Title: Assessment of bias and precision among simple closed population mark-recapture estimators

Author: Kyle Dettloff[1]

[1]National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami, FL 33149, USA
Email: kyle.dettloff@noaa.gov

Keywords: bias correction; closed population; mark-recapture; Monte Carlo simulation; Schumacher-Eschmeyer

Abstract

Mark-recapture methods have been heavily studied and employed in fisheries and other wildlife sciences over the past century to approximate population sizes for animal species of interest. This paper focuses on the comparative statistical performance through simulation of common closed population mark-recapture estimators, including those of Lincoln-Petersen, Chapman, Chao, Schnabel, and Schumacher-Eschmeyer. A new bias-adjusted version of the Schumacher-Eschmeyer estimator is proposed and is shown to exhibit superior performance at small sample sizes in comparison to the original estimator. Simulation results indicate that Chapman's method outperforms all other two-visit methods and that bias-adjusted versions of Schnabel and Schumacher-Eschmeyer differ slightly depending on bias or precision, but both perform well. Minimum sample sizes such that resulting estimates are approximately unbiased are proposed to advise practitioners on the most appropriate use of these estimators for simple closed population mark-recapture data.

1. Introduction

Reliable population size estimates are fundamental to understanding the ecology and conservation needs of animal populations. Procedures to derive these estimates differ fundamentally from that of traditional statistical survey methods in that there is no predetermined frame of individuals from which an investigator can sample. The first solution to this problem in the ecological literature dates back to 1896 when fishery scientist G.C.J. Petersen invented a brass tag to attach to fish, recognizing that population size could be estimated when sufficient numbers of these tags had been recaptured. The same method was later derived by ornithologist Frederick Lincoln (1930) to estimate the size of North American waterfowl populations using banding data. These approaches are collectively known as the Lincoln-Petersen method, the simplest and most well-known metric to estimate the size of a closed population, meaning a negligible effect of births, deaths, and movement during the study period. Animals are captured, marked, and released on the first visit, and, after being allowed to mix with the full population, randomly sampled without replacement on a second visit, noting the number of marked individuals present from the first sample in the second sample. This two-visit method was subsequently expanded to apply to k visits by Schnabel (1938) and further modified by Schumacher and Eschmeyer (1943) to be formulated as a linear regression. More recently, Brittain and Böhning (2009) used empirical data to compare the performance of the Chao (1987, 1989) estimator, which relaxes the assumption of independence between visits, with the more common Chapman (1951) estimator, a bias-adjusted version of the two-visit Lincoln-Petersen.

Though more sophisticated model-based methods have been proposed to relax the assumptions of independence between captures (Otis et al. 1978; Huggins 1989, 1991) and closed populations (Seber 1982), this paper focuses strictly on the theoretical performance of simple closed population estimators for which identification of individuals is not required. While sampling complexities such as variation in capture probabilities and uncertain sampling area are known to impact bias and variance in mark-recapture studies (Amstrup et al. 2005, McNair et al. 2018), this study aims to assess the validity of these existing metrics under conditions in which all assumptions are known to be met, allowing one to select a superior metric from a purely statistical basis, all other sampling considerations equal. For a study evaluating the performance of various mark-recapture models on field data with known reference population sizes, including those allowing individual heterogeneity in capture probabilities, see Grimm et al. (2014).

These simple estimators remain relevant within fisheries and conservation biology, returning over 1,650 articles since 2016 with reference to closed population mark-recapture in a Google Scholar search. Of these, 69 contain specific reference to the Schumacher-Eschmeyer estimator, which has been employed to estimate population size in a wide variety of fisheries, including endangered Atlantic sturgeon (Kahn et al. 2014, 2019; Hale et al. 2016), lake trout (Hansen et al. 2008), walleye (Spencer et al. 2002), American lobster (Rowe 2002), pirarucu (Castello 2004), and even for estimating angler counts in a creel survey (Hansen and Van Kirk 2018).

Comparing these historically important and still common estimators within a consistent framework and across a range of simulations offers clarity as to which methods are most appropriate under known sampling scenarios, with the equal benefit of indicating which estimators should be confidently discontinued from use in favor of superior alternatives. Consequently, these simulations are also able to suggest approximate minimum sample sizes needed to generate reliable estimates of population size among the best performing methods.

2. Materials and methods

A collection of common closed population mark-recapture estimators prevalent in the ecological literature were selected and compared in performance through simulation. Comparisons were conducted for both two-visit methods and multi-visit methods. All methods analyzed assume:

1. Closed populations (no change in population size between sampling events)
2. Independence between visits (marking does not influence the probability of recapture)
3. Independence between individuals (complete mixing occurs between sampling events)
4. Individuals are sampled without replacement
5. No marks are lost between sampling events

A brief overview of the various estimators is provided followed by the simulation study methodology.

2.1 Overview of estimators

2.1.1 Lincoln-Petersen estimator (2 visits)

Also known as the Lincoln Index, Lincoln-Petersen (Petersen 1896, Lincoln 1930) is the simplest and most intuitive of the estimators, and is the maximum likelihood estimator (MLE) of *N*. It is given by (Eq. 1),

$$\hat{N} = \frac{Mn}{m} \tag{1}$$

where:

M = number of animals marked on the first visit;
n = total number of animals captured on the second visit;
m = number of marked animals recaptured on the second visit.

This equation implies that the proportion of marked individuals captured in the second sample (m/n) is equal to the proportion of the total population (N) that has been marked in the first sample (M/N). Lincoln-Petersen forms the basis for all estimators that follow.

2.1.2 Chapman estimator (2 visits)

At small sample sizes, (Eq. 1) produces biased estimates of population size (Chapman 1951). Several modifications have been suggested to reduce this bias, the most common being the Chapman (1951) estimator, given by (Eq. 2).

$$\widehat{N} = \frac{(M+1)(n+1)}{(m+1)} - 1 \tag{2}$$

This estimator, based on the hypergeometric distribution, possesses finite moments, as the denominator cannot be zero, which is possible in Lincoln-Petersen when $M + n < N$. A method to obtain robust confidence intervals around Chapman estimates of population size is provided by Sadinle (2009).

2.1.3 Chao estimator (2 visits)

Chao (1987, 1989) proposed an estimator that relaxes the assumption of independence in capture probability between visits. Brittain and Böhning (2009) show that for the two-visit, equal capture probability scenario, this estimator can be formulated as (Eq. 3).

$$\widehat{N} = \frac{(M+n-2m)^2}{4m} + M + n - m \tag{3}$$

In the case of $M = n$, it can be seen that (Eq. 3) reduces to the Lincoln-Petersen estimator (Eq. 1).

