

The Siren Song of the Schaefer Estimator – no better than a pooled Petersen.

Carl James Schwarz¹

Department of Statistics and Actuarial Science

Simon Fraser University

8888 University Drive

Burnaby, BC V5A 1S6

cschwarz@stat.sfu.ca

A. Neil Arnason, C. W. Kirby

Department of Computer Science

University of Manitoba

Winnipeg, MB R3T 2N2

arnason@cs.umanitoba.ca

Keyword: mark-recapture; tag-recapture; salmon escapement; stratified-Petersen

¹ Author for correspondence

Abstract:

The Schaefer estimator is commonly used as an alternative to a simple Petersen in mark-recapture experiments when the assumptions of equal mixing, homogeneous capture, or homogeneous recovery probabilities may not hold. It has been derived in various ways without realizing that all formulations are algebraically identical. We review the conditions under which the estimator is consistent, find the size of the bias when assumptions are violated, and find the approximate variance of the estimator.

Surprisingly, these indicate that the Schaefer estimator performs essentially the same as the Petersen estimator. Consequently, there is no reason to prefer it over the more simpler and better understood Petersen estimate. However, there are still benefits to using the stratified-Petersen (the Darroch estimator) in cases where both the Schaefer and pooled-Petersen estimator are inconsistent.

1. Introduction

The Petersen estimator is one of the most widely used mark-recapture estimators in fisheries management. However, it can be severely biased if its underlying assumptions are violated – in particular, if complete mixing of tagged and untagged fish, or homogeneous capture probabilities, or homogeneous recapture probabilities does not occur.

In such cases, stratifying the captures and recaptures by location and/or time have been proposed. Two common estimators for these experiments have been used – the

stratified-Petersen first introduced by Darroch (1961) with subsequent work by Plante, Rivest and Trembley (1998), Banneheka, Routledge, and Schwarz (1997), and Schwarz and Taylor (1998); or the Schaefer estimator first introduced by Schaefer (1951) but with little follow-up work.

Eames, et al. (1983) and other have noted that the Schaefer and the simple Petersen estimator are often very similar. On the other hand, Ricker (1975, p. 104) notes that "such close agreement would be infrequent". Surprisingly, although the Schaefer estimator has been used in the literature in three common disguises, its properties have not been systematically investigated. The conditions under which it is consistent were known, but the variance of the estimator and the size of the bias when the assumptions fail is unknown.

In this paper, we show that the three common "intuitive" derivations of the Schaefer estimator are, in fact, identical. Next we review the conditions under which it is consistent, develop an approximation for the its bias and an approximation for its variance. The key, surprising, result is that despite its apparently more complex form and attempts to deal with stratification, the Schaefer estimator is basically no better than the simple pooled-Petersen which ignores stratification. It is not surprising then that the two estimates are often very similar.

In most biological situations, the pooled-Petersen will be biased and so the biologist needs some guidance on when the bias is likely to be so large that the estimate

is unacceptable. A recent study (Warren and Dempson 1995) looked at this using simulation. Their results were instructive for the case they considered, the Conne River experiment of Dempson and Stansbury (1991), but they lack generality because they only considered situations where the pooled-Petersen is unbiased. They did not consider situations in which selective stratification might outperform pooled-Petersen by decreasing bias enough to overcome the precision disadvantage. (All the estimates in the Warren and Dempson study are asymptotically unbiased because they assumed equal tagging rates across strata and this is a sufficient condition for consistency of the pooled-Petersen; thus their simulation results on bias compare only the small sample bias among the estimates achieved at different pooling levels). However, there are still benefits to using the stratified-Petersen (the Darroch estimator) in cases where both the Schaefer and pooled-Petersen estimator are inconsistent.

2. The Schaefer estimator and its properties

2.1 The experiment, its statistics, and notation

Two common scenarios where a stratified analysis would be contemplated are salmon returning to spawn (e.g. Schwarz and Taylor, 1998), and smolt returning to the ocean (Warren and Dempson, 1995). In both cases, fish are initially captured, tagged with individually numbered tags, and released. These typically take place over a long period of time, and consequently, the releases are stratified by time (e.g. daily or weekly). Later, fish are again captured, and a count is made of the total number captured, how many of

the fish captured are marked, and the tag numbers of the recovered fish. Again, as this takes place over a period of time, the recoveries can be stratified by time (e.g. daily or weekly). Because marked fish are individually numbered, their stratum of release and stratum of recovery can be determined. The data can be arranged in an array as shown in Figure 1.

The underlying population of interest can also be similarly stratified as shown in Figure 2 (of course the value of N_{ij} would be unknown). The objectives are to estimate (1) N – the total run size and (2) $N_{i\cdot}$ or $N_{\cdot j}$ the individual run components.

