary for $\hat{\zeta}$ to be approximately unbiased over a substantial portion of the parameter space. When $n = 20$ and $K \geq 40$, the moment estimator produces approximately unbiased estimates in the range $0.03 < \zeta < 0.08$ and $\phi < .50$, as well as for $\zeta \leq 0.03$ and $\phi < .10$. When $n = 24$ and $K \geq 40$, the moment estimator produces approximately unbiased estimates at $0.025 < \zeta < 0.10$ and $\phi < .65$ as well as $\zeta < 0.25$ and $\phi < .15$. Higher values of K tend to widen the portion of the parameter space where the moment estimator produces approximately unbiased estimates at smaller values of ζ, while slightly reducing the area at higher values of ζ.

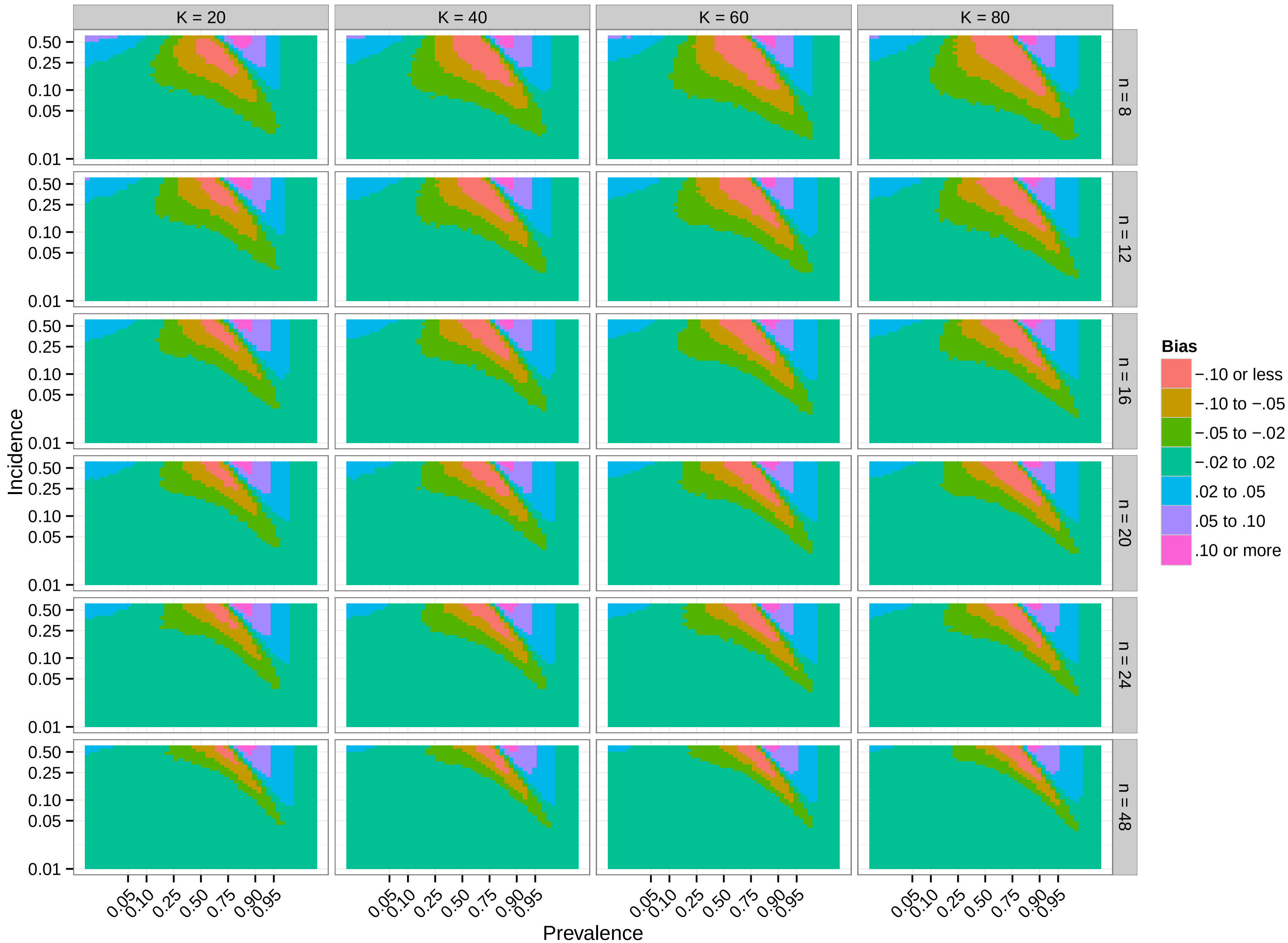

Figure 1: Bias of the estimate of prevalence. Each plot shows the bias across a range of values for ϕ and ζ. Columns of the lattice correspond to different values of K. Rows of the lattice correspond to different values of n.

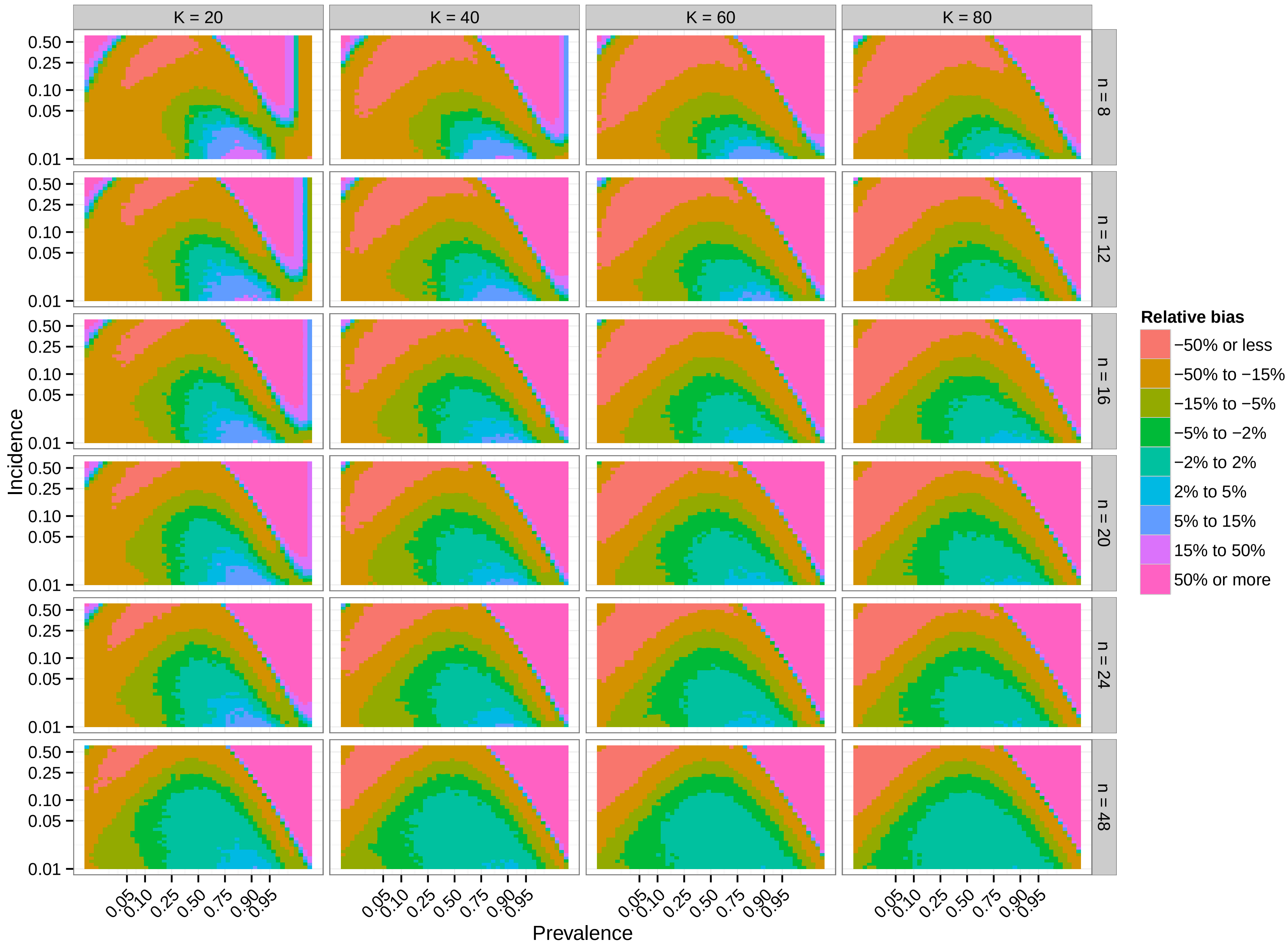

Figure 2: Relative bias of the prevalence odds. Each plot shows the bias across a range of values for ϕ and ζ. Columns of the lattice correspond to different values of K. Rows of the lattice correspond to different values of n.

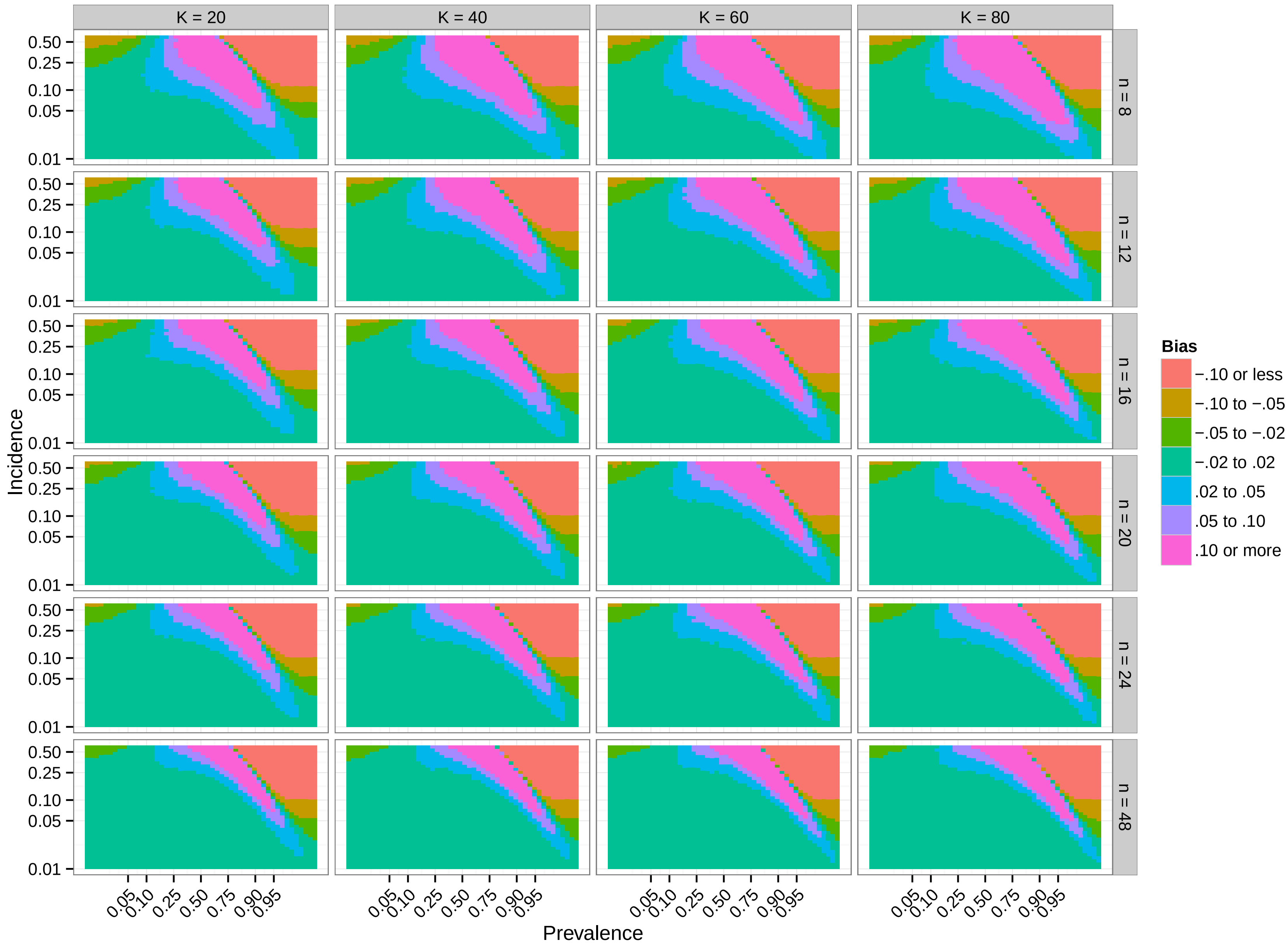

Figure 3: Bias of the estimate of incidence. Each plot shows the bias across a range of values for ϕ and ζ. Columns of the lattice correspond to different values of K. Rows of the lattice correspond to different values of n.

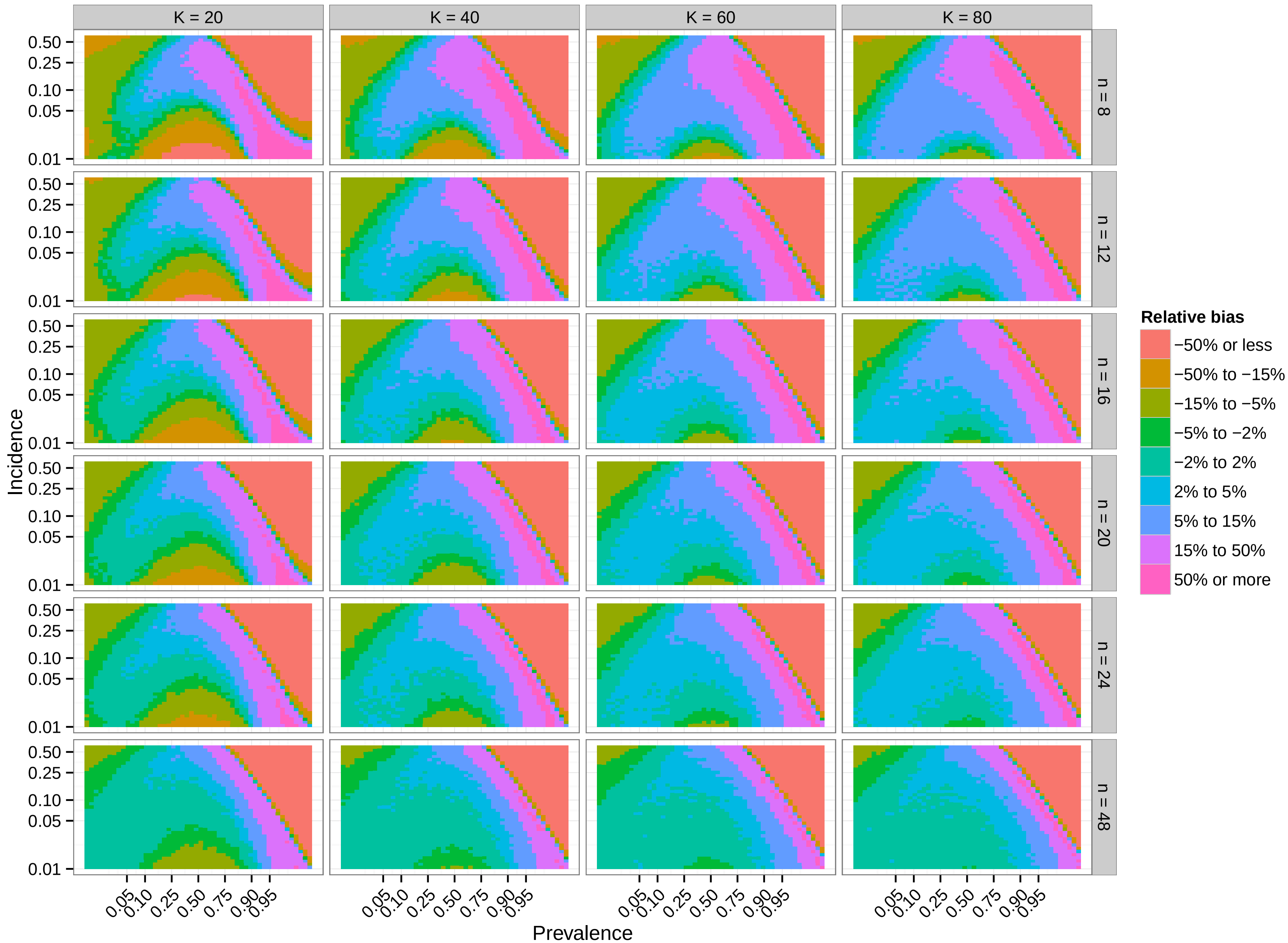

Figure 4: Relative bias of incidence. Each plot shows the bias across a range of values for ϕ and ζ. Columns of the lattice correspond to different values of K. Rows of the lattice correspond to different values of n.

Taken together, these results suggest that the moment estimators perform reasonably in terms of absolute bias, but less well if assessed in terms of relative bias. The relative measures are likely to be more important given that the main text focuses on ratio measures of change. In relative terms, the moment estimators require relatively large sample sizes in order to obtain reasonable estimates of either prevalence or incidence over a reasonable portion of the parameter space. Furthermore, the subset of the parameter space for which approximately unbiased estimates of *both* ϕ and ζ can be obtained is fairly small. When $n = 20$ and $K \geq 40$, the moment estimator produces approximately unbiased estimates for both parameters at $.25 < \phi < .75$ and $0.025 < \zeta < 0.10$. When $n = 24$ and $K \geq 40$, the moment estimator produces approximately unbiased estimates for both parameters at $.20 < \phi < .60$ and $0.025 < \zeta < 0.10$.

3 Moment estimator vs. the "naïve" estimate of prevalence

Although the moment estimators perform well in terms of relative bias for only a portion of the parameter space, there are very few alternatives for estimating prevalence available to researchers. One method considered in the main text is to simply treat the mean of the summary measurements as an estimate of prevalence. As a minimal standard, an estimator of ϕ should be more accurate than this "naïve" estimator if it is to be recommended for use. We therefore compared the efficiency of the naïve estimator to that of the moment estimator for prevalence. We calculated relative efficiency as $\sqrt{\mathrm{E}\left[(Y^P - \phi)^2\right] / \mathrm{E}\left[\left(\hat{\phi} - \phi\right)^2\right]}$, where Y^P is a PIR measurement. Relative efficiency greater than one favors the moment estimator while relative efficiency below one favor the naïve estimator.

Figure 5 presents the results for the relative efficiency of the moment estimator versus to the naïve estimator. Even at relatively modest sample sizes of $n = 12$, the moment estimator is nearly always more efficient than the naïve estimator when $\phi < .50$. In this space, the moment estimator performs at least 20% better than the naïve estimate of ϕ for all values of $\zeta > 0.10$; this range widens to include smaller values of ζ as the sample size increases. When $n = 12$ or more, $\phi < .25$, and $\zeta > 0.10$, the moment estimator is is nearly always 80% more efficient than the naïve estimator. For the larger sample size of $n = 24$, this range widens to include $\phi < .45$. Larger values of K increase the area of the parameter space where the moment estimator performs well. On this basis, we suggest that the minimum number of intervals that should be used with PIR is $K = 40$. In summary, although the moment estimator can be somewhat biased, particularly for small ϕ or large ζ, it still provides a substantial improvement over the naïve estimator across a wide range of true parameter values.

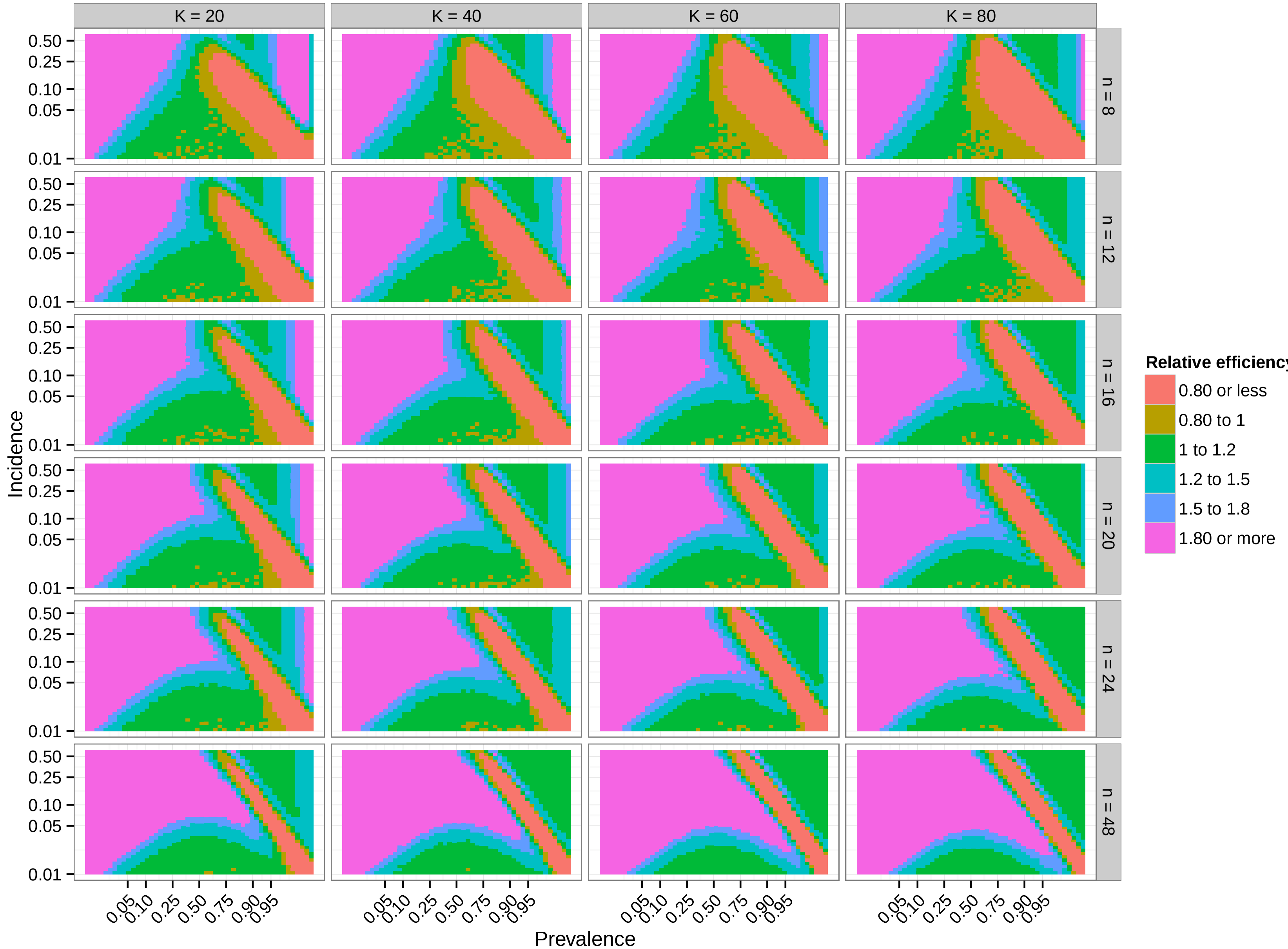

Figure 5: Relative efficiency of $\hat{\phi}$ versus Y^P. Each plot shows the efficiency across a range of values for ϕ and ζ. Columns of the lattice correspond to different values of K. Rows of the lattice correspond to different values of n. Higher values of the proportion favor the moment estimator over the na"ive estimator.

4 Bootstrap confidence interval performance

Finally, we assessed the actual coverage rates of nominal 95% confidence intervals for prevalence and incidence. Just as in our analysis of relative bias, we adopt liberal criteria for acceptable coverage due to the lack of viable alternative estimators. Specifically, we interpret coverage rates of 94-96% as accurate, 92.5%-97.5% as acceptable, and rates outside of this range as poor. Note that the coverage rates are invariant to monotonic transformation, and so results are identical for the log-odds of prevalence and the log of incidence.

Figure 6 depicts the actual coverage rates of the confidence interval for ϕ. The areas where the bootstrap produces acceptable or better coverage nearly mirrors that where the moment estimator produces approximately unbiased estimates, with slightly reduced values for ζ. At $n = 16$ and $K = 60$, the bootstrap confidence intervals have acceptable acceptable coverage when $.25 < \phi < .50$ and $\zeta < 0.75$. As n increases, this range widens to include smaller values of ϕ and larger values of ζ. Larger values of K reduce the area where the bootstrap produces acceptable coverage, but this may simply be a reflection of the area shifting to slightly lower values of ζ and slightly higher values of ϕ.

Figure 7 depicts the actual coverage rates of the confidence interval for ζ. At least acceptable coverage tends to mirror the area where the moment estimator produces approximately unbiased estimates. For $n = 20$ and $K \geq 40$, the bootstrap produces acceptable or better coverage when $0.03 < \zeta < 0.08$ and $\phi < .50$. For $n = 24$ and $K \geq 40$, the bootstrap produces acceptable or better coverage when $0.025 < \zeta < 0.5$ and $\phi < .50$ as well as $\zeta < 0.25$ and $\phi < .15$. In addition, the coverage is poor when ϕ is near zero and ζ is either low or high. Beyond that, there are non-uniform effects of K that tend to alter the coverage and some areas within the described parameter space where the coverage is less than acceptable. On the whole, the coverage rates of the bootstrap confidence interval for ζ tend to be non-uniform and rather difficult to characterize.

5 Discussion

In practice, we expect that observers will use partial interval recording to measure behavior for only certain classes of behavior, such as those for which prevalence is less than .50 and events occur neither too frequently nor too infrequently, perhaps in the range of $0.05 < \zeta < 0.50$. When prevalence is higher than this, whole interval recording may be a more appropriate method of direct observation. If the value of incidence is outside the suggested range, the behaviors of interest happen either more frequently than once every 2 intervals on average or less frequently than once every 20 intervals on average. In either case, the length of the active interval may need to be tuned to more appropriate match the frequency of the behavior. Our discussion on use of the moment estimators is limited to these areas of the parameter space.

When ϕ is the primary quantity of interest, a sample size of at least $n = 20$ with at least $K = 40$ intervals per session is needed to obtain an approximately unbiased estimate and a confidence interval with an acceptable level of coverage over a reasonably large subset of the parameter space. When $n = 20$, this area of the parameter space is $\phi > .25$ and $\zeta < 0.10$. When $n = 24$, this area of the parameter space grows to include $\phi > .20$ and $\zeta < 0.10$.

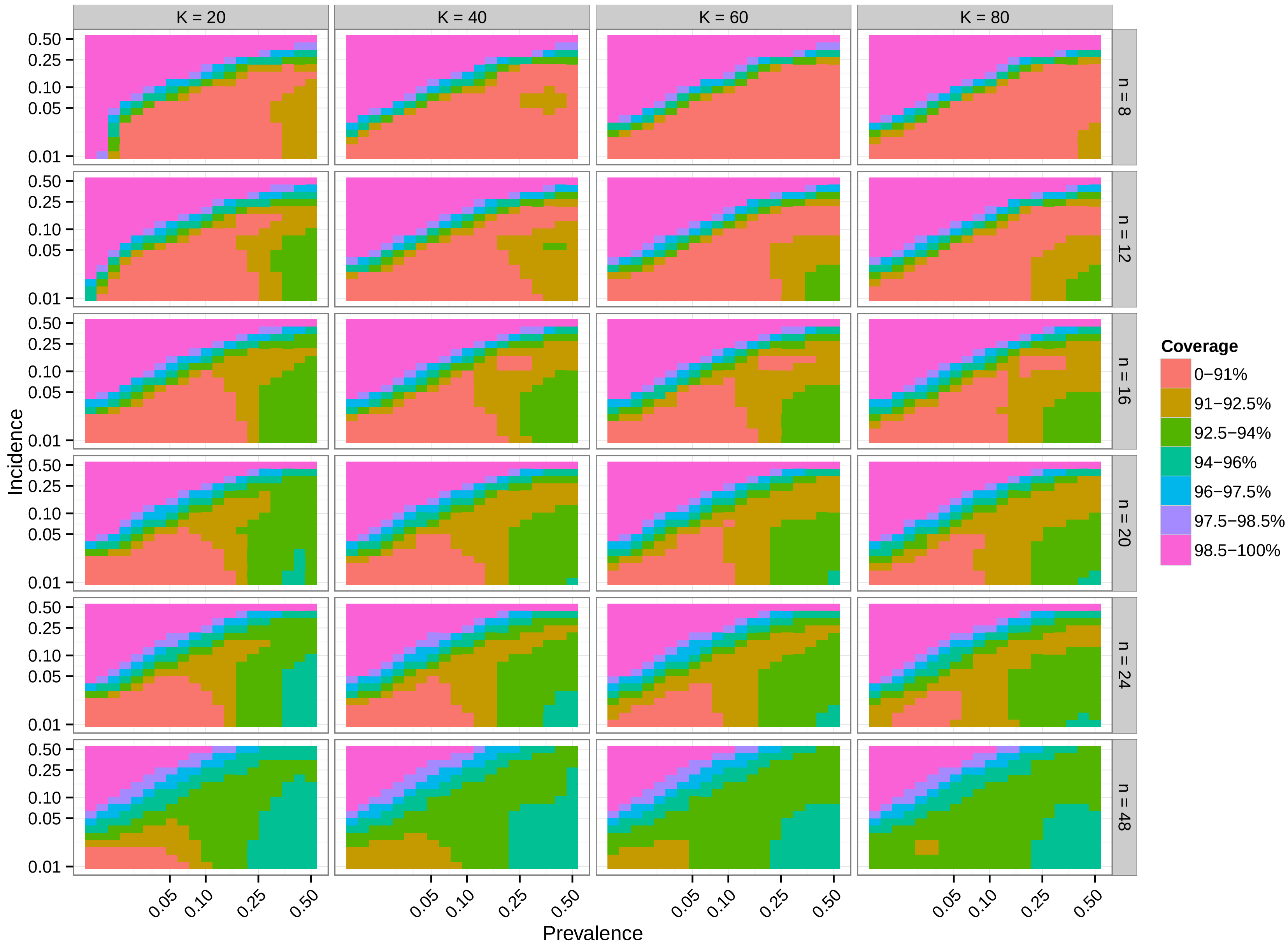

Figure 6: Performance of the bootstrap confidence interval for ϕ. Each plot shows the actual coverage rate of a nominal 95% confidence interval across a range of values for ϕ and ζ. Columns of the lattice correspond to different values of K. Rows of the lattice correspond to different values of n.

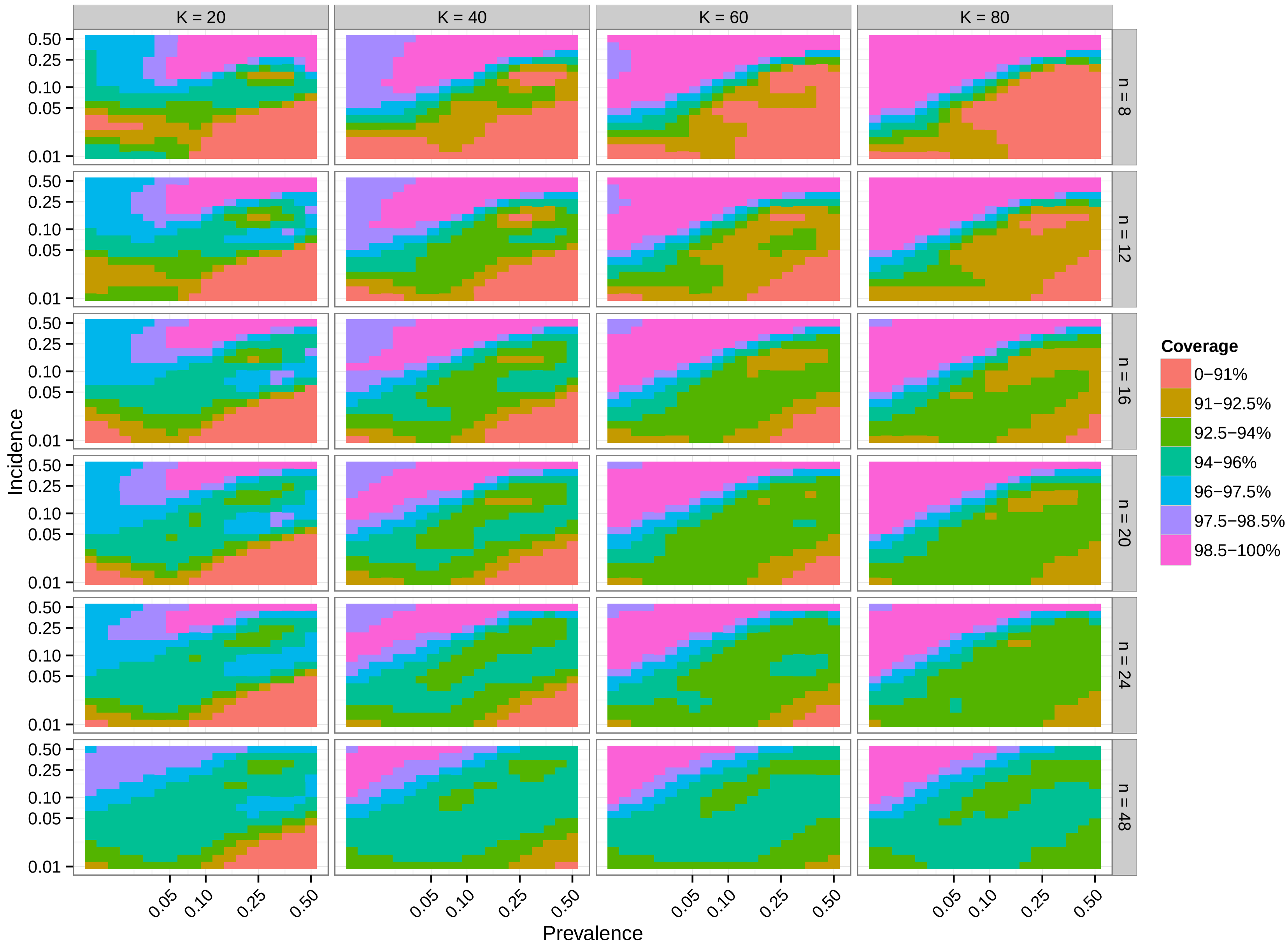

Figure 7: Performance of the bootstrap confidence interval for ζ. Each plot shows the actual coverage rate of a nominal 95% confidence interval across a range of values for ϕ and ζ. Columns of the lattice correspond to different values of K. Rows of the lattice correspond to different values of n.

However, note that the moment estimator for prevalence outperforms the naïve estimator over a large portion of the parameter space when samples sizes are as small as $n = 12$.

When ζ is the primary parameter of interest, a sample size of at least $n = 20$ with at least $K = 40$ intervals per session is needed to obtain good point estimates and confidence intervals over some portion of the parameter space. However, this subset includes only a fairly narrow range of incidence values, $0.03 < \zeta < 0.08$ when $\phi < .50$, as well as $\zeta \leq 0.03$ when $\phi < .10$. When $n = 24$ and $K \geq 40$, the moment estimator produces approximately unbiased estimates at $0.025 < \zeta < 0.10$ and $\phi < .50$ as well as $\zeta < 0.25$ and $\phi < .15$. Higher values of K tend to widen the portion of the parameter space where the moment estimator produces approximately unbiased estimates at smaller values of ζ, while slightly reducing the area at higher values of ζ. In addition, the coverage is poor at low and high values of ζ and low values of ϕ.

There are only small parts of the parameter space where both moment estimators are close to unbiased (even by our liberal criterion). In general, the areas of the parameter space where good estimates and confidence intervals can be obtained for prevalence are areas where they cannot be obtained for incidence, and vice versa. Thus, even though the moment estimators are the only one of the four methods discussed in the main text that provide information about both prevalence and incidence, their use should still be restricted to contexts in which only one or the other quantity is of primary interest. On the whole, the simulation results presented here indicate that the moment estimators should be used for tentative, exploratory purposes, but should not be treated as definitive unless based on very large samples. Analysts might also consider using bootstrap bias correction for the point estimates or bias-corrected, accelerated bootstrap confidence intervals. These elaborations on the moment estimators may have improved operating characteristics, though this remains to be investigated in further simulations.