2.1.4 Bayesian estimator (2 visits)

The final two-visit estimator considered is a Bayesian formulation which estimates the posterior mean based on the hypergeometric distribution, analogous to the Chapman estimator. The derivation is presented in Webster and Kemp (2013), and results in (Eq. 4).

$$\widehat{N} = \frac{(M-1)(n-1)}{(m-2)} \text{ for } m > 2 \tag{4}$$

2.1.5 Schnabel estimator ($\geq$ 2 visits)

Schnabel (1938) published the first mark-recapture estimator designed for more than two sampling visits, generalizing the traditional two-visit Lincoln-Petersen approach. The equation is formulated as a weighted average of Lincoln-Petersen estimates across the series of visits (Eq. 5),

$$\widehat{N} = \frac{\sum_k M_k n_k}{\sum_k m_k} \quad \text{for } k \geq 2 \tag{5}$$

where:

M_k = total number of marked animals in the population prior to visit k;

n_k = total number of animals captured on visit k;

m_k = total number of marked animals recaptured on visit k.

All individuals captured on each visit are marked and released into the population, with no need to distinguish between marks made on different visits. Note that the estimator becomes equivalent to Lincoln-Petersen (Eq. 1) in the case of $k = 2$ visits.

An improved small sample bias correction given in (Eq. 6) was proposed by Chapman (1952), noting that each m_k is approximately Poisson distributed with parameter $M_k n_k/N$. This correction has been recommended by multiple sources, and its performance is evaluated here alongside the original estimator.

$$\widehat{N} = \frac{\sum_k M_k n_k}{\sum_k m_k + 1} \qquad (6)$$

2.1.6 Schumacher-Eschmeyer estimator ($\geq$ 2 visits)

A similar estimator which employs the same sampling methodology as Schnabel to handle multiple recapture events was proposed by Schumacher and Eschmeyer (1943), taking the form of (Eq. 7).

$$\widehat{N} = \frac{\sum_k M_k^2 n_k}{\sum_k M_k m_k} \text{ for } k \geq 2 \qquad (7)$$

The logic behind this formula is that the proportion of marked individuals on the k^{th} visit (m_k/n_k) plotted against the number of individuals previously marked (M_k) should be linear and pass through the origin with a slope of N^{-1} under the basic assumptions outlined above. (Eq. 7) uses linear regression techniques to estimate N based on this rationale.

Following the bias-adjusted formulation of the Schnabel estimator in (Eq. 6), a similar small sample bias correction to the Schumacher-Eschmeyer equation is proposed here as (Eq. 8) based on the Chapman correction (Eq. 2).

$$\widehat{N} = \frac{\sum_{k=2}^{k}(M_k+1)^2(n_k+1)}{\sum_k M_k(m_k+1)} - 2 \qquad (8)$$

The performance of this estimator is evaluated alongside the original through simulation in the present study.

2.2 Simulation study

A series of Monte-Carlo simulations were run using the R programming language (v4.2.0; R Core Team, 2022) to evaluate the behavior of each estimator outlined above. Code is provided in Supporting Information for readers to run these simulations under their own selected parameters.

2.2.1 Estimator performance

Fixed populations of known size N were generated and randomly sampled without replacement using the *sample* function, according to the capture and marking methodology for each procedure. For each estimator, this was repeated 10,000 times for all sample sizes M and n increasing from very small to large for both cases of $M = n_{2...k}$ and $M \neq n_{2...k}$. Estimates $\widehat{N}$ and SE($\widehat{N}$) were calculated by taking the sample mean and standard deviation, respectively, from all simulations producing finite values. The resulting estimates were plotted against the geometric means of the marked individuals and the total numbers of captured individuals during subsequent visits to evaluate how each estimator's bias and variance changes with increasing sample size. Likewise, the scaled root-mean-square error (RMSE) (9) of each estimator was plotted over the range of sample sizes to visualize performance in terms of a metric that combines the effect of bias and variance.

$$RMSE_{scaled} = \frac{\sqrt{var(\widehat{N}) + bias(\widehat{N})^2}}{N} \tag{9}$$

2.2.2 Minimum sample sizes

The Chapman estimator has been shown to be exactly unbiased when the sum of the sample sizes is at least as large as the population size, or $M + n \geq N$ (Robson and Regier, 1964; Wittes 1972). A less stringent condition for the estimator to be approximately unbiased, with negative bias less than 2%, was noted by Robson and Regier (1964) in cases when:

$$\sqrt{Mn} \geq 2\sqrt{N} \tag{10}$$

That is, when the geometric mean of the marks and captures is at least twice the square root of the population size. The derivation of the degree of bias in the estimator leading to this approximate threshold is provided in Chapman (1951) using Stirling's formula with the hypergeometric distribution. Based on the results of 2.2.1, the bias-corrected version of Schnabel (Eq. 6) was analyzed across various simulation scenarios, suggesting a generalization of (Eq. 10) produces a similarly acceptable rule of thumb for the minimum sample size needed relative to the population size for resulting estimates to be approximately unbiased (Eq. 11).

$$\sqrt[k]{M \prod_{i=2}^{k} n_i} \geq \frac{2\sqrt{N}}{k-1} \tag{11}$$

Percent error, calculated according to (Eq. 12), was used to evaluate stability in bias at the proposed approximate minimum sample sizes, based on 100,000 simulations with known population sizes N ranging from 10^2 to 10^6.

$$Pct.Error = 100\left(\frac{\widehat{N} - N}{N}\right) \tag{12}$$

Similar exploration was conducted for the bias-adjusted Schumacher-Eschmeyer (Eq. 8).

3. Results

3.1 Estimator performance

While simulations were conducted for a wide range of population sizes, results below are only presented for the case of $N = 1{,}000$, as similar patterns held across all sizes evaluated.

3.1.1 Two-visit methods

Fig. 1 displays simulation results for the Lincoln-Petersen, Chapman, Chao, and Bayesian estimators in the case of $M < n$, or fewer individuals marked than recaptured. All four are biased low at very small sample sizes, which can be easily confirmed using simple arithmetic on hypothetically small M and n. This is due to the fact that either there will be no recaptures and thus no valid estimate, or if there is an unlikely recapture, the resulting estimate will be a gross underestimate of population size. While it is already well known that Chapman is an improvement over Lincoln-Petersen in terms of bias, it was seen clearly here that Lincoln-Petersen becomes biased high for an intermediate range of sample sizes after exhibiting this known low bias at very small sample sizes. Lincoln-Petersen also exhibits a higher variance than Chapman even at sample sizes where Chapman is already approximately unbiased. Likewise, the Chao and Bayesian estimators all eventually become biased high before approaching approximately unbiased states at larger sample sizes. The Chao estimator retains a high bias for much longer than any of the others in this case, while also being more variable than Lincoln-Petersen. The Bayesian estimator clearly has higher bias and variance at small sample sizes than Lincoln-Petersen or Chao. In contrast, the Chapman estimator never becomes biased high, and approaches the value of the true population size much faster than the other three estimators while also exhibiting lower variance. Similar outcomes were observed with $n < M$ and $n = M$, so simulation results are only presented for $M < n$ to illustrate a scenario where the Lincoln-Petersen and Chao estimates differ.