2.2 The three “derivations”

In the first derivation, a Petersen estimator is derived for the population in each individual release stratum. This requires three statistics – the number tagged and released in the stratum, the total fish recovered from that stratum, and the number of marks recovered from that stratum. The first and third statistics are available from the information in Figure 1; the second statistic is unavailable and must be “estimated”. This is done by estimating how many of the total recoveries (n_j^r) would have come from release stratum i by apportioning the releases in the same ratio as the marked fish:

$$\hat{r}_i = \sum_{j=1}^t n_j^r \frac{m_{ij}}{m_{\cdot j}}$$

A Petersen estimator is then formed for each row as:

$$\hat{N}_{i\cdot} = \frac{n_i^c \hat{r}_i}{m_{i\cdot}} = \sum_{j=1}^t \frac{n_i^c n_j^r m_{ij}}{m_{i\cdot} m_{\cdot j}}$$

and an estimate of the overall total run is found by summing the individual row estimates:

$$\hat{N}_{Schaefer} = \sum_{i=1}^s \hat{N}_{i\bullet} = \sum_{i=1}^s \sum_{j=1}^t \frac{n_i^c n_j^r m_{ij}}{m_{i\bullet} m_{\bullet j}}$$

In the second derivation, a Petersen estimator is derived for the population in each recovery stratum. Now the number of fish initially tagged and belonging to each recovery stratum must be estimated. In much the same way as above, the rows of Figure 1 are used to “redistribute” the n_i^c among the columns in the same proportion as the observed marks:

$$\hat{t}_j = \sum_{i=1}^s n_i^c \frac{m_{ij}}{m_{i\bullet}}.$$

A Petersen estimate is formed for each column as:

$$\hat{N}_{\bullet j} = \frac{\hat{t}_j n_j^r}{m_{\bullet j}} = \sum_{i=1}^s \frac{n_i^c n_j^r m_{ij}}{m_{i\bullet} m_{\bullet j}}$$

and an estimate of the overall run is found by summing the individual row estimates giving:

$$\hat{N}_{Schaefer} = \sum_{j=1}^t \hat{N}_{\bullet j} = \sum_{i=1}^s \sum_{j=1}^t \frac{n_i^c n_j^r m_{ij}}{m_{i\bullet} m_{\bullet j}}$$

This was the derivation used by Macdonald and Smith (1980) in their simpler estimator, but they did not apparently realize that their estimator was equivalent to the Schaefer estimator.

In the third derivation, a Petersen estimator is found for each of the st sub-populations. For each estimate, both the number initially tagged and the number recovered from each cell must be estimated by apportioning the number released and recovered in the same proportion as the marks:

$$\hat{t}_{ij} = n_i^c \frac{m_{ij}}{m_{i\cdot}} \text{ and } \hat{r}_{ij} = n_j^r \frac{m_{ij}}{m_{\cdot j}}.$$

Then a Petersen estimate is formed for each cell as:

$$\hat{N}_{ij} = \frac{\hat{t}_{ij} \hat{r}_{ij}}{m_{ij}} = \frac{n_i^c \frac{m_{ij}}{m_{i\cdot}} n_j^r \frac{m_{ij}}{m_{\cdot j}}}{m_{ij}} = \frac{n_i^c n_j^r m_{ij}}{m_{i\cdot} m_{\cdot j}}$$

The overall population is then estimated by:

$$\hat{N}_{Schaefer} = \sum_{i=1}^s \sum_{j=1}^t \hat{N}_{ij} = \sum_{i=1}^s \sum_{j=1}^t \frac{n_i^c n_j^r m_{ij}}{m_{i\cdot} m_{\cdot j}}$$

This was the derivation presented by Warren and Dempson (1995).

All three “derivations” are algebraically identical and proposed by Schaefer (1951).

The pooled-Petersen is formed by ignoring the stratification and forming the estimator:

$$\hat{N}_{PooledPetersen} = \frac{n_{\cdot}^c n_{\cdot}^r}{m_{\cdot\cdot}}.$$

2.3 The properties of the Schaefer estimator

In order to derive the properties of the Schaefer estimator, parameters representing the capture, recovery, and “movement” processes are defined in Figure 3. “Movement” refers to the probability that a fish initially present in stratum i at the time of tagging will be present in stratum j at the time of recovery.

To assess the consistency of \hat{N} , Chapman and Junge (1956) replaced all statistics by their expected value giving:

$$E[\hat{N}_{Schaefer}] \approx \sum_{i=1}^s \sum_{j=1}^t \frac{(N_{i\cdot} p_i^c) \left(\sum_{i=1}^s N_{i\cdot} \theta_{ij} p_j^r \right) (N_{i\cdot} p_i^c \theta_{ij} p_j^r)}{\left(\sum_{j=1}^t N_{i\cdot} p_i^c \theta_{ij} p_j^r \right) \left(\sum_{i=1}^s N_{i\cdot} p_i^c \theta_{ij} p_j^r \right)}$$

The estimator will be consistent, (i.e. will equal N), in an infinite number of ways, but the following special cases have an easy interpretation:

- (1) $p_i^c = \bar{p}^c$, i.e. the probability of capture is equal among all tagging strata;
- (2) $p_j^r = \bar{p}^r$ and $\theta_{i\cdot} = \theta$, i.e. the probability of recovery is equal among all recovery strata and each tagging stratum has the same overall probability of recovery.
- (3) $\theta_{ij} = \theta_j$, i.e. homogeneous movement in all release strata – complete mixing.

These are the same conditions as required for consistency of the pooled-Petersen as outlined by Seber (1982). And, as in the case of the pooled-Petersen, closure is assumed. Death or permanent emigration ($\theta_{i\cdot} < 1$) implies that only the population in the tagging strata can be estimated; “birth” or immigration imply that only the population at the end of the experiment can be estimated; both forms of non-closure imply it estimates

some function of total fish ever present (the initial population + immigrants who survive to the recovery periods).

Even if the Schaefer estimator is consistent under similar conditions as the pooled-Petersen, perhaps its “bias” is smaller when assumptions are violated. Schwarz and Taylor (1998) showed that the relative bias in the pooled-Petersen from heterogeneity of the capture or recovery probabilities can be approximated by:

$$\hat{B}_{PooledPetersen}^{rel} \cong -C(\text{tagging}, \text{recovery}) \frac{\sqrt{V(\text{tagging})V(\text{recovery})}}{E[\text{tagging} \times \text{recovery}]}$$

where C, V, and E refer to the correlation, variance, and expectation of the tagging and recovery probabilities taken over the entire population. By using similar methods as outlined in Macdonald and Smith (1980), the relative bias in the Schaefer estimator (Appendix I) is found to be:

$$\hat{B}_{Schaefer}^{rel} \cong -C(\text{tagging}, \text{recovery}) \frac{\sqrt{V(\text{tagging})V(\text{recovery})}}{E[\text{tagging}]E[\text{recovery}]}$$

which is essentially the same as that for the pooled-Petersen except for the denominator.

There are three cases to consider. If the tagging and recovery probabilities are independent, then $C(\text{tagging}, \text{recovery})=0$ and both estimators are unbiased. If tagging and recovery probabilities are positively correlated, as is often the case as outlined in Schwarz and Taylor (1998), then $E[\text{tagging} \times \text{recovery}] > E[\text{tagging}]E[\text{recovery}]$ and the pooled estimator's bias is closer to 0 than the Schaefer estimator. If the tagging and recovery probabilities are negatively correlated, then the Schaefer estimator will have a smaller bias.

Seber (1982) noted that the variance of $\hat{N}_{Schaeffer}$ is unknown and we are unaware of any subsequent derivations. In Appendix II, we derive an approximation for the variance and show that the Schaefer estimator has approximately the same coefficient of variation as the pooled-Petersen, i.e.

$$CV_{Schaefer} \equiv CV_{PooledPetersen} \equiv \frac{1}{\sqrt{E[m..]}}$$

Unlike, the Petersen estimator, the Schaefer estimator also provides estimates of the individual run components. Macdonald and Smith (1980) show that the individual components are also biased, i.e.

$$E[\hat{N}_{\cdot j}] \equiv N_{\cdot j} \frac{p_j^r}{\bar{p}^r}$$

and by a similar argument, it can be shown that:

$$E[\hat{N}_{i \cdot}] \equiv N_{i \cdot} \frac{p_i^c}{\bar{p}^c}$$

In many experiments, the total effort is the limiting quantity in any stratum, i.e. the catch is essentially fixed, and so the probability of capture or recovery is small when the corresponding segment of the run is large and vice-versa. Consequently, estimates of the run components will be biased downwards when the run is large, and biased upwards when the run is small. Furthermore, as they noted in their paper, the estimates will tend to follow the constraint $\frac{n_j^r}{\hat{N}_{\cdot j}} \equiv \bar{p}^r$ so that they will not provide any assistance in assessing if the capture or recovery probabilities are heterogeneous. If the population is not closed,

Macdonald and Smith (1980) also showed that the estimates of the run components in the recovery strata are also inflated by an additional factor.

3. Simulation Study

We conducted three simulation studies. The first was used to verify the above results for the Schaefer estimator in small samples. The second and third investigated the performance of the Schaefer and pooled-Petersen vs the more complex stratified-Petersen estimator.

In the first, simulation, we conducted approximately 80 sets of studies where the correlation of the tagging and recovery probabilities ranged from -0.9 to 0.9 and the coefficient of variation in capture probabilities among individuals ranged from 0 to 100%. In each simulation set, five hundred replicate populations of size 2000 were generated with mean tagging and recovery probabilities of 0.20. For each replicate, a 3x3 stratification was induced and the Schaefer and pooled-Petersen estimate computed. The empirical bias and empirical standard errors were computed over these replicates. The results are shown in Figures 4 and 5.

The biases of the two estimates are very similar except in cases of extreme correlation. In these cases, the matrix of marked recoveries will be strongly diagonal which usually indicated non-mixing cases. In the case of complete no mixing, e.g. stratification by gender, the Petersen estimates is formed for each sub-population and the

overall estimate is formed from the sum. There is no gain by considering a stratified estimator.

The variability of the two estimates is also comparable. However, the Schaefer estimator is more sensitive to sparse data. If some of the $E[m_{ij}]$ are small, then an observed count of 0, 1, or 2 in that cell is quite likely, which makes quite a difference for its contribution to the overall total. The pooled-Petersen is robust against these sparse cells.