Fig. 2 demonstrates how the RMSE for each method changes with increasing sample size plotted on the x-axis as the geometric mean of M and n for a hypothetical population of size $N = 1{,}000$. The Bayesian estimator has the highest RMSE at low sample sizes, eventually approaching that of the other methods at larger sample sizes, while the Chao estimator retains a much higher RMSE even as the sample size increases. It is clear again that the Chapman estimator exhibits superior performance, having a much lower RMSE than the others at sample sizes even above the point at which the extreme downward-bias present in all methods begins to disappear (represented by the dashed vertical line in Fig. 2).

3.1.2 Multi-visit methods

Simulation results are presented for both the original and bias-adjusted variations of Schnabel and Schumacher-Eschmeyer with $M = n_2 = n_3$ (Fig. 3). For brevity, results are only displayed for $k = 3$ visits, as similar scaled patterns were observed under increased k. As with the two-visit methods, all estimates are biased low at very small sample sizes with negative bias decreasing as sample size increases. Similar to Lincoln-Petersen, both of the original unadjusted estimators eventually become biased high before turning approximately unbiased at larger sample sizes. Between the original versions, the unadjusted Schnabel estimator possesses noticeably lower variability and slightly less small sample bias than the unadjusted Schumacher-Eschmeyer (Figs. 3 and 4).

Performance was very similar between the adjusted versions, however, offering improvements in both bias and variability. In each, bias approaches zero as the sample size increases without ever becoming positive. At small sample sizes, the adjusted Schnabel becomes unbiased slightly faster, while the adjusted Schumacher-Eschmeyer is less variable, resulting in a lower RMSE for the adjusted Schumacher-Eschmeyer within a small range of sample sizes before the two become effectively equivalent at larger sample sizes (Fig. 4). While of limited practical impact, simulation results revealed that taking the ceiling of the adjusted Schnabel estimator provides a more appropriate estimate of population size, with slightly reduced bias across all sample sizes. The adjusted Schumacher-Eschmeyer estimator presented in (Eq. 8) exhibited the fastest reduction in bias at small sample sizes while remaining exactly unbiased at large sample sizes among a variety of alternate formulations considered. That is, it was seen to satisfy that:

$$E_N[\widehat{N}] \rightarrow N \text{ } as \text{ } \sqrt[k]{M \prod_{i=2}^{k} n_i} \rightarrow N \tag{13}$$

All patterns were essentially the same in the case of $M > n_2 = n_3$, and the only apparent difference when $M < n_2 = n_3$ is even higher variability in the unadjusted Schumacher-Eschmeyer estimator at small sample sizes. Notice that the slightly lower RMSE of the adjusted Schumacher-Eschmeyer is driven by lower variability at small sample sizes, even though the absolute bias of the adjusted Schnabel decreases slightly faster (Fig. 5).

3.2 Minimum sample size

Simulations confirmed the minimum sample sizes (Eq. 10) suggested by Robson and Regier (1964) for the Chapman estimator to produce approximately unbiased estimates. This threshold is represented by a dashed vertical line in Fig. 2. Relative bias (Eq. 11) at these minimum suggested sample sizes was less than 2% for cases $M = n$, $M < n$, and $M > n$ across all population sizes N from 10^2 to 10^6.

Likewise, the approximate minimum sample size rule (Eq. 11) proposed for the bias-adjusted Schnabel estimator, following the logic in (Eq. 10) and confirmed based on examination of simulation results, reveals that the geometric mean of the sample sizes across all visits should be at least twice the square root of the population size divided by one less the number of visits. Relative bias at this sample size was observed to generally range between approximately -10% and -3% for population sizes N ranging from 10^2 to 10^6, respectively, increasing slightly with the number of visits (Table 1). A corresponding simple rule of thumb to achieve a consistent bias threshold relative to population size was not readily apparent for the bias-corrected Schumacher-Eschmeyer estimator.

4. Discussion

The methods described in this paper have been well studied over the decades since they have been published and are still widely encountered today, including in introductory ecology textbooks (Krebs 1999) and statistical programming packages (Nelson, 2023, >100 K total downloads; Ogle et al., 2023, >550 K total downloads). The comparison of these multiple common methods side by side under known conditions provides insight into the exact behavior of the estimators across a complete range of sample sizes.

It is clear that among the two-visit methods, the Chapman estimator is superior in terms of both bias and variance. The estimator of Bailey (1951, 1952), originally proposed for sampling with replacement and not presented here due to its similarity to Chapman, behaves much the same but retains a slightly larger downward bias across all sample sizes, which can be inferred from a simple examination of the equation. The Chao and Bayesian estimators can easily be eliminated as viable options for data of this type in that the Chao estimator possesses a large upward bias when $M \neq n$ and the Bayesian estimator possesses much higher bias and variance at small sample sizes than the other methods. The unidirectional bias of the Chapman estimator is appealing, but it should be noted that a strong downward bias is still present at very small sample sizes, and therefore it is not recommended to be used when sample sizes are smaller than that approximated by (Eq. 10). This recommendation is paradoxical in that the approximate minimum sample size depends on an estimate of the population size, which one is aiming to estimate to begin with. Therefore, it should be seen as a way of avoiding inaccurate estimates from absurdly small sample sizes based on an educated guess of the order of magnitude of the size of the population being sampled.

Likewise among multi-visit methods, the bias-adjusted estimators of Schnabel and Schumacher-Eschmeyer clearly exhibited superior performance at small sample sizes, with no reduction in performance at large sample sizes. Krebs (1999) suggested thresholds for when the small sample adjusted Schnabel should be used, however based on the present simulation results it is recommended that the adjusted estimators are used in place of the originals in all scenarios. The bias-adjusted form of the Schumacher-Eschmeyer estimator presented here can be especially impactful on population size estimates of species with low sample sizes and numbers of recaptures (e.g., only two recaptured Atlantic sturgeon among 17 marked in a study by Kahn et al. (2014)). Application can also extend to calculating more robust closed population estimates for threatened species either directly or indirectly impacted by fishing pressure, such as Indo-Pacific humpback dolphin (Zhou et al. 2007), or juvenile marine turtle species observed to exhibit high fidelity to distinctive foraging areas (Wildermann et al. 2019).