The second simulation study examined the question if there are biologically realistic situations in which the pooled-Petersen is badly biased by stratum variation in tagging and capture rates and if so, what is the cost in precision in using the unbiased stratified-Petersen estimator? A complete factorial simulation study was carried out on theoretical stratified populations with $s = t = 12$ and sampling plans involving all combinations of the following factors:

1. Two total run sizes, N , were used: small (12,000) and large (120,000).
2. Two entry models: distribute N over the initial strata uniformly and according to an approximate normal shape (peaking in strata 6 and 7)
3. Three migration models: three transition matrices (θ_{ij}) were used, all being upper triangular matrices as are typical of temporal stratification. The matrices can be ordered by the degree of spread: (A) strongly diagonal dominant with elements only on the diagonal ($\theta_{ii}=0.75$) and on the super-diagonal ($\theta_{i,i+1}=0.25$); (B) animals can migrate to strata $j=i$ through $i+3$ with probabilities 0.4, 0.3, 0.2, 0.1; (C) complete upper triangular with uniform migration to strata $j=i$ through $j=12$. The upper triangular transition matrices ensure complete mixing is never met

4. Six tagging probabilities vectors: the 6 vectors had fixed mean (0.0625 or 0.125 depending on factor 1) across the 12 strata but increasing variation (from a coefficient of variation of 16% at factor level 1 to a coefficient of variation of 72% at factor level 6). Achieving these distributions did not require choosing unrealistic capture probabilities. No stratum had a capture probability below 0.01 or above 0.12.
5. Six recovery probability vectors: the same six levels were used as in 4.
6. Two average sampling rates: low (average tagging and recovery rates of 6.25%) and high (average 12.5%). Tagging rate assignment to strata is discussed under factor 4.

Factors 3, 4 and 5 are responsible for lack of consistency in the pooled-Petersen estimate. All the assumptions for the Darroch estimator were met (including the extra assumption that marked and unmarked exhibit the same migration patterns) and so the Darroch estimates are asymptotically unbiased at all factor combinations. For each factor combination, the expected initial and final stratum sample sizes and expected numbers of marked recoveries were computed, and these were used to compute the expected value of the pooled-Petersen estimate, $E[\hat{N}]$. The asymptotic relative bias (ARB) is then computed as $100 \cdot (N - E[\hat{N}]) / N$. The expected statistics can be used in the formula for the standard error for the pooled-Petersen to approximate the expected standard error, $E[SE(\hat{N})]$ and percent CV, $100 \cdot E[SE(\hat{N})] / N$. It is known that bias only begins to have serious effects on confidence interval coverage when the ARB exceeds about half the CV in magnitude (Cochran, 1977). It is important to point out that the expected standard error formed for the pooled-Petersen does not estimate the true standard error when the pooled-Petersen is biased; rather it estimates the average standard error that the experimenter will obtain in this (biased) case. However, stochastic simulations carried out by one of us (CWK) confirm that the expected standard error is close to the true value and that the coefficient

of variation is virtually identical to the true value.

The results on the ARB are reported in Table 1. The table shows results for the 2×2 most extreme values of factors 4 and 5 (out of the $6 \times 6 = 36$ possible factor combinations) but for all combinations of the first three factors. It can be seen (Table 1a) that total population size (factor 1) has no effect on the ARB nor did the average capture probability level (factor 6). Thus all the ARB results are general to all N and average sampling rates. Note however that the ARB does not account for small sample bias which will be affected by sample size. The magnitude of small sample biases is illustrated by Dempson and Warren (1995) and is generally less than 3% in small populations ($N = 10,000$) and less than 1% in large populations ($N > 30,000$). Very large positive biases are easily produced when the consistency conditions are not met but the bias only becomes very large when both sets of capture probabilities exhibit high variability across strata. The coefficient of variation for the factor combinations presented in Table 1 never exceeded 5%, so most of the biases will produce a large reduction in C.I. coverage, i.e. the nominal 95% confidence interval computed using a pooled-Petersen will fail to cover the true population value substantially more than 5% of the time.

The relative root mean square error (MSE) results in Table 1 (b and c) show that the trade-off between precision and unbiasedness when going from the (biased but more precise) pooled-Petersen to the (unbiased but less precise) Darroch estimator generally favours the Darroch estimator. The root MSE is the square root of the bias squared plus the variance of the estimate, and is a widely used measure of this trade-off; “relative” means the root MSE is expressed as a percentage of N . These results are not invariant to factor 1 (N) and 6 (average p). In small populations, the performance of the two estimators is comparable unless there is high variability in both capture probability

vectors (factors 4 and 5) but in larger populations (or equivalently, at higher average capture rates), the loss in precision is not so great in the Darroch estimator and it will outperform the biased pooled-Petersen. Because the bias is 0 for the Darroch estimator, the results in Table 1(c) are in fact the coefficient of the estimate, which shows that quite respectable precision is obtainable even in small populations with relatively low (6.25%) average sampling rates.

A third simulation was to investigate the consistency of the estimators in situations where "births" and/or "deaths" occur between the tagging and recovery samples. ["Births" is the generic term for any introduction of new animals; "deaths" is the generic term for all disappearances of animals from the study population.] It was assumed that survival rates depend only on the stratum of origin and that survival rates and migration rates apply equally to the marked and unmarked animals of the initial stratum. Survival rates and numbers of new entries (births) were varied across strata to avoid seeing consistency resulting from mere constancy of these values across strata. Consistency was examined in the Schaefer, pooled-Petersen and Darroch estimators of N using 4×4 migration matrices that were diagonally dominant and non-singular but not upper triangular. Consistency was also examined for the Darroch estimators of N and the individual stratum sizes with $s > t$ and $s < t$.

Table 2 summarizes the cases where bias exists and, where it exists, its direction. The first 4 lines (closed models) recapitulate results that we already know from the analytic consistency theory reported in Seber (1982) and as confirmed by our second simulation. The next 4 lines (death only) show that the Schaefer and pooled-Petersen estimate may now show bias even when consistency conditions hold (the p^c are constant). This is to be expected because the differential mortality of the marked animals produces unequal

marked ratios in the recovery strata even though the initial marked fraction is the same in all strata. The Darroch estimator is not biased in this situation, as we would expect, at least in cases where initial stratum estimates are available ($s \leq t$). When $s > t$, the estimates of final strata are (severely) biased as we would expect from the fact that the Petersen estimate cannot be used to estimate final sizes when there are losses. The negative bias in N for Unequal p^c when $s > t$ is a surprise for which no explanation is readily apparent.

The Birth-only results in the next four lines of Table 2 are the predictable complements of the Death-only results. The Petersen estimate can estimate population sizes at the final time when births are occurring, the pooled-Petersen and Schaefer are biased when the p^r are unequal and the Darroch estimate is biased only where final stratum estimates are not available ($s < t$).

Finally, the presence of both births and deaths causes positive bias, as we would expect from the known result for the Petersen estimate in this situation. The bias can be enormous.

4. Discussion

Surprisingly, the Schaefer estimator is essentially equivalent to the pooled-Petersen estimator in terms of conditions for consistency, bias when assumptions are violated, and precision. Consequently, when estimating the overall run size, there is no reason why it should be preferred over the pooled estimator.

The Schaefer estimator provides “sensible looking” estimates of the individual run components, but these can be severely biased under heterogeneous capture or recovery probabilities. In addition, the estimates themselves do not provide any means of checking these crucial assumptions. We again cannot recommend the use of the Schaefer estimator under these circumstances.

This still leaves the problem of how to deal with experiments of this sort. Darroch’s (1961) stratified Petersen estimator provides consistent estimates of the run totals and run components. Schwarz and Taylor (1998) provide a thorough discussion on the usage of the stratified-Petersen and Arnason *et al.* (1996) have developed user friendly software to allow users to analyze such experiments. This software also has a simulation and planning capacity to assess the likely degree of bias and precision to be obtained in experiments.

Unfortunately, estimation of the individual run components is much more difficult – even the stratified-Petersen estimates are often unsatisfactory in small samples with poor precision and occasionally “negative” estimates (refer to the example in Schwarz and Taylor, 1998). The problem is that the Schaefer and pooled-Petersen are on one end of the spectrum of models compared to the stratified-Petersen model. The latter simply has too many parameters to be individually estimated well, but because of high correlation among the estimates, estimates of the overall total perform well. Additional research is underway to derive a more constrained stratified-Petersen estimator.

Acknowledgement

This work was supported by Research Grants from the Natural Science and Engineering Research Council of Canada (NSERC). Development of software was supported by Manitoba Hydro.

References

- Arnason, A. N., Kirby, C.W., Schwarz, C.J., and Irvine, J. R. 1996. Computer analysis of data from stratified mark-recovery experiments for estimating salmon escapement and other populations. Cdn. Tech. Rep. Fish. Aquat. Sci. **2106**.
- Banneheka, S. G., Routledge, R.D. and Schwarz, C. J. 1997. Stratified two-sample tag-recovery census of closed populations. Biometrics **53**: 1212-1224.
- Chapman, D. G. and Junge, C. O. 1956. The estimation of the size of a stratified animal population. Ann. Math. Stat. **27**: 375-389.
- Cochran, G. 1977. Sampling techniques. 3rd ed. John Wiley and Sons, New York.
- Darroch, J. N. 1961 The two-sample capture-recapture census when tagging and sampling are stratified. Biometrics **48**: 241-260.
- Eames, M. J., Quinn, T. J., and Hino, M. 1983. 1977 Northern Puget Sound adult coho and chum tagging studies. Wash. Dept. Fish. Prog. Rep. **148**: 1-22.
- Macdonald, P. D. M. and Smith, H. D. 1980. Mark-recapture estimation of salmon smolt runs. Biometrics **36**: 401-417.
- Plante, N., Rivest, L.-P., and Tremblay, G. 1998. Stratified capture-recapture estimation of the size of a closed population. Biometrics **54**: 47-60.

- Ricker, W. E. 1975. Computation and Interpretation of Biological Statistics of Fish Populations. Bulletin of the Fisheries Research Board of Canada, **191**.
- Schaefer, M.B. 1951. Estimation of the size of animal populations by marking experiments. U.S. Fish and Wildlife Service Bulletin **69**: 191-203.
- Schwarz, C. J. and Taylor, C.G. 1998. Use of the stratified-Petersen estimator in fisheries management: estimating the number of pink salmon (*Oncorhynchus gorbuscha*) spawners in the Fraser River. Cdn. J. Fish. Aquat. Sci. **55**: 281-296.
- Seber, G. A. F. 1982. The estimation of animal abundance. 2nd edition. Griffen:London.
- Warren, W. G. and Dempson, J. B. 1995. Does temporal stratification improve the accuracy of mark-recapture estimates of smolt production: a case study based on the Conne River. N. Amer. J. Fish. Manage. **15**: 126-131.