Deviation from assumptions in field scenarios can lead to additional limitations of these estimators to accurately estimate population size that extend beyond small sample bias. First and foremost, if a population is assumed to be closed but in fact immigration and emigration are occurring, population size may be overestimated since there are fewer marked individuals present to be captured at any given time. Conversely, large numbers of individuals with low capture probabilities within a closed population are assumed to lead to underestimation of population size among all estimators evaluated (Amstrup et al. 2005). However, Grimm et al. (2014) observed that a multiple Lincoln-Petersen estimator (with Chapman correction), similar to the multi-visit methods evaluated here but with data from all subsequent visits pooled, performed well on field data even with heterogeneous capture probabilities among individuals. This led them to conclude it is a viable estimator of minimum population size provided a sufficient number of individuals with high catchability are sampled, as capture probability increases and heterogeneity decreases over multiple capture periods. Additionally, Seber (1982) noted the regression-based Schumacher-Eschmeyer is expected to be the most robust multi-visit method regarding violations of assumptions at the expense of loss in efficiency, which was observed in the present study in terms of higher small sample variability when assumptions were exactly met.

This lends support to favoring the adjusted Schumacher-Eschmeyer both in terms of theoretical performance and robustness to violations of assumptions.

A generalization of (Eq. 10), (Eq. 11) can provide a rough approximation of the minimum sample size necessary to avoid a strong downward theoretical bias in estimates from the adjusted multi-visit methods. Note, however, that the downward bias at this precise threshold will be slightly larger for the adjusted Schnabel (Table 1) than for the two-visit Chapman estimator with the trade-off of decreased variability at equal total sample sizes, and larger still for the adjusted Schumacher-Eschmeyer (Fig. 5). This may be something to consider when selecting an estimator depending on the relative risks associated with over vs. under-estimating a given population size, noting that the approximate sample size provided by (Eq. 11) should be seen even more as a lower bound when using the adjusted Schumacher-Eschmeyer.

Otis et al. (1978) and Evans, Kim, and O'Brien (1996) outline the problems with using traditional normal approximation based methods to construct confidence intervals for mark-recapture estimates, especially at small sample sizes, which fail to capture asymmetry around estimates and are known to result in coverage below nominal levels. Buckland and Garthwaite (1991) give general approaches to obtain confidence intervals for recapture data using parametric bootstrapping techniques, which are shown to provide more robust estimates. Tyers (2021) implements a bootstrap method shown to be robust for two-visit data, in which data m for the second sampling event are resampled with replacement using a binomial distribution with size parameter n and probability parameter m/n. These or similar bootstrap techniques can be extended to the multi-visit methods outlined here, which also have the advantage of being valid for the exact form of the estimator used.

Other sources have provided rules of thumb as to the number of recaptures needed to achieve approximately unbiased estimates for closed population mark-recapture methods (Bailey 1951, Ricker 1975, Krebs 1999), which may simplify sampling in certain scenarios. However, all results presented here treat the sample sizes as fixed and the number of recaptures as random, in accordance with the procedures evaluated. It should also be noted that the present evaluation of performance relates to the intrinsic theoretical bias of the estimators when assumptions are perfectly met, and should not be extended to sources of bias or variability arising from sampling methodology.

Acknowledgements

I would like to thank five anonymous reviewers for constructive suggestions to improve this manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Amstrup, S.C., McDonald, T.L., Manly, B.F.J., 2005. Handbook of capture-recapture analyses. Princeton University Press. Princeton, New Jersey, USA.

Bailey, N.T., 1951. On estimating the size of mobile populations from recapture data. Biometrika 38(3-4), 293-306. https://doi.org/10.2307/2332575

Bailey, N.T., 1952. Improvements in the interpretation of recapture data. J. Anim. Ecol. 21(1), 120-127. https://doi.org/10.2307/1913

Brittain, S., Böhning, D., 2009. Estimators in capture-recapture studies with two sources. Adv. Stat. Anal. 93(1), 23-47.

Buckland, S.T., Garthwaite, P.H., 1991. Quantifying precision of mark-recapture estimates using the bootstrap and related methods. Biometrics 47(1), 255-268. https://doi.org/10.2307/2532510

Castello, L. 2004. A method to count pirarucu *Arapaima gigas*: fishers, assessment, and management. N. Am. J. Fish. Manag. 24(2), 379-389. https://doi.org/10.1577/M02-024.1

Chao, A., 1987. Estimating the population size for capture-recapture data with unequal catchability. Biometrics 43(4), 783-791. https://doi.org/10.2307/2531532

Chao, A., 1989. Estimating population size for sparse data in capture-recapture experiments. Biometrics 45(2), 427-438. https://doi.org/10.2307/2531487

Chapman, D.G., 1951. Some properties of the hypergeometric distribution with applications to zoological sample censuses. Univ. Calif. Publ., Stat. 1(7), 131-160.

Chapman, D.G., 1952. Inverse, multiple and sequential sample censuses. Biometrics 8(4), 286-306. https://doi.org/10.2307/3001864

Evans, M.A., Kim, H.M., O'Brien, T.E, 1996. An application of profile-likelihood based confidence interval to capture: recapture estimators. J. Agric. Biol. Environ. Stat. 1(1), 131-140. https://doi.org/10.2307/1400565

Grimm, A., Gruber, B., Henle, K., 2014. Reliability of different mark-recapture methods for population size estimation tested against reference population sizes constructed from field data. PLoS One 9(6), e98840. https://doi.org/10.1371/journal.pone.0098840

Hale, E.A., Park, I.A., Fisher, M.T., Wong, R.A., Stangl, M.J., Clark, J.H, 2016. Abundance estimate for and habitat use by early juvenile Atlantic sturgeon within the Delaware River Estuary. Trans. Am. Fish. Soc. 145(6), 1193-1201. https://doi.org/10.1080/00028487.2016.1214177

Hansen, J.M., Van Kirk, R.W., 2018. A mark-recapture-based approach for estimating angler harvest. N. Am. J. Fish. Manag. 38(2), 400-410. https://doi.org/10.1002/nafm.10038

Hansen, M.J., Horner, N.J., Liter, M., Peterson, M.P., Maiolie, M.A., 2008. Dynamics of an increasing lake trout population in Lake Pend Oreille, Idaho. N. Am. J. Fish. Manag. 28(4), 1160-1171. https://doi.org/10.1577/M07-149.1

Huggins, R.M., 1989. On the statistical analysis of capture-recapture experiments. Biometrika 76, 133-140. https://doi.org/10.1093/biomet/76.1.133

Huggins, R.M., 1991. Some practical aspects of a conditional likelihood approach to capture experiments. Biometrics 47, 725-732. https://doi.org/10.2307/2532158

Kahn, J.E., Hager, C., Watterson, J.C., Russo, J., Moore, K., Hartman, K., 2014. Atlantic Sturgeon annual spawning run estimate in the Pamunkey River, Virginia. Trans. Am. Fish. Soc. 143(6), 1508-1514. https://doi.org/10.1080/00028487.2014.945661

Kahn, J.E., Hager, C., Watterson, J.C., Mathies, N., Hartman, K.J., 2019. Comparing abundance estimates from closed population mark-recapture models of endangered adult Atlantic sturgeon. Endanger. Species Res. 39, 63-76. https://doi.org/10.3354/esr00957

Krebs, C.J, 1999. Ecological methodology. 2nd ed. Harper & Row, New York, New York, USA.

Lincoln, F.C., 1930. Calculating waterfowl abundance on the basis of banding returns. U.S. Department of Agriculture Circular 118, 1-4.

McNair, J.N., Ruetz III, C.R., Carlson, A., Suh, J, 2018. Reducing effects of dispersal on the bias of 2-sample mark-recapture estimators of stream fish abundance. PLoS ONE 13(8), e0200733. https://doi.org/10.1371/journal.pone.0200733

Nelson, G.A, 2023. fishmethods: Fishery Science Methods and Models. R package version 1.11-3. URL https://CRAN.R-project.org/package=fishmethods

Ogle, D.H., Doll, J.C., Wheeler, P., Dinno, A., 2023. FSA: Fisheries Stock Analysis. R package version 0.9.4. URL https://github.com/fishR-Core-Team/FSA

Otis, D.L., Burnham K.P., White, G.C., Anderson D.R, 1978. Statistical inference from capture data on closed animal populations. Wildl. Monogr. 62, 1-135.

Petersen, C.G.J, 1896. The yearly immigration of young plaice into the Limfjord from the German Sea. Report of the Danish Biological Station to the Board of Agriculture 6, 1-48.

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Ricker, W.E., 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191, 1-382.

Robson, D.S., Regier, H.A., 1964. Sample size in Petersen mark-recapture experiments. Trans. Am. Fish. Soc. 93(3), 215-226. https://doi.org/10.1577/1548-8659(1964)93[215:SSIPME]2.0.CO;2

Rowe, S., 2002. Population parameters of American lobster inside and outside no-take reserves in Bonavista Bay, Newfoundland. Fish. Res. 56(2), 167-175. https://doi.org/10.1016/S0165-7836(01)00321-6

Sadinle, M., 2009. Transformed logit confidence intervals for small populations in single capture-recapture estimation. Commun. Stat. Simul. Comput. 38(9), 1909-1924. https://doi.org/10.1080/03610910903168595

Schnabel, Z.E., 1938. The estimation of the total fish population of a lake. Am. Math. Mon. 45(6), 348-352. https://doi.org/10.2307/2304025

Schumacher, F.X., Eschmeyer, R.W., 1943. The estimation of fish populations in lakes or ponds. J. Tenn. Acad. Sci. 18, 228-249.

Seber, G.A.F., 1982. The estimation of animal abundance and related parameters. 2nd ed. Macmillan, New York, New York, USA.

Spencer, S.C., Colby, P.J., Momot, W.T., Fruetel, M. 2002. Response of a walleye population to pulse fishing in Henderson Lake, Ontario. N. Am. J. Fish. Manag. 22(3), 842-851. https://doi.org/10.1577/1548-8675(2002)022<0842:ROAWPT>2.0.CO;2

Tyers, M., 2021. recapr: Two event mark-recapture experiment. R package version 0.4.4. URL https://CRAN.R-project.org/package=recapr

Webster, A.J., Kemp, R. 2013. Estimating omissions from searches. Am. Stat. 67(2), 82-89. https://doi.org/10.1080/00031305.2013.783881

Wildermann, N.E., Sasso, C.R., Stokes, L.W., Snodgrass, D., Fuentes, M.M., 2019. Habitat use and behavior of multiple species of marine turtles at a foraging area in the Northeastern Gulf of Mexico. Frontiers in Marine Science 6, 155. https://doi.org/10.3389/fmars.2019.00155

Wittes, J.T., 1972. Note: on the bias and estimated variance of Chapman's two-sample capture-recapture population estimate. Biometrics 28(2), 592-597. https://doi.org/10.2307/2556173

Zhou, K., Xu, X., Tian, C. 2007. Distribution and abundance of Indo-Pacific humpback dolphins in Leizhou Bay, China. N. Z. J. Zool. 34(1), 35-42. https://doi.org/10.1080/03014220709510061

546 Fig. 1. Two-visit simulated estimates of population size $\widehat{N}$ +/- standard errors across sample
547 sizes, denoted as the geometric mean over all visits, from a population of $N = 1{,}000$ individuals,
548 with M (i.e., n_1) < n_2.

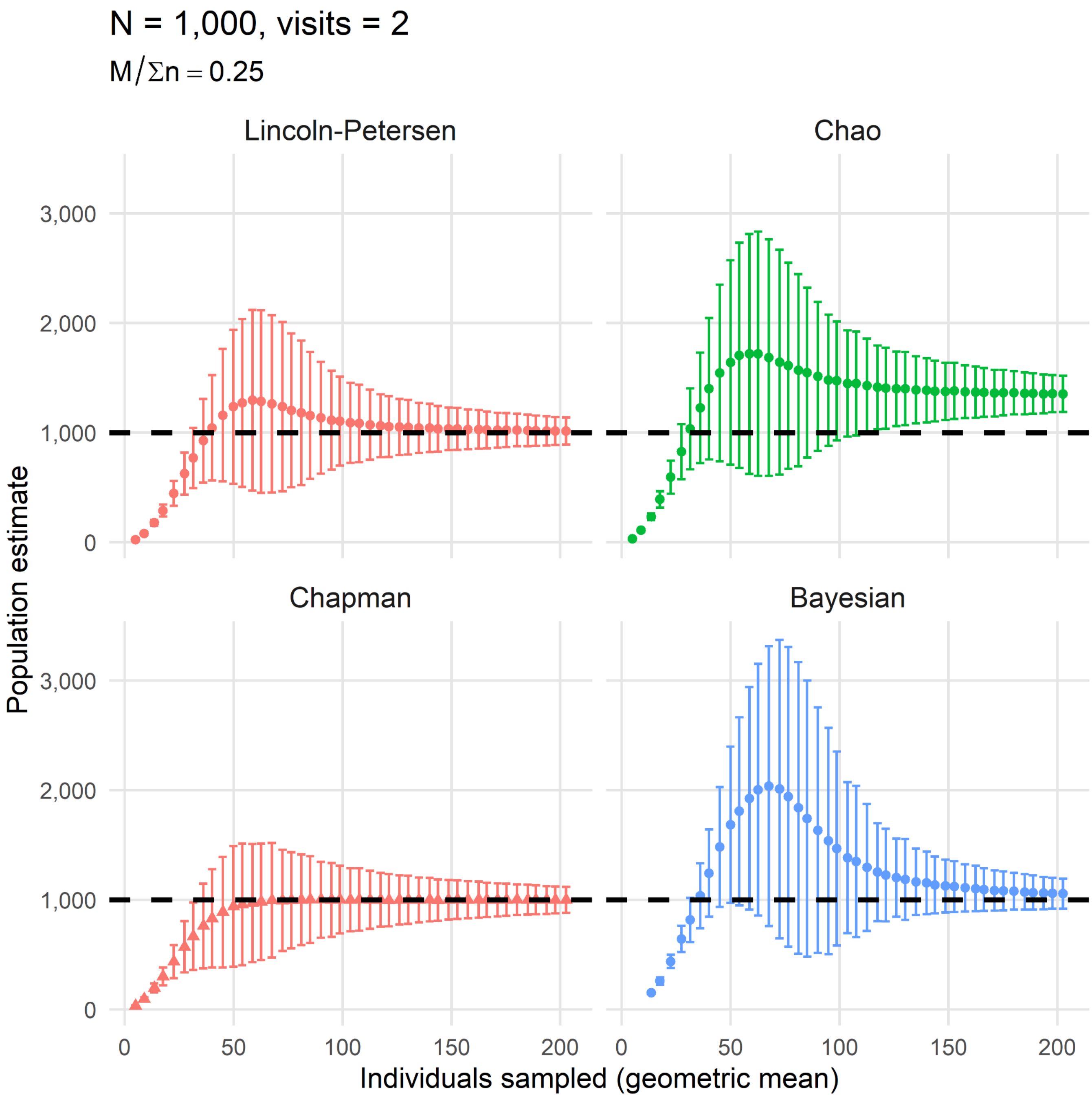


Fig. 2. Scaled root mean square error (RMSE) of the two-visit estimators across sample sizes, denoted as the geometric mean over all visits, from a population of $N = 1{,}000$ and M (i.e., n_1) < n_2. Dashed vertical line represents the sample size at which the Chapman estimator becomes approximately unbiased.

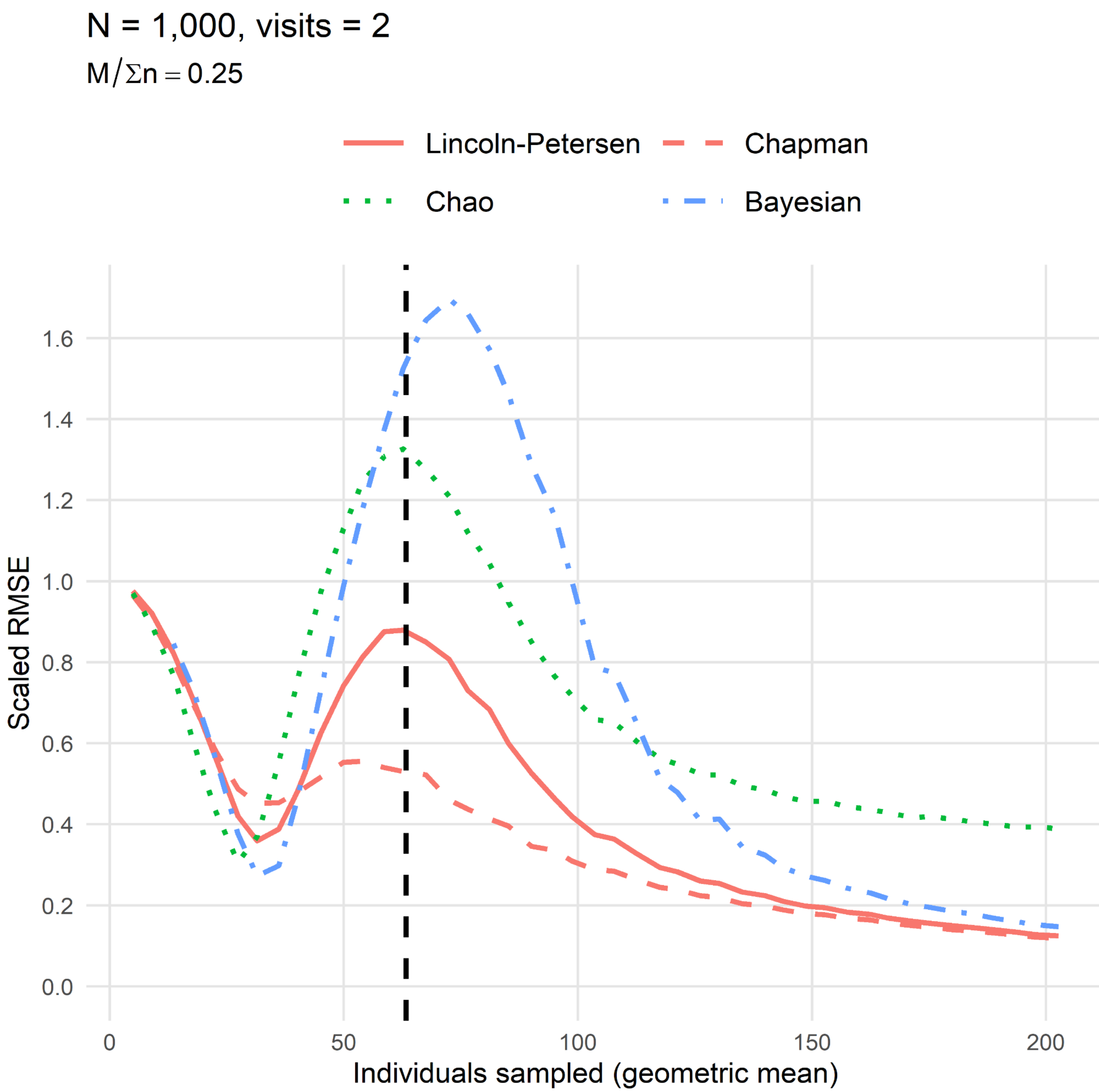


Fig. 3. Multi-visit ($k = 3$) simulated estimates of population size $\widehat{N}$ +/- standard errors across sample sizes, denoted as the geometric mean over all visits, from a population of $N = 1,000$ individuals, with M (i.e., n_1) = n_2 = n_3.

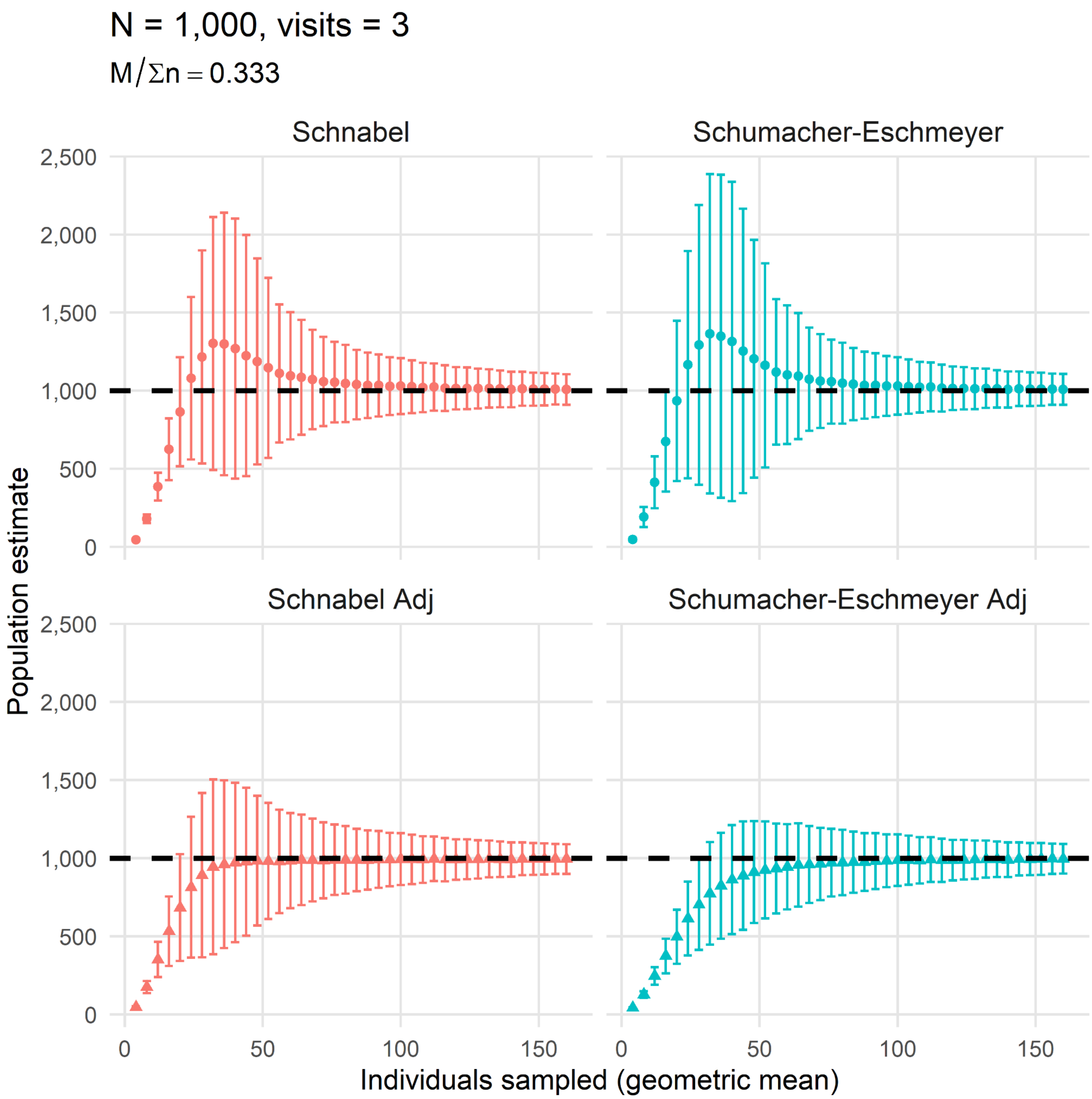


Fig. 4. Scaled root mean square error (RMSE) of the multi-visit estimators across sample sizes, denoted as the geometric mean over all visits, from a population of $N = 1{,}000$ and M (i.e., n_1) = $n_2 = n_3$. Dashed vertical line represents the sample size at which the adjusted Schnabel estimator becomes approximately unbiased.

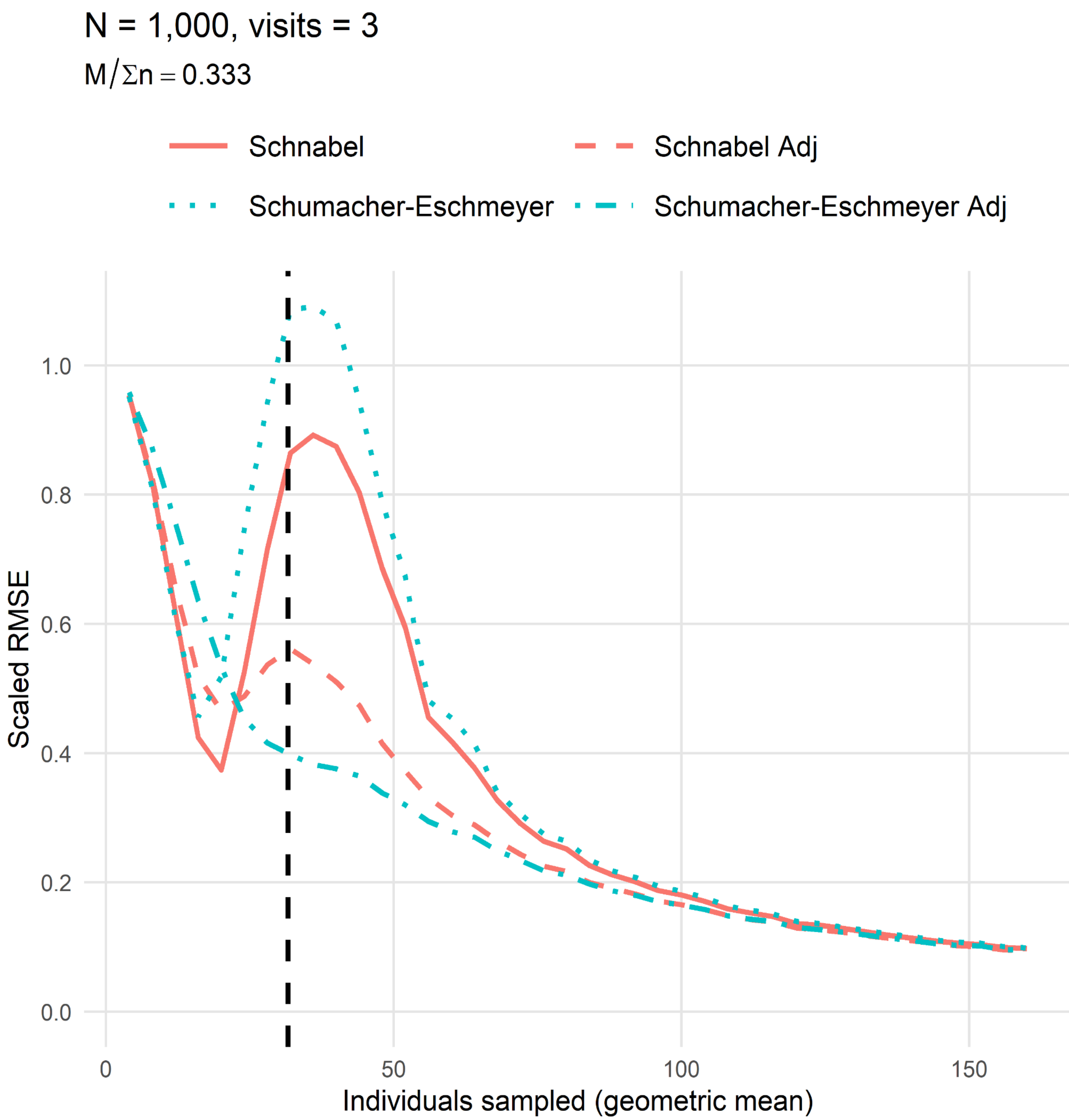
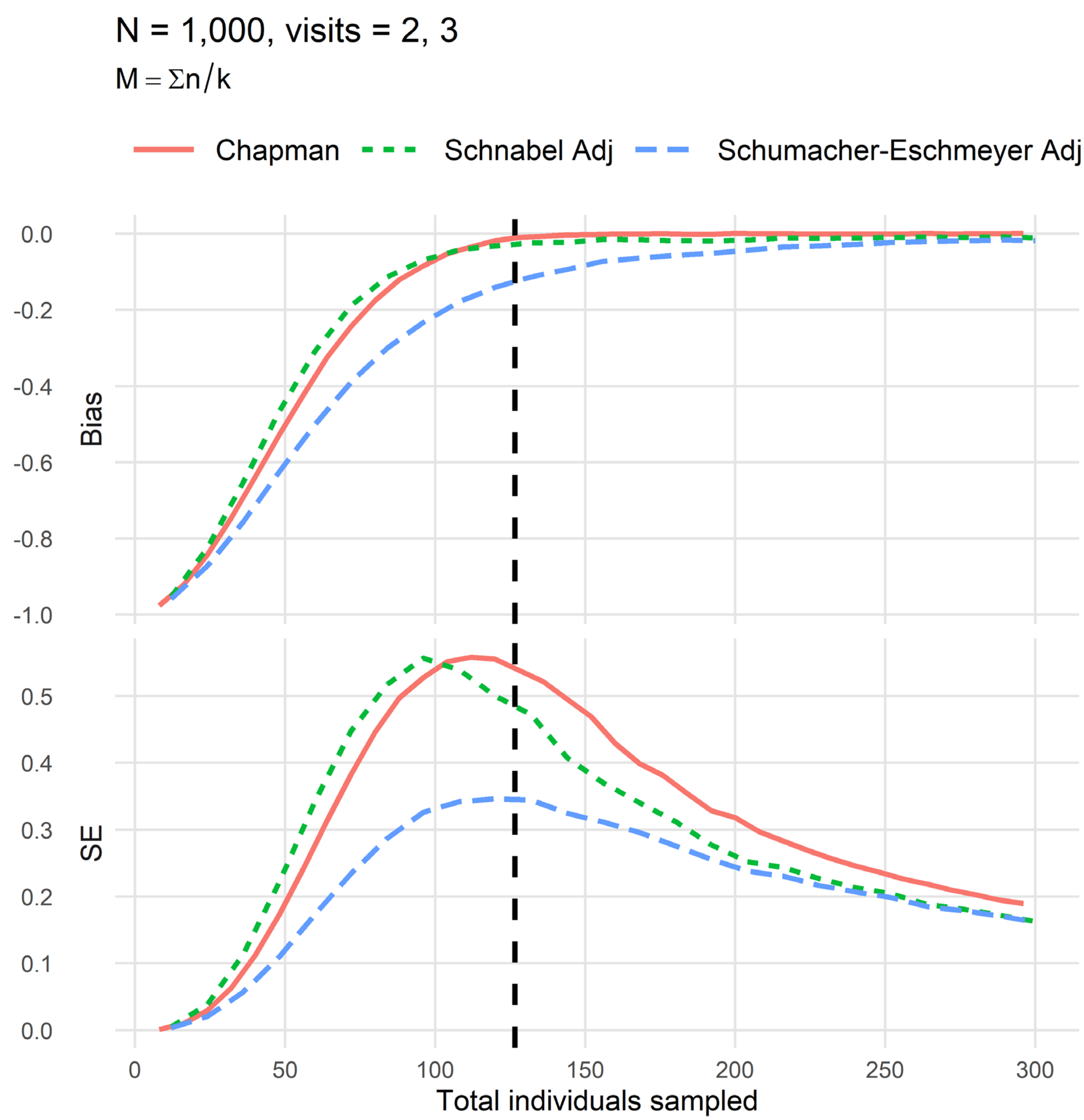


Fig. 5. Comparison of relative bias and standard error (SE) of the Chapman estimator with those of the adjusted multi-visit estimators across total sample size over all visits ($k = 2$ for Chapman, $k = 3$ for multi-visit estimators), from a population of $N = 1{,}000$ and M (i.e., n_1) $= n_2$ ($= n_3$). Dashed vertical line represents the sample size at which the 2-visit Chapman estimator becomes approximately unbiased (Eq. 10). Note the Chapman estimator possesses slightly less downward bias at this threshold than the adjusted Schnabel for an equal number of total individuals sampled spread over 3 visits, with the advantage of lower variability in the adjusted Schnabel. The adjusted Schumacher-Eschmeyer possesses the lowest variability among methods at the expense of taking longest to become unbiased.



	$k = 3$		$k = 4$		$k = 5$	
N	$\bar{\bar{N}}$	Pct. Error	$\bar{\bar{N}}$	Pct. Error	$\bar{\bar{N}}$	Pct. Error
100	89	-10.6	89	-10.8	87	-13.2
1,000	935	-6.5	912	-8.8	904	-9.6
10,000	9,394	-6.1	9,293	-7.1	9,138	-8.6
100,000	94,792	-5.2	93,062	-6.9	91,832	-8.2
1,000,000	948,038	-5.2	930,035	-7.0	916,636	-8.3

Tbl. 1. Relative bias in the adjusted Schnabel estimator (Eq. 6) for M (i.e., n_1) = $n_{2 \ldots k}$ at recommended minimum sample sizes from (Eq. 11) for $k = 3$, 4 and 5 visits at various population sizes N, based on 100,000 simulations.