Appendix I. Derivation of bias of Schaefer estimator

A large sample approximate bias for the Schaefer estimator starts by replacing all statistics by their expected value:

$$\hat{N}_{Schaefer} \approx \sum_{i=1}^s \sum_{j=1}^t \frac{(N_{i\bullet} p_i^c) \left(\sum_{i=1}^s N_{i\bullet} \theta_{ij} p_j^r \right) (N_{i\bullet} p_i^c \theta_{ij} p_j^r)}{\left(\sum_{j=1}^t N_{i\bullet} p_i^c \theta_{ij} p_j^r \right) \left(\sum_{i=1}^s N_{i\bullet} p_i^c \theta_{ij} p_j^r \right)} = \sum_{i=1}^s \sum_{j=1}^t N_{ij} \frac{p_i^c p_j^r}{\bar{p}_i^c \bar{p}_j^r}$$

where

$$\bar{p}_j^c = \frac{\left(\sum_{i=1}^s N_{i\bullet} p_i^c \theta_{ij} \right)}{\left(\sum_{i=1}^s N_{i\bullet} \theta_{ij} \right)} = \frac{\left(\sum_{i=1}^s N_{ij} p_i^c \right)}{N_{\bullet j}}$$

$$\bar{p}_i^r = \frac{\left(\sum_{j=1}^t N_{i\bullet} \theta_{ij} p_j^r \right)}{N_{i\bullet}} = \frac{\left(\sum_{j=1}^t N_{ij} p_j^r \right)}{N_{i\bullet}}$$

are the average recovery and recapture probabilities in each column and row respectively.

Now let

$$p_i^c = \frac{\bar{p}^c}{(1 + \delta_i)}, p_j^r = \frac{\bar{p}^r}{(1 + \varepsilon_j)}$$

and using $1/(1+x) \approx 1-x$, substitute into above to give

$$\begin{aligned}
\hat{N}_{Schaefer} &\approx \sum_{i=1}^s \sum_{j=1}^t \frac{N_{ij}}{(1 + \delta_i + \varepsilon_j + \delta_i \varepsilon_j) \left(1 - \frac{\sum_{i=1}^s N_{ij} \delta_i}{N_{\bullet j}} \right) \left(1 - \frac{\sum_{j=1}^t N_{ij} \varepsilon_j}{N_{i\bullet}} \right)} \\
&\approx \sum_{i=1}^s \sum_{j=1}^t N_{ij} \left(1 - \delta_i - \varepsilon_j - \delta_i \varepsilon_j + \frac{\sum_{i=1}^s N_{ij} \delta_i}{N_{\bullet j}} + \frac{\sum_{j=1}^t N_{ij} \varepsilon_j}{N_{i\bullet}} \right) \\
&\approx \sum_{i=1}^s \sum_{j=1}^t N_{ij} \delta_i \varepsilon_j \\
&\approx \frac{\text{cov}(\text{tagging}, \text{recovery})}{E[\text{tagging}]E[\text{recovery}]} = -C(\text{tagging}, \text{recovery}) \frac{\sqrt{V(\text{tagging})V(\text{recovery})}}{E[\text{tagging}]E[\text{recovery}]}
\end{aligned}$$

Appendix II. Variance of the Schaefer estimator

Starting from

$$\hat{N}_{Schaefer} = \sum_{i=1}^s \sum_{j=1}^t \frac{n_i^c n_j^r m_{ij}}{m_{i.} m_{.j}}$$

note that in many studies, $m_{i.} \ll n_i^c, m_{.j} \ll n_j^r$, and following Macdonald and Smith (1980), the m_{ij} have approximate independent Poisson distributions. The number released and recovered can be treated as fixed. Consequently,

$$V(\hat{N}_{Schaefer}) \cong \sum_{i=1}^s \sum_{j=1}^t V(m_{ij}) \left(\frac{\partial \hat{N}}{\partial m_{ij}} \right)^2$$

After taking the partial derivatives, and in order that the estimator is nearly unbiased, assume that $p_i^c \cong \bar{p}^c, p_j^r \cong \bar{p}^r$. Then the messy expression reduces to:

$$V(\hat{N}_{Schaefer}) \cong \sum_{i=1}^s \sum_{j=1}^t \frac{E(m_{ij})}{(\bar{p}^c \bar{p}^r)^2} \approx \frac{N}{\bar{p}^c \bar{p}^r}$$

The coefficient of variation is found as:

$$CV(\hat{N}_{Schaefer}) \cong \frac{\sqrt{V(\hat{N})}}{N} \cong \frac{1}{\sqrt{N \bar{p}^c \bar{p}^r}} \cong \frac{1}{E[m_{..}]}$$

Table 1: Percent relative bias and Root MSE in the pooled-Petersen and Root MSE of the Darroch estimate for a $s = t = 12$ stratum experiment. The entry models had either equal initial stratum sizes (Uniform) or peaked (Normal) at the middle stratum. Sampling probabilities at capture (p^c) and recovery times (p^r) went from low variability across strata (1) to high (6) but with fixed mean capture rates (0.0625) applied to a Small (12,000) and Large (120,000) population. The transition models (ordered A,B,C) are all upper triangular but increasingly spread out (less diagonally dominant). The pooled-Petersen underestimates N in all cases.

<i>Entry Model</i>	<i>Pc</i>	<i>Pr</i>	<i>Small</i>			<i>Large</i>		
			<i>A</i>	<i>B</i>	<i>C</i>	<i>A</i>	<i>B</i>	<i>C</i>

(a) Pooled Peterson Relative Bias

<i>Uniform</i>	1	1	-2.06	-1.48	-0.53	-2.06	-1.48	-0.53
	1	6	-7.28	-6.14	-1.87	-7.28	-6.14	-1.87
	6	1	-7.28	-6.36	-2.29	-7.28	-6.36	-2.29
	6	6	-30.97	-25.73	-9.41	-30.97	-25.73	-9.41
<i>Normal</i>	1	1	-2.53	-1.64	-0.54	-2.53	-1.64	-0.54
	1	6	-9.46	-7.61	-1.94	-9.46	-7.61	-1.94
	6	1	-9.45	-7.91	-2.70	-9.45	-7.91	-2.70
	6	6	-42.39	-34.54	-11.61	-42.39	-34.54	-11.61

(b) Pooled Peterson Relative Root MSE

<i>Uniform</i>	1	1	4.66	4.46	4.23	2.45	1.99	1.43
	1	6	8.23	7.27	4.28	7.38	6.26	2.23
	6	1	8.23	7.45	4.68	7.38	6.48	2.63
	6	6	31.07	25.87	10.00	30.98	25.74	9.47
<i>Normal</i>	1	1	5.09	4.73	4.41	2.89	2.16	1.49
	1	6	10.50	8.84	4.55	9.57	7.74	2.34
	6	1	10.50	9.18	5.61	9.56	8.05	3.12
	6	6	42.47	34.68	12.31	42.40	34.56	11.68

(c) Darroch Relative Root MSE

<i>Uniform</i>	1	1	4.54	4.50	4.39	1.43	1.42	1.39
	1	6	6.78	6.67	5.87	2.15	2.11	1.86
	6	1	6.78	6.67	6.34	2.15	2.11	2.01
	6	6	11.77	11.13	8.94	3.72	3.52	2.83
<i>Normal</i>	1	1	4.87	4.79	4.59	1.54	1.52	1.45
	1	6	8.08	7.72	6.36	2.56	2.44	2.01
	6	1	8.11	7.92	7.46	2.56	2.50	2.36
	6	6	14.59	13.64	10.69	4.61	4.31	3.38

Table 2: Consistency of the Schaefer (Sch.), pooled-Petersen (PPE) and Darroch (D) estimators. Three cases of the Darroch are considered: estimates of total size N when $s = t$; estimates of N and initial stratum sizes N_i when $s < t$; estimates of N and final stratum sizes N_j when $s > t$. Consistency is tested under closure assumptions for death and birth (Pr = Present; Ab = Absent) and when initial (p^c) and final (p^r) sampling probabilities are equal (E) or unequal (U) across strata. A blank cell indicates the estimate is consistent; + or – indicate positive or negative bias, and o indicates cases that were not investigated.

Closure Effect		Sampling Prob.		N			D $s < t$		D $s > t$	
Death	Birth	p^c	p^r	Sch.	PPE	D $s = t$	N	N_i	N	N_j
Ab	Ab	E	E							
		E	U							
		U	E							
		U	U	+	+					
Pr	Ab	E	E							x
		E	U							x
		U	E	+	+				–	x
		U	U	+	+				–	x
Ab	Pr	E	E					x		
		E	U	+	+			x		
		U	E				–	x		
		U	U	–	–		–	x		

Pr	Pr	E	E	+	+	+	o	o	o	o
		E	U	+	+	+	o	o	o	o
		U	E	+	+	+	o	o	o	o
		U	U	+	+	+	o	o	o	o

Figure 1: Statistics collected from the experiment arranged into an array.

Tagging Stratum	Fish tagged	Recovery Stratum				not recovered
		1	2	...	t	
1	n_1^c	m_{11}	m_{12}	...	m_{1t}	$n_1^c - m_{1\bullet}$
2	n_2^c	m_{21}	m_{22}	...	m_{2t}	$n_2^c - m_{2\bullet}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
s	n_s^c	m_{s1}	m_{s2}	...	m_{st}	$n_s^c - m_{s\bullet}$
Total of fish recovered		n_1^r	n_2^r	...	n_{t1}^r	

n_i^c are the number of fish captured, tagged, and released in release stratum i , $i=1, \dots, s$

n_j^r are the number of fish recovered in recovery stratum j , $j=1, \dots, t$.

m_{ij} are the number fish tagged and released in stratum i that are subsequently recovered in stratum j . $i=1, \dots, s$; $j=1, \dots, t$.

A \bullet in a subscript implies summation over that subscript, i.e. $m_{\bullet j} = \sum_{i=1}^s m_{ij}$, $m_{i\bullet} = \sum_{j=1}^t m_{ij}$

Figure 2: The population stratified in a similar fashion as the statistics..

Tagging Stratum	Total Fish	Recovery Stratum				Died or did not move to a recovery stratum
		1	2	...	t	
1	N_1^c	N_{11}	N_{12}	...	N_{1t}	$N_1^c - N_{1\bullet}$
2	N_2^c	N_{21}	N_{22}	...	N_{2t}	$N_2^c - N_{2\bullet}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\underline{s}	N_s^c	N_{s1}	N_{s2}	...	N_{st}	$N_s^c - N_{s\bullet}$
Total	N_\bullet^c	$N_{\bullet 1} = N_1^r$	$N_{\bullet 2} = N_2^r$...	$N_{\bullet t} = N_t^r$	

N_{ij} are the number fish that were present in stratum i at release time and in stratum j at recovery time. $i=1,\dots,s; j=1,\dots,t$.

A \bullet in a subscript implies summation over that subscript, i.e. $N_{\bullet j} = \sum_{i=1}^s N_{ij}$, $N_{i\bullet} = \sum_{j=1}^t N_{ij}$

and $N = \sum_{i=1}^s \sum_{j=1}^t N_{ij}$

Figure 3: Capture probabilities, recovery probabilities, and “movement” probabilities for the stratified experiment.

"Movement" from tagging to recovery strata						
Tagging Stratum	Tagging probability	1	2	...	t	Died or did not move to a recovery stratum
1	p_1^c	θ_{11}	θ_{12}	...	θ_{1t}	$1 - \theta_{1\cdot}$
2	p_2^c	θ_{21}	θ_{22}	...	θ_{2t}	$1 - \theta_{2\cdot}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
s	p_s^c	θ_{s1}	θ_{s2}	...	θ_{st}	$1 - \theta_{s\cdot}$
Recovery		p_1^r	p_2^r	...	p_t^r	

probabilities =>

p_i^c is the probability that a fish present in tagging stratum i will be captured, tagged, and released.

p_j^r is the probability that a fish present in recovery stratum j will be captured.

θ_{ij} is the probability that a fish present in release stratum i will also be present in recovery stratum j . $i=1, \dots, s$; $j=1, \dots, t$.

Figure 4:

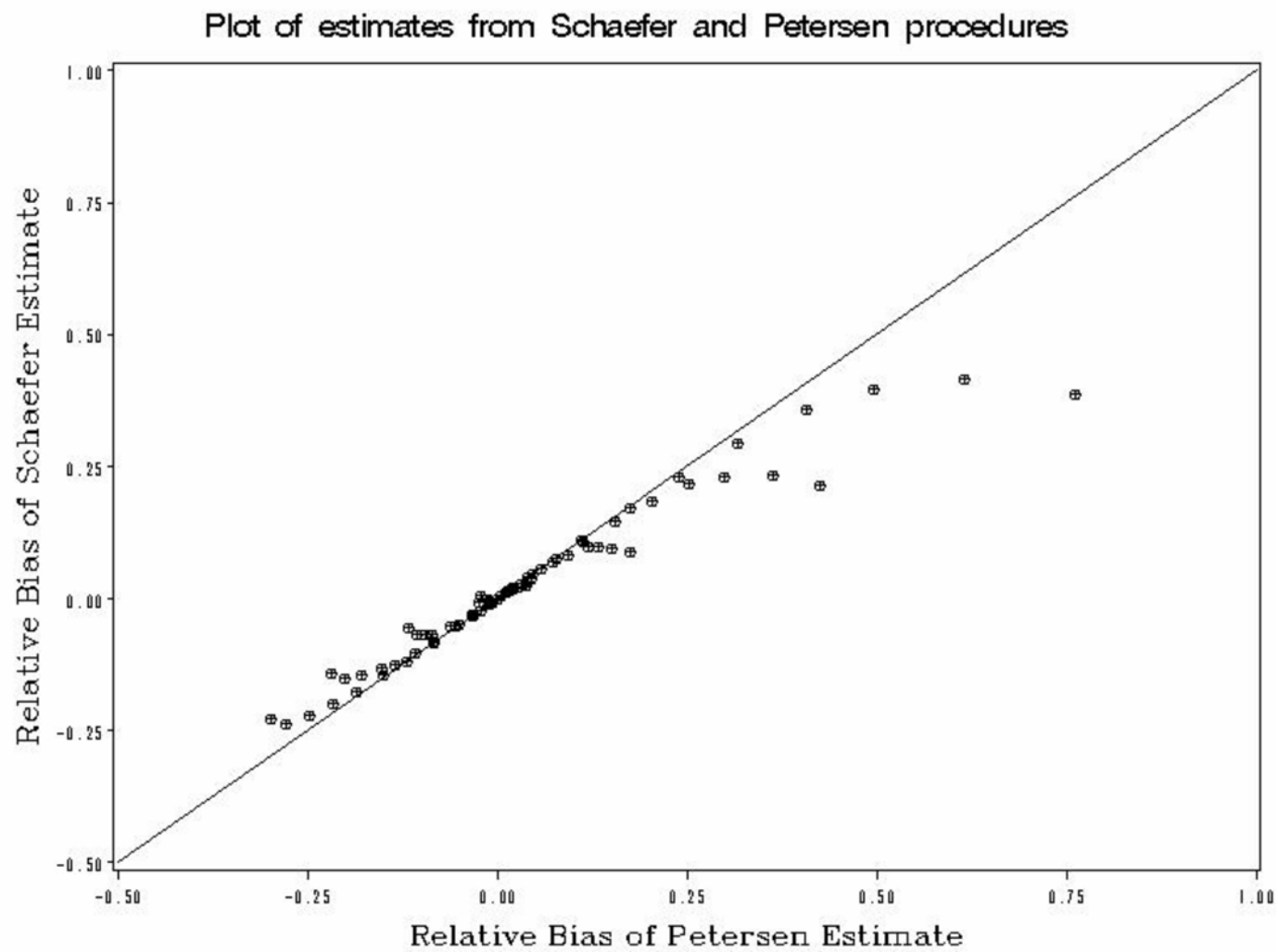


Figure 5:

