Multistage Sampling Designs in Fisheries Research: Applications in Small Streams

David G. Hankin

Department of Fisheries, Humboldt State University, Arcata, CA 95521, USA

Hankin, D. G. 1984. Multistage sampling designs in fisheries research: applications in small streams. Can. J. Fish. Aquat. Sci. 41: 1575–1591.

A common, although generally unrecognized, use of multistage sampling designs in freshwater fisheries research is for estimation of the total number of fish in small streams. Here there are two stages of sampling. At the first stage one selects a sample of stream sections, usually of equal length, and at the second stage one estimates the total number of fish present in each selected section. This paper argues that the conventional practice of selecting stream sections of equal length is ill-advised on both biological and statistical grounds, and that errors of estimation of fish numbers within selected sections will usually be small compared with errors of estimation resulting from expansion of sampled sections to an entire stream. If stream sections are instead allowed to vary in size according to natural habitat units, then alternative two-stage sampling designs may take advantage of the probable strong correlation between habitat unit sizes and fish numbers. When stream sections of unequal sizes are selected with probabilities proportional to their size (PPS), or measures of the sizes of selected sections are incorporated into estimators, one may substantially increase precision of estimation of the total number of fish in small streams. Relative performances of four alternative two-stage designs are contrasted in terms of precision, relative cost, and overall cost-effectiveness. Choice among alternative designs depends primarily on the correlation between fish numbers and habitat unit sizes, on the total number of stream sections, and on sample size. Recommendations for choices among the designs are presented based on these criteria.

Les plans d'échantillonnage à étapes multiples, servent souvent, sans toutefois être généralement reconnus, dans le domaine de la recherche sur les pêches en eau douce pour estimer le nombre total de poissons dans les petits cours d'eau. Ici, l'échantillonnage se fait en deux étapes. Dans un premier temps, on choisit un échantillonnage de sections d'un cours d'eau, habituellement de longueurs égales puis, dans un deuxième temps, on estime le nombre de poissons présents dans chacune des sections choisies. Le présent document avance que la méthode voulant qu'un choisisse des sections du cours d'eau de longueurs égales est mal inspirée tant du point de vue biologique que statistique et que les erreurs que comporte l'estimation du nombre de poissons dans les sections choisies seront généralement petites comparativement à celles résultant de l'extrapolation des résultats à la totalité du cours d'eau. Si, au contraire, on fait varier la longueur des sections en fonction des habitats naturels, alors les plans d'échantillonnage à deux étapes pourront tenir compte de la forte corrélation qui existe probablement entre la taille de l'habitat et le nombre de poissons. Il est possible d'augmenter de façon substantielle la précision du calcul du nombre total de poissons dans les petits cours d'eau en choisissant des sections du cours d'eau de longueurs inégales présentant une probabilité à leur taille (PPT), ou en incorporant dans les estimateurs les mesures relatives à la taille des sections choisies. On compare la performance relative de quatre plans à deux étapes pour ce qui est de la précision, du prix de revient relatif et du rendement global en termes de coût-efficacité. Le choix du plan dépendra surtout de la corrélation entre le nombre de poissons et la taille de l'habitat, du nombre total de sections du cours d'eau considérées et de la taille de l'échantillon. En se basant sur ces critères, on a formulé des recommandations sur la façon de choisir le plan à utiliser.

Received October 27, 1983
Accepted July 20, 1984
(J7590)

Reçu le 27 octobre 1984
Accepté le 20 juillet 1984

Many applications of statistical survey (sampling) theory in marine fisheries research have involved use of multistage sampling designs (Abramson 1968; Tomlinson 1971; Southward 1976; Schweigert and Sibert 1983). In multistage sampling there are two or more levels (stages) of sample selection. For example, suppose one wished to estimate the mean length of a commercial fish species landed at a particular port (over some specified period of time). The first stage of sampling might be selection of a particular sample of fishing vessels which lands catch of that species at that port. Given a first stage selection of vessels, one must sample from the vessels' catches. If fish are unloaded in large bins, then one selects a particular sample of bins at the second stage of sampling (from each of those vessels selected at the first stage). If the number of fish within selected bins is large, then subsampling of bins may be required; within each selected bin a third stage sample of fish may be drawn and on these fish actual measurements of fish length would be made.

In the simplest case, equal probability selection methods are used at each stage of sampling: each vessel entering a port has an equal chance of inclusion at the first stage, each bin from a selected vessel has an equal chance of inclusion at the second stage, and each fish within a selected bin has an equal chance of being measured at the third stage. Either simple random sampling (SRS) or systematic sampling could be used at each stage, although systematic sampling (e.g. one in every k vessels) may often prove to be more practical. Each stage of sampling therefore requires choice of selection method. These choices determine which multistage estimators are appropriate and may, in general, strongly influence the variance and/or bias of resulting estimates (here, of mean length). The term multistage sampling design is used to define a specific collection of selection procedures (used at each stage of sampling) and their associated estimators of attributes and of the variances of those estimated attributes.

One must also choose the sample sizes at each stage of selection; in sampling theory jargon, one must choose the sampling fractions at each stage. What fraction of the vessels should be included at the first stage? What fraction of a selected vessel's bins should be included at the second stage? What fraction of those fish within selected bins should be measured? Providing answers to these kinds of questions has often been the principal focus of papers devoted to the use of multistage sampling designs in fisheries research (Schweigert and Sibert 1983; also, Cuff and Coleman 1979 for an example using stratified sampling). Given estimates of variance at each stage of sampling, one may determine optimal sampling fractions at each stage based on criteria of (a) minimizing the variance of resulting estimates at fixed cost or (b) achieving a specified variance at minimum cost (see Cochran 1977, p. 313–316).

Less attention has been given to selection method per se or to the use of auxiliary variables in fisheries research (see Lenarz and Adams 1980 for a comparison of SRS, systematic, and stratified SRS in groundfish trawl surveys). Unequal probability selection methods, which often rely on information provided by some inexpensively gathered auxiliary variable, seem especially worthy of attention. In the preceding example, selection of vessels according to (the auxiliary variable) vessel size may increase the precision of resulting estimates and do so in a cost-effective manner. Because the number of fish landed may be highly correlated with a vessel's size, landings from a large vessel will usually have more influence on mean fish length than will landings from a small vessel. By setting first stage selection probabilities proportional to a measure of vessel size (PPS), one creates a selection procedure whereby larger vessels are more likely to be included in first stage selections than are smaller vessels (see Tomlinson 1971). SRS is a special case of such unequal probability sampling (Kendall and Stuart 1983, p. 189–195).

Multistage sampling designs have seen less frequent use in freshwater fisheries research. However, a very common, although generally unrecognized, use is for estimation of the total number of fish in small streams. In this context there are two stages of sampling. At the first stage, a particular set of stream sections, usually of equal length, is selected (usually by SRS). Within any selected stream section, some population estimation techniques, most frequently a removal method estimator (Seber 1982, Sect. 7.2) based on electrofishing, is used to estimate the total number of fish present. Because there are no discrete sampling units at the second stage of sampling (bins were obvious discrete sampling units in the commercial fishery example), multistage sampling theory has not been previously used to explicitly address estimation of the total number of fish in small streams. However, one may make an analogy between (a) estimating the total number of fish present in a stream section (a first stage or primary unit) by electrofishing and (b) estimating the mean length of fish from a particular vessel by subsampling of a vessel's bins.

To make this analogy obvious, assume that there are only two stages of sampling for the commercial fishery example: all fish within selected bins are measured. Then, the bounds of error of estimation of mean fish length from a particular vessel arises from sampling variance: the variation among all those possible estimates of mean fish length derived from all of the possible samples (finite in number) of n bins selected from the N bins on that vessel. Because this variation results at the second stage of sampling, it is termed second stage variance (within a selected primary unit). When electrofishing is used to estimate the total number of fish in a selected primary unit (the primary unit total), second-stage variance depends on the variance of the sampling distribution of the removal method estimator. This sampling distribution may be visualized as a plot of the relative frequency of particular estimates of the primary unit total. However, to obtain such a plot one must imagine repeating the electrofishing/removal method process an infinite number of times, while assuming a known primary unit total and a known and constant capture probability (the chance that a fish which is present at the time of electrofishing will be captured by electrofishing). Determination of second stage variance for this kind of sampling distribution, in contrast with the case of sampling from a discrete number of bins, requires construction of a formal stochastic model and use of specific probability distribution assumptions (Seber 1982, Sect. 7.2; Schnute 1983).

Bohlin (1981) recently treated the usual two-stage stream survey problem in terms of sources of error: (a) error arising from variation among (estimated) primary unit totals and (b) "measurement" error within selected primary units. The variance formulae he derived are equivalent to those presented by Raj (1968, p. 116–119) and Cochran (1977, p. 300–303) for two-stage sampling of equal-sized primary units if electrofishing sampling variance is equated with second stage variance. However, Bohlin did not view this problem in terms of multistage sampling theory.

In this paper it is argued that (a) the usual practice of selecting sections of equal length is ill-advised on both biological and statistical grounds and that (b) a preferable practice is to allow stream sections to vary in size according to natural habitat units. If primary units (stream sections) are of unequal sizes, then several alternative two-stage sampling designs may be used in the context of stream surveys. In particular, (a) primary units may be selected with probabilities proportional to primary unit sizes and (b) the use of an auxiliary variable (here, primary unit size) may result in substantial improvements in the precision of estimation of the total number of fish in small streams. Average costs for alternative designs may be computed and the cost-effectiveness of alternative designs compared. Although alternative selection procedures and the benefits of using auxiliary variables will be demonstrated in the narrow context of small stream surveys, the general principles and procedures used should be applicable in a broad range of fisheries contexts, both freshwater and marine.

Sampling Designs

Notation and definitions for all sampling designs presented in this paper are summarized below:

Notation	Definition
N	total number of primary units in the sampling universe
n	total number of primary units in the sample
Y_i	total in primary unit i; estimated as $\hat{Y}_i$
$Y = \sum^{N} Y_i$	total across all primary units
$\bar{Y} = Y/N$	mean of all primary unit totals; estimated as $\hat{\bar{Y}} = \sum^{n} \hat{Y}_i/n$
σ_i^2	second stage variance within primary unit i; estimated as $\hat{\sigma}_i^2$
M_i	size of primary unit i
$M_0 = \sum^{N} M_i$	total size of all primary units
$\bar{Y}_i = Y_i/M_i$	mean per unit of size in primary unit i; estimated as $\hat{\bar{Y}}_i = \hat{Y}_i/M_i$
$\bar{\bar{Y}} = Y/M_0$	overall mean per unit of size; estimated as $\hat{\bar{\bar{Y}}} = \sum^{n} \hat{Y}_i \Big/ \sum^{n} M_i$
$p_i = M_i/M_0$	probability of selecting the ith unit on a given draw when units are selected by PPS with replacement
π_i	probability that the ith primary unit is in a sample of size n drawn from N by PPS without replacement
π_{ij}	probability that both primary units i and j are in a sample of size n drawn from N by PPS without replacement

An estimate is in all cases distinguished from a true value by a circumflex above. For variances, $\hat{V}(\hat{Y})$ distinguishes a sample-based estimator of variance from the true sampling variance of the estimated total, $V(\hat{Y})$. Similar notation is used to distinguish sample-based estimators and true mean square error (= variance + bias^2). Unless otherwise specified, summations are implicitly over primary units in the sample ($i = 1, 2, \ldots, n$) or in the sampling universe ($i = 1, 2, \ldots, N$). The sampling universe consists of the total number of primary units and their respective attributes from which the sample is drawn by some selection method.

The Usual Case: Primary Units of Equal Sizes Selected by SRS

If primary units of equal sizes are selected by SRS, a removal method estimator is used to estimate selected primary unit totals, and sampling variance of the removal method estimator is equated with second stage variance, then (Cochran 1977, p. 300–303; Raj 1968, p. 116, 119):

$$(1) \quad \hat{Y} = \frac{N}{n} \sum^{n} \hat{Y}_i$$

$$(2) \quad V(\hat{Y}) = \frac{N(N-n) \sum^{N} (Y_i - \bar{Y})^2}{n(N-1)} + \frac{N \sum^{N} \sigma_i^2}{n}$$

$$(3) \quad \hat{V}(\hat{Y}) = \frac{N(N-n) \sum^{n} (\hat{Y}_i - \hat{\bar{Y}})^2}{n(n-1)} + \frac{N \sum^{n} \hat{\sigma}_i^2}{n}.$$

The first term in equation 2 accounts for variation among primary unit totals (first stage variance), whereas the second term accounts for variation arising from (sub)sampling within selected primary units (second stage variance). Total sampling variance is the sum of first and second stage variances. When unbiased estimators are used at the second stage of sampling, equations 1–3 give exact and unbiased results. However, because removal method estimation of a primary unit total ($\hat{Y}_i$) is only asymptotically unbiased, and estimators for the variance of an estimated total ($\hat{\sigma}_i^2$) are only approximate (Appendix A), equations 1 and 3 are also only approximate. Equations 1 and 3 are equivalent to formulas derived by Bohlin (1981).

Although the above formulas have (approximate) statistical validity, there are serious biological problems with the convention of selecting primary units of equal sizes. If primary units are equal-length sections of stream, then a selected unit may contain more than one habitat type, and only portions of discrete, natural habitat units (e.g. portions of a pool and/or of a riffle) may be included in a selected unit. For removal method estimation using electrofishing, one delimits the primary unit boundaries by first placing block nets at the upstream and downstream ends of the selected stream section. Placement of these block nets may result in considerable displacement of fish from the primary unit to be sampled when the ends of a selected unit fall, say, midway in a pool or riffle. Also, it will sometimes be physically impossible to place block nets at the selected locations: a pool may simply be too deep. Such nonsampling error may seriously bias estimation of primary unit totals and may seriously compromise the validity of any further inferences.

There are at least two obvious ways by which to circumvent the above problems. Having identified the selected unit and having recognized that the upper end of the unit passes through the deepest part of a pool, say, one could "move the section upstream" until the entire pool, or most of it, was included. Because this alternative involves a purposive action, it destroys both the intent and statistical validity of any primary unit selection method. Alternatively, one could at least ensure that only a single habitat type would be included in any selected primary unit. If a stream were mapped and then stratified by habitat type into, say, pools and riffles, then these two strata could be independently sampled. However, this alternative would not eliminate those nonsampling errors identified above: a section could still fall midway within a pool or riffle.

If primary units were instead allowed to vary in size according to the sizes of natural habitat units, then several advantages would result. First, those nonsampling errors identified above would be minimized. Placement of block nets at the upper and lower ends of entire pools or riffles would be far less likely to displace significant numbers of fish from the unit to be sampled. Second, the numbers and sizes of fish estimated to be present in particular pools or riffles could be related to the sizes of these natural habitat units. Finally, when primary units are of unequal size, multistage sampling theory offers a variety of alternative two-stage designs with which to estimate the total number of fish in small streams.

The remainder of this paper is devoted to a presentation of four such alternative sampling designs and to a consideration of their advantages, limitations, and costs. It is assumed that stream habitat has first been mapped and stratified into habitat strata. The four sampling designs thus deal only with estimation of the total number of fish within a particular habitat stratum. Because all stratum estimates are independent of one another, this introduces no complications or restrictions to the procedures presented and contrasted. Given independent estimates of the total number of fish in the hth stratum, $\hat{Y}_h$, and associated sample-based variances for the estimated stratum totals, $\hat{V}(\hat{Y}_h)$, estimates of stratum totals and variances are strictly additive (Raj 1968, p. 123):

$$\hat{Y} = \sum^{L} \hat{Y}_h;\ \hat{V}(\hat{Y}) = \sum^{L} \hat{V}(\hat{Y}_h);\ h = 1, 2, \ldots, L.$$

In the following, all stratum-specific notation has been avoided; notation follows that presented in the preceding summary of notation and definitions.

Alternative Designs: Primary Units of Unequal Sizes

The four alternative two-stage designs, all based on primary units of unequal sizes, may be classified by selection method and by their use (or lack of use) of an auxiliary variable (here, primary unit size):

Design A: SRS — no auxiliary variable used

Design B: SRS/ratio estimation — uses an auxiliary variable

Design C: PPS — sampling with replacement

Design D: PPS — sampling without replacement

For both PPS designs, first stage selection probabilities are based on the auxiliary variable, primary unit size. It is assumed for all designs that (a) primary unit totals are estimated by the removal method based on electrofishing and that (b) electrofishing sampling variance is equated with second stage variance. Note, however, that other estimators of primary unit totals, e.g. mark–recapture estimators, could be used at the second stage of sampling. All formulas would remain valid.

Design A: SRS

If primary units of unequal sizes are selected by SRS, but one does not employ an auxiliary variable, then equations 1–3 are again appropriate (Cochran 1977, p. 300–303). If one had unbiased estimators at the second stage of sampling, equations 1 and 3 would again be exact and unbiased.

When primary unit totals are highly correlated with primary unit sizes, first stage variance will be large for this design. In this case, individual primary unit totals, Y_i, will be very different from the average primary unit total, $\bar{Y}$, and the squared differences between Y_i and $\bar{Y}$ will make the first term in equation 2 large.

Design B: SRS/ratio estimation

If primary units of unequal sizes are selected by SRS and one also incorporates a measure of the sizes of selected primary units (M_i) in estimators, then a two-stage ratio estimator may be used (Cochran 1977, p. 300–305):

$$(4) \quad \hat{Y} = M_0 \sum^{n} \hat{Y}_i \Big/ \sum^{n} M_i$$

$$(5) \quad \mathrm{MSE}(\hat{Y}) \approx \frac{N(N-n) \sum^{N} M_i^2(\bar{Y}_i - \bar{\bar{Y}})^2}{n(N-1)} + \frac{N \sum^{N} \sigma_i^2}{n}$$

$$(6) \quad \widehat{\mathrm{MSE}}(\hat{Y}) = \frac{N(N-n) \sum^{n} M_i^2(\hat{\bar{Y}}_i - \hat{\bar{\bar{Y}}})^2}{n(n-1)} + \frac{N \sum^{n} \hat{\sigma}_i^2}{n}.$$

Because ratio estimation is biased, mean square error is used as a measure of precision rather than variance. However, use of the auxiliary variable, M_i, in equations 4–6 may dramatically increase precision of estimation over equations 1–3 when (a) primary unit totals are proportional to primary unit sizes and (b) the variation in primary unit totals increases in proportion to primary unit size. In biological terms, condition (a) means that the total number of fish in a selected habitat unit should be proportional to the size of the habitat unit; the larger a pool, the greater the number of fish. When the above conditions are (exactly) met, the ratio estimator is the best linear unbiased (BLUE) estimator. When the above conditions are approximately met, then the first term in equation 5 becomes small because the average number of fish per unit of habitat size within any selected primary unit ($\bar{Y}_i$) is a very stable quantity, nearly independent of primary unit size, and is approximately equal to the average number of fish per unit of habitat size in the sampling universe ($\bar{\bar{Y}}$). However, Cochran (1977, p. 162–164) noted that sample-based estimates of mean square error (equation 6) may have serious negative bias when the number of selected primary units is small ($n < 12$). Equations 5 and 6 are both large-sample approximate results.

Design C: PPS — Sampling with replacement

For this design, primary units are selected with replacement with probabilities $p_i = M_i/M_0$. Raj (1968, p. 119–121) derived estimators for this case that require independent sampling of a primary unit if it is selected more than once in the sample; no single estimate of a primary unit total may be used more than once:

$$(7) \quad \hat{Y} = (1/n) \sum^{n} \hat{Y}_i/p_i$$

$$(8) \quad V(\hat{Y}) = (1/n) \sum^{N} p_i(Y_i/p_i - Y)^2 + (1/n) \sum^{N} \sigma_i^2/p_i$$

$$(9) \quad \hat{V}(\hat{Y}) = [1/n(n-1)] \sum^{n} (\hat{Y}_i/p_i - \hat{Y})^2.$$

Equations 7–9 are unbiased when one uses unbiased estimators at the second stage of sampling; if electrofishing is used at the second stage, equations 7 and 9 will be approximate.

PPS with replacement designs will perform best when Y_i are proportional to M_i. Then, $p_i = M_i/M_0 = Y_i/Y$, and the expected value of $Y_i/p_i = Y$ for all i; hence, the first term in equation 8 becomes zero. However, the possibility that the same primary unit may be included more than once in a sample means that PPS with replacement selection of primary units will usually be less efficient than PPS without replacement selection. When the sampling universe is large and the number of selected primary units is small, efficiency will be comparable with PPS without replacement designs. Chief virtues of PPS with replacement are (1) computation of selection probabilities is simple because the p_i are independent of sample size and (2) the sample-based estimator of variance (equation 9) does not require estimates of second stage variance.

Design D: PPS — Sampling without replacement

For this design, primary units are selected without replacement with probabilities proportional to their sizes according to one of many possible selection methods (see Hanif and Brewer 1980 for a review of 50 published unequal probability selection methods). Let π_i = probability that the ith primary unit will be in a sample of size n, and let π_{ij} = probability that units i and j will be in the sample ($i \neq j$). Assuming that unbiased estimators are used at the second stage of sampling, Raj (1968, p. 118–119) proved that unbiased two-stage estimators are

(10) $\hat{Y} = \sum^{n} \hat{Y}_i/\pi_i$

(11) $V(\hat{Y}) = \sum_{i}^{N-1} \sum_{j>i}^{N} (\pi_i\pi_j - \pi_{ij})(Y_i/\pi_i - Y_j/\pi_j)^2 + \sum^{N} \sigma_i^2/\pi_i$

(12) $\hat{V}(\hat{Y}) = \sum_{i}^{n-1} \sum_{j>i}^{n} \frac{(\pi_i\pi_j - \pi_{ij})}{\pi_{ij}} (\hat{Y}_i/\pi_i - \hat{Y}_j/\pi_j)^2 + \sum^{n} \hat{\sigma}_i^2/\pi_i$

with the restrictions that all $\pi_{ij} > 0$ and that

$$\sum^{N} \pi_i = n \quad \text{and} \quad \sum_{i}^{N-1} \sum_{j>i}^{N} \pi_{ij} = n(n-1)/2.$$

Summations are over all possible distinct pairs of primary units in the sampling universe (equation 11) or in a particular sample of size n (equation 12). When electrofishing is used at the second stage of sampling, equations 10 and 12 are approximate.

Because primary units are selected without replacement, the efficiency of this design will almost always exceed that of a PPS with replacement design applied to the same sampling universe. This design will also perform best when Y_i are proportional to M_i. If the selection method ensures that π_i are proportional to M_i (not all methods will ensure this (see Appendix B)), then $Y_i/\pi_i = Y_j/\pi_j$ for all i, j ($i \neq j$) and the first term in equation 11 becomes zero. For sample sizes exceeding 2, computation of the π_{ij} is not a simple task and requires use of a computer. Appendix B presents details of two PPS without replacement selection methods, whose performances are contrasted in the following section, and of formal properties of equations 11 and 12.

Performance of Alternative Designs

Relative performances of the four alternative designs depend primarily on the following conditions: (1) the number of primary units that are selected, n, and the size of the sampling universe, N, (2) the degree to which primary unit totals are proportional to primary unit sizes, measured roughly by the correlation between Y_i and M_i, (3) the range of sizes of primary units in the sampling universe, and (4) survey costs associated with electrofishing within selected primary units, as opposed to those costs that are independent of the sizes of those primary units that are selected. No single design will perform best under all possible conditions. Design A (SRS) may perform best when primary unit totals are poorly correlated with primary unit sizes, whereas design D (PPS without replacement) may perform best when primary unit totals are highly correlated with primary unit sizes. Comparison of the relative performances of alternative designs therefore requires their application to specific sampling universes; results depend on the characteristics of those sampling universes used as bases for comparison.

Construction of sampling universes

Sampling universes formally consist of unique sets of primary units with sizes and totals known exactly and are usually of two types: natural or artificial. Natural sampling universes have a real-life existence, whereas artificial sampling universes are imaginary and constructed. It is impossible to satisfy the formal requirements for a natural sampling universe in this stream survey context because (1) numbers of fish in primary units are always estimated rather than enumerated, (2) only a small fraction of primary units within a stream is ever sampled, and (3) selected primary units are almost always of equal sizes. However, one may arrive at a compromise between a natural and an artificial sampling universe by treating available estimates of primary unit totals and sizes as if they fulfilled the requirements of a natural sampling universe.

TABLE 1. Pool sizes (m^2) and estimated population sizes for pools sampled during 1981 and 1982 in Knowles Creek, Oregon, by the U.S. Forest Service.

1981		1982	
Pool size	Estimated population	Pool size	Estimated population
6686	2875	2641	5547
4757	1142	649	779
520	52	340	68
302	175	255	153
219	129	210	42
186	39	180	144
179	159	152	76
108	130	145	217
79.4	127	70.0	70
44.2	42	46.4	51
31.3	47	40.0	28
16.0	16	32.7	49
12.0	12	22.0	22
9.60	23	6.96	16
9.44	17	4.29	9

Estimated population sizes ($\hat{Y}_i$) of yearling coho salmon (*Oncorhynchus kisutch*) and surface areas (M_i) of entire pools sampled in a small Oregon coastal stream in each of two years (Table 1) were used to construct three pairs of such compromise sampling universes. For each pair of constructed sampling universes (1981 and 1982), the total number of sampled primary units (pools) was treated as the total size (N) of a small sampling universe, and estimated primary unit totals and sizes were treated as if they were exactly known totals and sizes. Constructed sampling universes were deliberately designed to differ in two principal respects, (1) the correlation between Y_i and M_i and (2) the range of Y_i and M_i, in order to present a useful contrast of the relative performances of alternative designs under differing conditions. Because these three pairs of constructed sampling universes were all small ($N \leqslant 15$), an additional, large ($N = 50$) artificial sampling universe was also constructed, based in part on Table 1. Primary unit totals and sizes for this large sampling universe are presented in Table 2.

Calculation of net relative efficiencies

The most useful single measure of the performance of alternative designs is net relative efficiency. Letting $V(\hat{Y})$ denote sampling variance and C denote total costs, then, for a fixed sampling fraction, n/N, the net relative efficiency of design B compared with design A is (Jessen 1978, p. 80, 112)

$$\text{NRE(B/A)} = V_A(\hat{Y})C_A/V_B(\hat{Y})C_B.$$

Net relative efficiency balances the relative efficiency of design B compared to design A (RE(B/A) = $V_A(\hat{Y})/V_B(\hat{Y})$) against the relative cost of design B compared with design A (RC(B/A) = C_B/C_A). For example, if design B is twice as precise as design A (RE(B/A) = 2), but total costs for design B are three times that for design A (RC(B/A) = 3), then the net relative efficiency of design B compared with design A is 2/3. Net relative efficiency is less than 1 because improvements in precision are more than

TABLE 2. Artificial sampling universe giving pool sizes and population sizes for all pools (constructed from those data presented in Table 1).

Pool no.	Pool size	Population
1	649	779
2	520	52
3	430	377
4	340	68
5	321	144
6	302	175
7	279	100
8	255	153
9	237	212
10	219	129
11	215	120
12	210	42
13	198	133
14	186	39
15	180	137
16	180	144
17	179	159
18	166	76
19	152	76
20	149	46
21	145	217
22	127	57
23	108	130
24	93.7	86
25	79.4	127
26	74.7	46
27	70.0	70
28	58.2	43
29	46.4	51
30	45.3	48
31	44.2	42
32	42.1	34
33	40.0	28
34	36.4	22
35	32.7	49
36	32.0	12
37	31.3	47
38	26.7	18
39	22.0	22
40	19.0	14
41	16.0	16
42	14.0	21
43	12.0	12
44	10.8	15
45	9.60	23
46	9.52	17
47	9.44	17
48	8.20	11
49	6.96	16
50	4.29	9

offset by increases in cost for design B: design B is less cost-effective than design A.

Calculation of net relative efficiency therefore requires calculation of (a) sampling variance and (b) total survey costs. In the example results that follow, equations 2, 5, 8, and 11 were used to calculate sampling variances for designs A, B, C, and D, respectively. For the PPS without replacement design (D), sampling variances were calculated for each of two different selection methods. For one of these selection methods, the magnitude of necessary computations limited maximum sample size to six for the small sampling universes and to three for the large sampling universe (see Appendix B). Calculations of second stage variances (σ_i^2) assumed use of a removal method estimator based on two passes of equal effort with electrofishing capture probability, q, set equal to 0.5 (see Appendix A; and Seber 1982, p. 318) and assumed to be independent of the sizes of the primary units. The possibility that q may be inversely related to primary unit size is also considered in Appendix A.

Based on conversations with biologists responsible for collecting those data presented in Table 1, relative costs for alternative designs were based on an assumption that when SRS was used to select primary units, half of survey costs were attributed to actual time spent electrofishing. Remaining costs, for housing and per diem for personnel, for time spent travelling to and from the study stream and between selected primary units, and for time spent setting up and taking down block nets prior to and after electrofishing, were assumed to be independent of the sizes of selected primary units. The time and cost spent to electrofish primary unit i was assumed to be proportional to M_i^b. For $b = 1$, time spent electrofishing primary unit i is directly proportional to actual primary unit size. For $b > 1$, larger units would require relatively more time than would be indicated by their actual sizes, M_i. Because in most situations b probably exceeds 1, b was set equal to either 1 or 1.5 for comparative purposes. For $b = 1.5$, M_i^b gives the *effective* size of primary unit i with respect to necessary electrofishing time. Thus, for example, the effective sizes of two primary units of actual sizes 100 and 1000 would be 1000 and 31 623; the ratio of actual sizes is 1:10, but that of effective sizes is about 1:32. Because M_i have been measured in terms of square metres, $M_i^{1.5}$ is a rough proxy for the volume of water that must be electrofished in primary unit i. Details of cost calculations are presented in Appendix C.

Because PPS designs (C and D) assign higher selection probabilities to larger primary units, the expected total sizes of n selected primary units for the PPS designs will always exceed those for the two SRS designs (A and B). Thus, for net relative efficiency of PPS designs compared with SRS designs to exceed 1, it is not sufficient that relative efficiency exceed 1; improvements in precision must more than compensate for increases in expected total costs for these designs. Net relative efficiencies (with respect to design A) were calculated for both actual and effective primary unit sizes for all designs and for both PPS without replacement selection methods, with one exception. For the large sampling universe, net relative efficiencies were calculated only for actual primary unit sizes ($b = 1$).

Results

For brevity in presentation of the relative performances of the four alternative two-stage sampling designs, the following notation has been adopted: (1) $V_A(\hat{Y})$, $V_B(\hat{Y})$, $V_C(\hat{Y})$, and $V_D(\hat{Y})$ denote sampling variances for design A (SRS), design B (SRS/ratio estimation), design C (PPS with replacement), and design D (PPS without replacement); (2) RC(B), RC(C), and RC(D) denote relative costs for designs B, C, and D (with respect to design A); and (3) NRE(B), NRE(C), and NRE(D) denote net relative efficiencies of designs B, C, and D (with respect to design A).

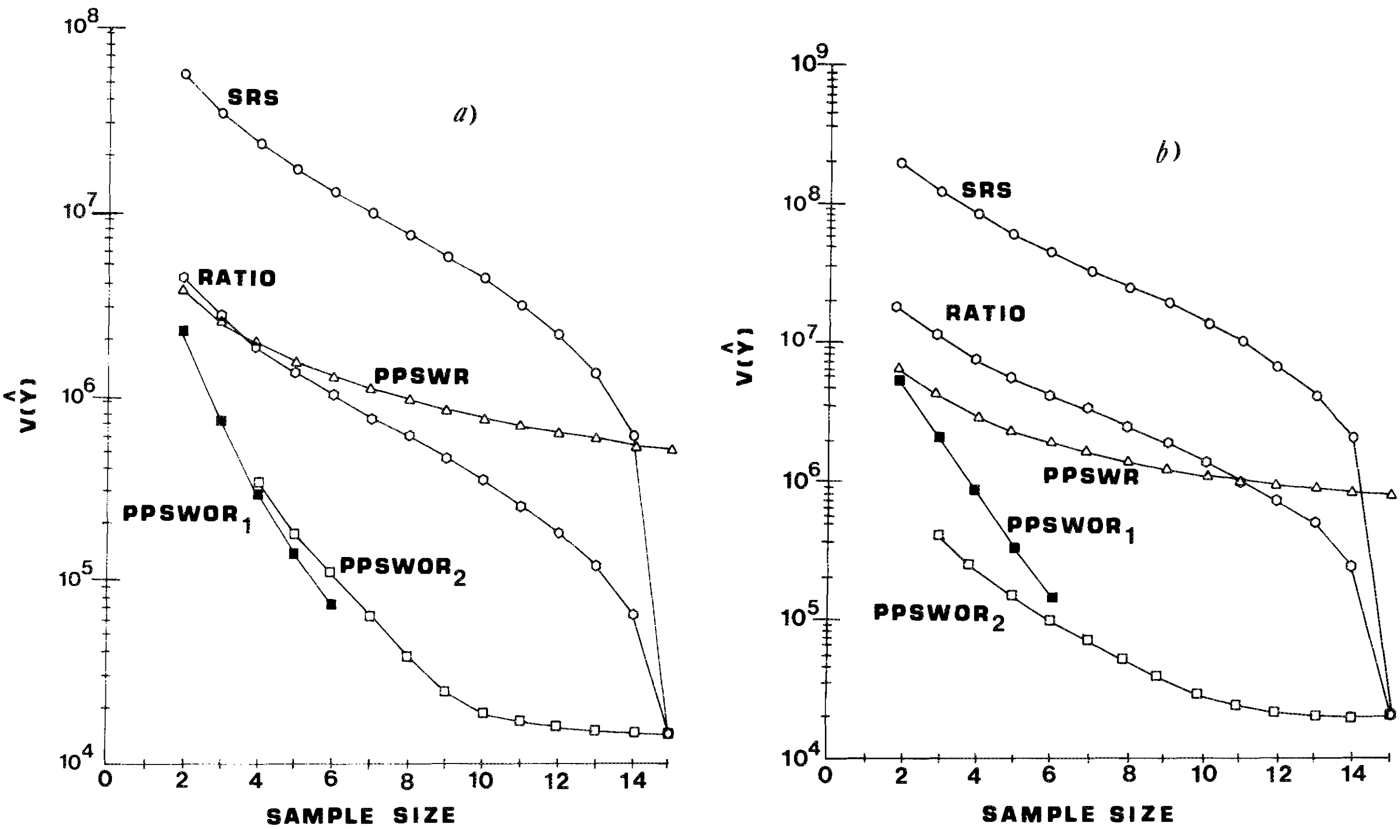

FIG. 1. Sampling variances, $V(\hat{Y})$, for SRS (SRS), SRS/ratio estimation (RATIO), PPS with replacement (PPSWR), and PPS without replacement (PPSWOR) designs plotted against sample size for example 1 in (a) 1981 and (b) 1982. Subscripts on PPSWOR indicate selection methods 1 and 2. Sample sizes for which more than one $\pi_{ij} = 0$ for the second selection method are not plotted. For the second selection method: 1981: $\pi_{10,11} = 0$ ($n = 10$), $\pi_{12,13} = 0$ ($n = 12$); 1982: $\pi_{3,4} = 0$ ($n = 3$), $\pi_{9,10} = 0$ ($n = 9$), $\pi_{13,14} = 0$ ($n = 13$), $\pi_{14,15} = 0$ ($n = 14$).

Example 1

When all data in Table 1 are used to construct two sampling universes, then $N = 15$ (for each) and the correlations between Y_i and M_i are 0.962 and 0.988 for 1981 and 1982. These strong correlations and the extreme ranges in primary unit sizes, 9.44–6686 and 4.29–2641 m^2, a priori suggest that designs B, C, and D should outperform design A.

Figure 1 shows that, for both 1981 and 1982, $V_A(\hat{Y})$ exceeds that for all other designs for $n < N$. $V_B(\hat{Y})$ is considerably less than $V_A(\hat{Y})$ and, for large n relative to N, $V_B(\hat{Y}) < V_C(\hat{Y})$. Design D (for both selection methods) is most precise for $n \geq 2$, and sampling variance declines most rapidly with increasing sample size for this design.

The penalty paid for the improved precision of the PPS designs (C and D) is a substantial increase in relative costs (Fig. 2). However, relative costs for the PPS without replacement design decline rapidly with increasing sample size. As a result of the high and constant relative cost of design C (see Appendix C), NRE(B) > NRE(C) for nearly all sample sizes for both years and for both actual and effective sizes of primary units. NRE(C) declines linearly with increasing sample size and, for $n \approx N$, is less than 1. For $n > 2$, NRE(D) exceeds that of any other design for both years and for both actual and effective primary unit sizes. Further, NRE(D) increases with increasing sample size until more than half of the primary units have been sampled; for all other designs, net relative efficiencies decline with increasing sample size (Fig. 3). Relative performances of the alternative designs are very similar for both 1981 and 1982, as is the case for all other examples, suggesting that relative performances may be fairly stable across years. For these two sampling universes, then, Design D (PPS without replacement) would be the clear design of choice (among the four alternatives) for $n > 2$. For $n = 2$, designs B and D have similar net relative efficiencies. (However, design B is not recommended for small n; see Discussion.)

Example 2

If the three largest primary units are removed from data sets in Table 1 for both years, then $N = 12$, the correlations between Y_i and M_i are reduced to 0.794 and 0.694, and the ranges of primary unit sizes decrease to about two rather than three orders of magnitude. $V_B(\hat{Y})$ is again less than $V_A(\hat{Y})$ for all $n < N$, but $V_C(\hat{Y})$ always exceeds $V_B(\hat{Y})$ for 1981 and does so for $n > 3$ in 1982. $V_D(\hat{Y})$ is smallest for all sample sizes (Fig. 4).

Because the range in sizes of primary units has been reduced, the relative costs for the PPS designs are much lower than for example 1 (Fig. 5). NRE(C) again declines linearly with increasing sample size and is less than 1 for $n > 3$ in most cases. NRE(B) exceeds 1 for all $n < N$ for both years and for both actual and effective primary unit sizes. NRE(D) exceeds NRE(B) for all sample sizes exceeding 4. For $n \leq 4$, either design B or design D has greatest net relative efficiency (Fig. 6). (However, design B is not recommended for small n; see Discussion.)

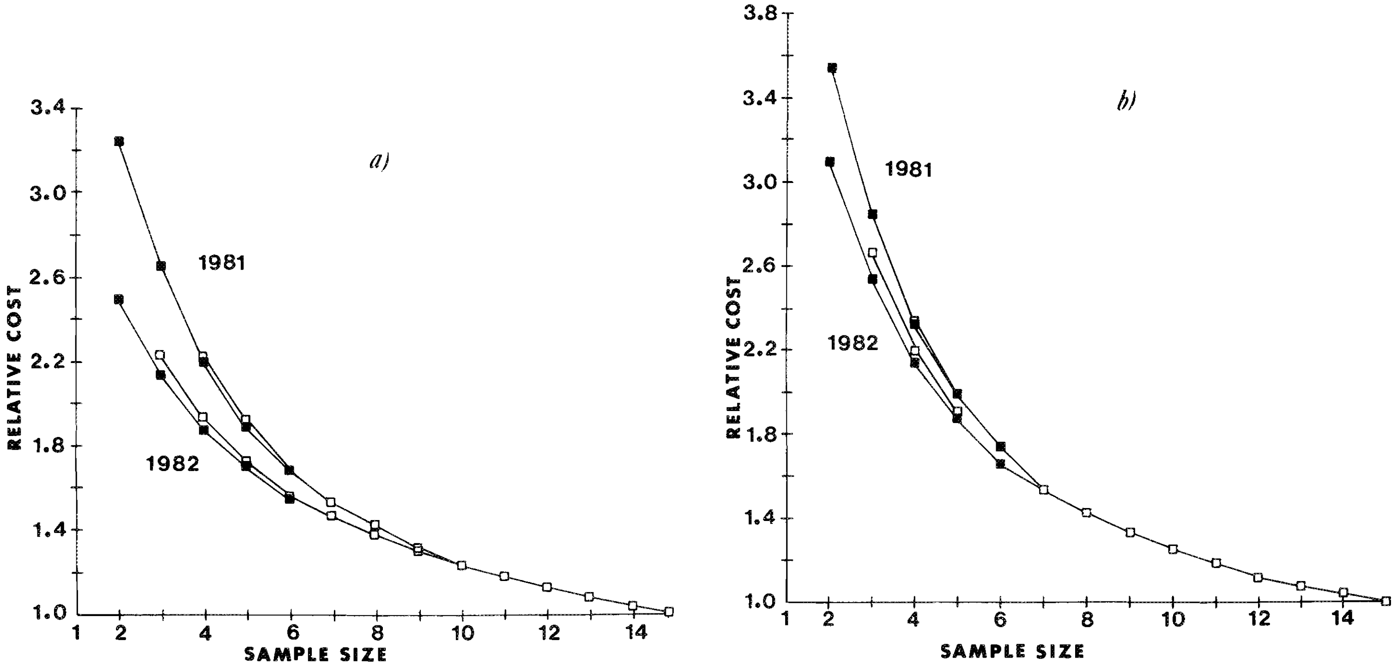

FIG. 2. Relative costs for the PPSWOR design, selection methods 1 (solid symbols) and 2 (open symbols), plotted against sample size for example 1 in 1981 and 1982. (a) Actual sizes of primary units; (b) effective sizes. Sample sizes for which one or more $\pi_{ij} = 0$ for the second selection method are listed in the caption to Fig. 1. Relative costs for PPSWR design in 1981 and 1982 were 3.44 and 3.03 for actual sizes and 4.24 and 5.20 for effective sizes. See text for explanation of actual and effective sizes of primary units.

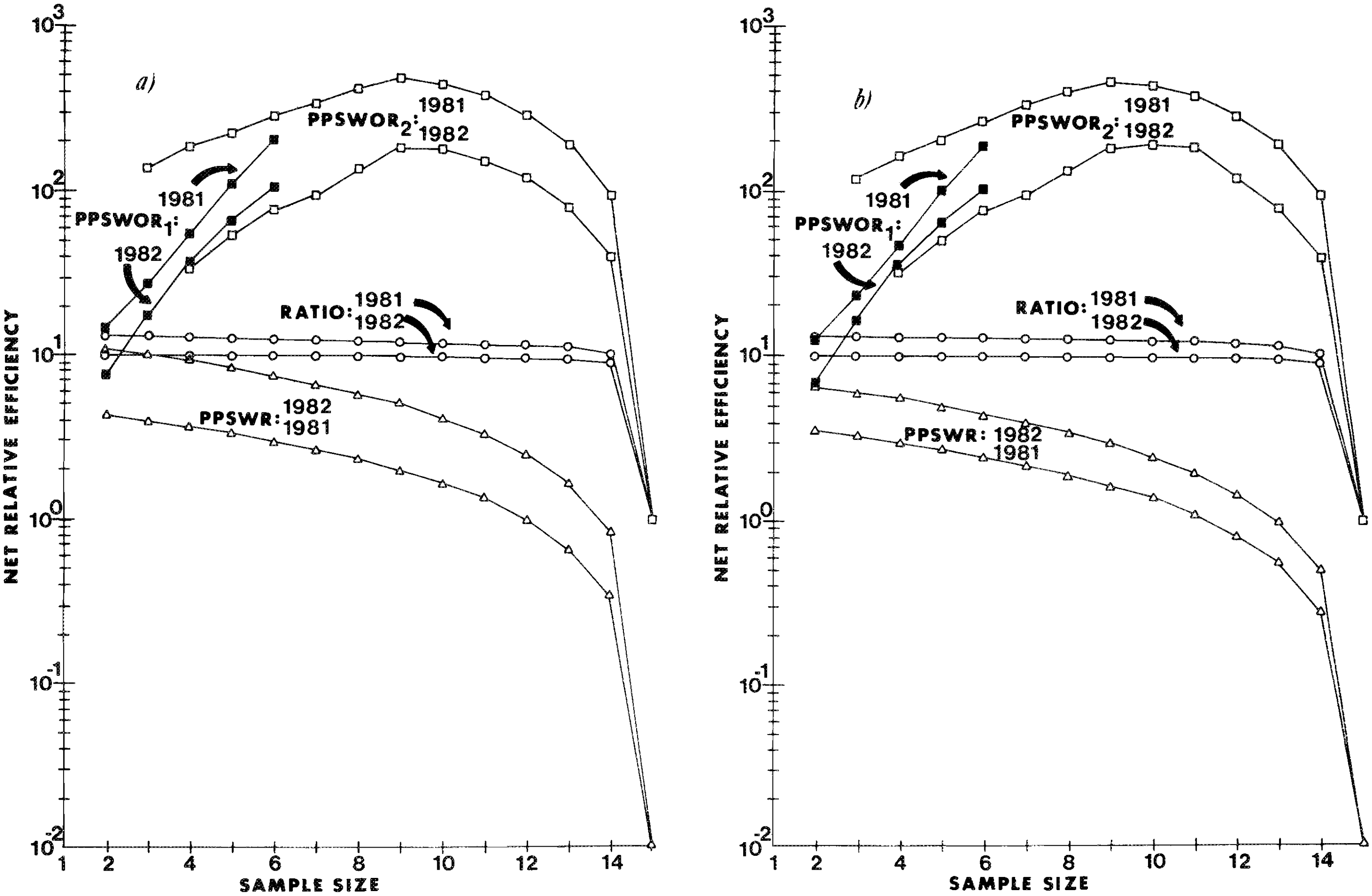

FIG. 3. Net relative efficiencies for RATIO, PPSWR, and PPSWOR designs plotted against sample size for example 1 in 1981 and 1982. (a) Actual sizes of primary units; (b) effective sizes. Subscripts on PPSWOR indicate selection methods 1 and 2. Sample sizes for which one or more $\pi_{ij} = 0$ for the second selection method are listed in the caption to Fig. 1.

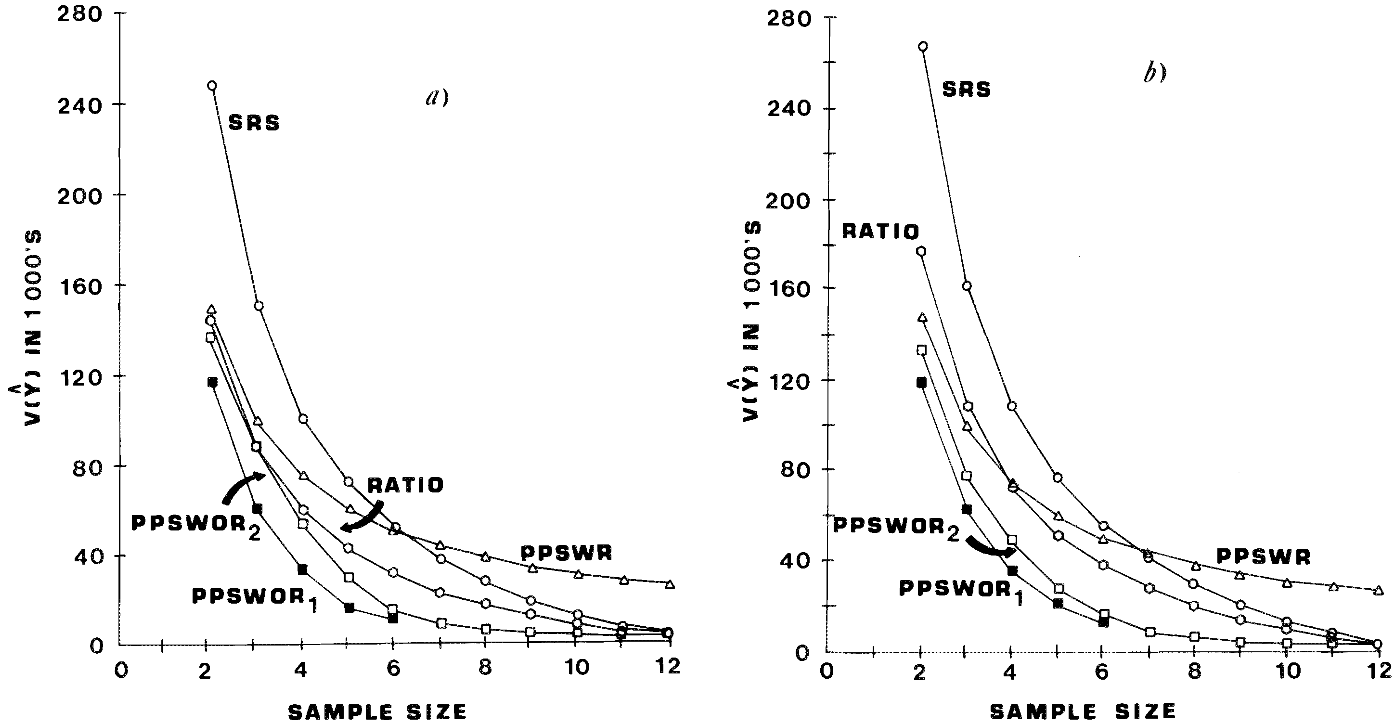

FIG. 4. Sampling variances, $V(\hat{Y})$, for SRS, RATIO, PPSWR, and PPSWOR designs plotted against sample size for example 2 in (a) 1981 and (b) 1982. Subscripts on PPSWOR indicate selection methods 1 and 2. For the second selection method: 1981: $\pi_{6,7} = 0$ ($n = 6$), $\pi_{10,11} = 0$ ($n = 10$), $\pi_{11,12} = 0$ ($n = 11$); 1982: $\pi_{7,8} = 0$ ($n = 7$), $\pi_{9,10} = 0$ ($n = 9$).

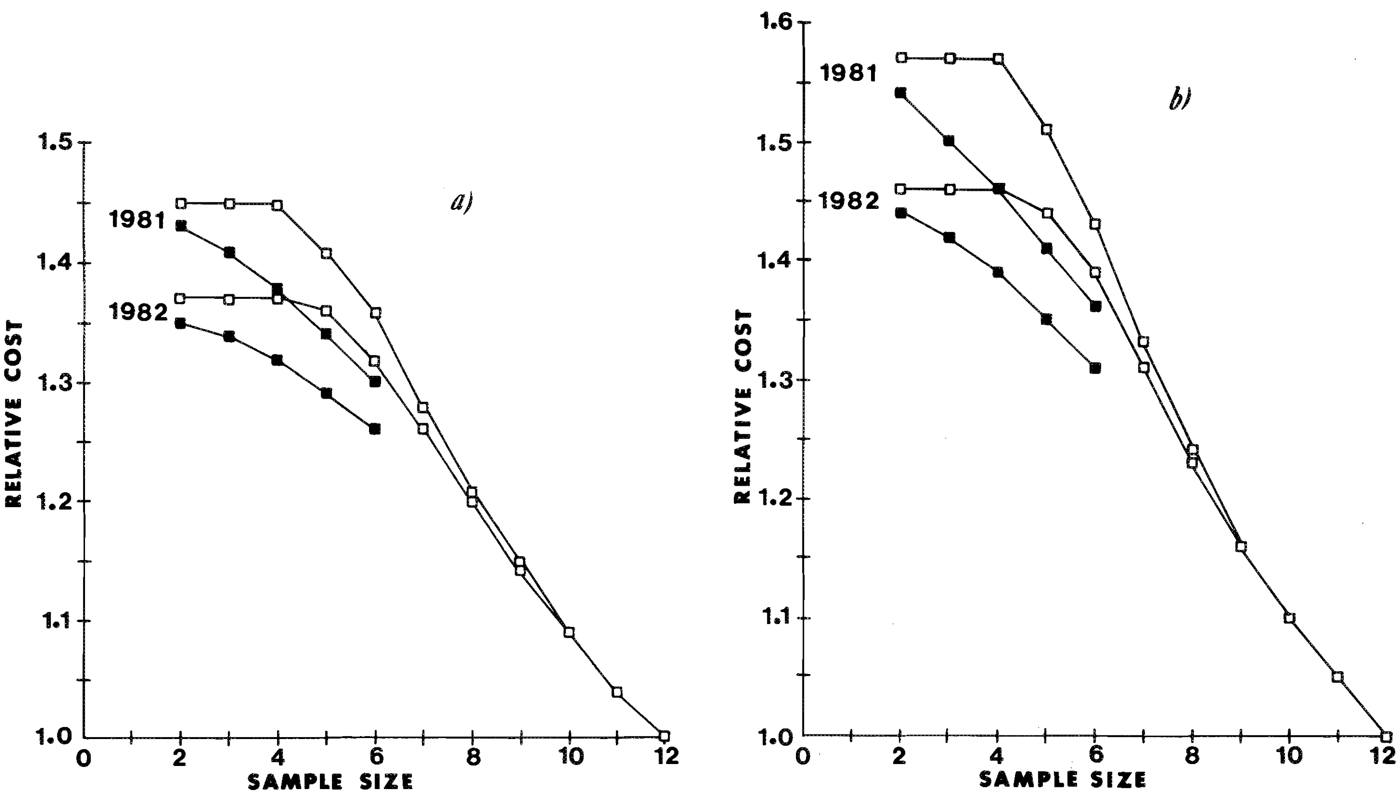

FIG. 5. Relative costs for the PPSWOR design, selection methods 1 (solid symbols) and 2 (open symbols), plotted against sample size for example 2 in 1981 and 1982. (a) Actual sizes of primary units; (b) effective sizes. Sample sizes for which one $\pi_{ij} = 0$ are listed in the caption to Fig. 4. Relative costs for PPSWR design in 1981 and 1982 were 1.45 and 1.37 for actual sizes and 1.74 and 1.59 for effective sizes.

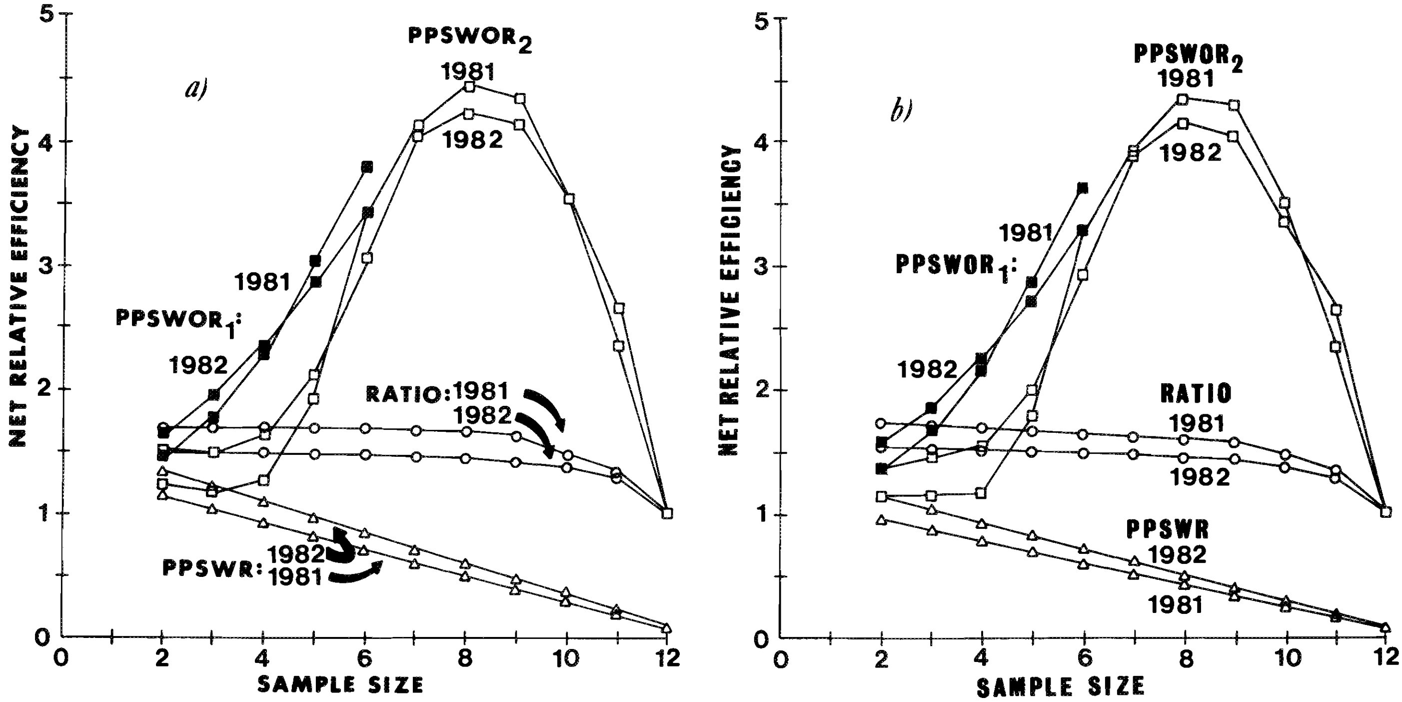

FIG. 6. Net relative efficiencies for RATIO, PPSWR, and PPSWOR designs plotted against sample size for example 2 in 1981 and 1982. (a) Actual sizes of primary units; (b) effective sizes. Subscripts on PPSWOR indicate selection methods 1 and 2. Sample sizes for which one $\pi_{ij} = 0$ for the second selection method are listed in the caption to Fig. 4.

Example 3

If the two smallest and two largest primary units are removed from data sets for each year presented in Table 1, then $N = 11$ and the correlations between Y_i and M_i are reduced to 0.310 and 0.398 for 1981 and 1982; these correlations would not be judged statistically significant ($p > 0.05$). Also, the range in primary unit sizes is reduced to just slightly more than one order of magnitude.

For these two sampling universes, $V_B(\hat{Y})$ exceeds $V_A(\hat{Y})$ for all sample sizes and all cases; for 1981 the contrast is strong. $V_C(\hat{Y})$ is slightly less than $V_B(\hat{Y})$ for $n \leq 3$, but for larger sample sizes $V_C(\hat{Y})$ exceeds $V_B(\hat{Y})$; $V_C(\hat{Y})$ exceeds $V_A(\hat{Y})$ for all sample sizes. For $n > 2$, $V_D(\hat{Y})$ is less than for any other design for all cases (Fig. 7).

Relative costs for the PPS without replacement design again decrease rapidly with increasing sample size, and for $n > 5$, NRE(D) always exceeds 1 (Fig. 8 and 9). Thus, for $n > 5$, the PPS without replacement design would again be the design of choice. For $n \leq 4$, the best choice would probably be the simplest: SRS (design A).

Example 4

For the artificial sampling universe presented in Table 2, $N = 50$, the correlation between Y_i and M_i is 0.760, and primary unit sizes range from 4.29 to 649 m^2. This universe thus has characteristics that are intermediate between examples 1 and 2 (in terms of correlation and range of primary unit sizes), but the sampling universe is much larger.

For this sampling universe, $V_A(\hat{Y})$ exceeds that of all other designs with one exception: for $n > 40$, $V_A(\hat{Y}) < V_C(\hat{Y})$. However, for small sample sizes, design C performs very well: for $n < 27$, $V_C(\hat{Y}) < V_B(\hat{Y})$; and for $n \leq 3$, $V_C(\hat{Y}) \approx V_D(\hat{Y})$. For sample sizes exceeding 3, $V_D(\hat{Y})$ is less than that for any other design (Fig. 10).

For $n < 8$, NRE(C) exceeds that for design B, and for $n \leq 3$, both PPS designs have comparable net relative efficiencies. NRE(D) increases dramatically with increasing sample size from $n = 2$ through at least $n = 28$ (Fig. 11). With the exception of very small sample sizes, then, design D (PPS without replacement) would again be the design of choice. For very small sample sizes ($n \leq 3$), design C might be the preferred choice due to the simplicity of the PPS with replacement calculations and the fact that equation 9 does not require estimates of second stage variances ($\hat{\sigma}_i^2$).

Figure 12 shows that the striking performance of the PPS without replacement design is achieved entirely through rapid reduction in first stage variance as sample size increases. For $n \leq 30$, second stage variance for the PPS design exceeds that for the SRS design (A), and for $n \geq 27$, second stage variance actually exceeds first stage variance for the PPS without replacement design. In contrast, first stage variance is at least an order of magnitude larger than second stage variance for $n < 40$ and is always large compared with second stage variance for the SRS design (A).

Discussion

The preceding examples illustrate that choice among alternative sampling designs for small streams depends sensitively on characteristics of sampling universes and on sampling fractions. Choice depends far less sensitively on electrofishing capture probability, q (even if q is more realistically allowed to depend on primary unit sizes; see Appendix A). For $q = 0.5$ and assumed independent of primary unit size, second stage variance was almost always small compared with first stage variance. The exception to this general result was for design D (PPS without replacement). For large sampling fractions (>50%), second stage variance exceeded first stage variance for this design (Fig. 12). However, one would rarely expect to

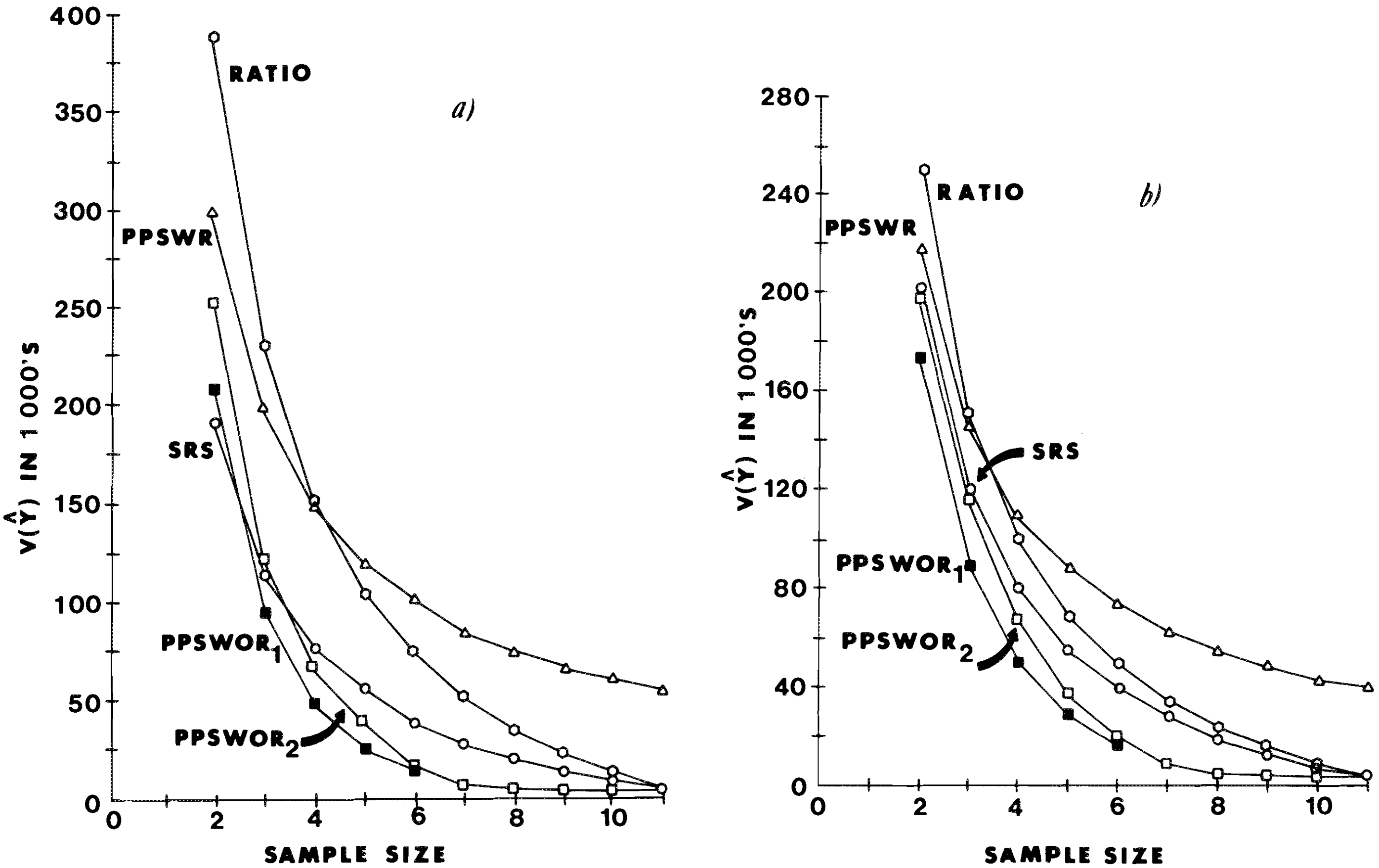

FIG. 7. Sampling variances, $V(\hat{Y})$, for SRS, RATIO, PPSWR, and PPSWOR designs plotted against sample size for example 3 in (a) 1981 and (b) 1982. Subscripts on PPSWOR indicate selection methods 1 and 2. For the second selection method: 1981: $\pi_{8,9} = 0$ ($n = 8$), $\pi_{10,11} = 0$ ($n = 10$); 1982: $\pi_{7,8} = 0$ ($n = 7$).

achieve such a large first stage sampling fraction in fisheries work.

Interestingly, far more attention has been paid to electrofishing capture probability (and hence, second stage variance) in small streams than to reduction in first stage variance through choice of two-stage sampling design (e.g. Bohlin 1982). Preoccupation with electrofishing capture probability has no doubt resulted from the convention of selecting primary units of equal sizes. When stream sections (primary units) are of equal sizes, there are no substantive alternatives to SRS or systematic selection of primary units and there are usually no obvious means whereby an auxiliary variable, such as primary unit size, can be used to increase the precision of resulting estimates. When primary units are allowed to vary in size, alternative two-stage sampling designs can substantially reduce first stage variance and substantially increase precision of estimation of the total number of fish in small streams.

Although results presented in this paper assumed that a removal method estimator, based on electrofishing, was used at the second stage of sampling, presented formulas and designs would remain valid and unaltered were some other population estimator, e.g. a mark–recapture estimator, used at the second stage of sampling. Ideally, estimators used at the second stage of sampling should be unbiased. Given unbiased second stage estimators, the formulas presented in this paper are unquestionably valid and, with the exception of design B (SRS/ratio estimation), formally unbiased. However, the author is not aware of any existing population estimators that are unbiased and also applicable to the small stream context. Of course, the use of an explosive (primacord) could result in direct enumeration of primary unit totals; second stage error is entirely eliminated by this method. When sample sizes are very small, such an approach may have merit. However, if one restricts oneself to the usual stream population estimators, mark–recapture and removal method, one must recognize that estimates of primary units are biased; therefore, presented formulas must be regarded as approximations.

When electrofishing is used at the second stage and capture probability exceeds 0.5, then the relative magnitude of first stage variance compared with second stage variance may present an even more dramatic contrast than that indicated by Fig. 12. Thus, when primary units are of unequal size, the relative performances of alternative two-stage sampling designs are diluted very little in the stream survey context from what their relative performances would have been in a one-stage setting. It is worth noting that this would not usually be the case in more typical commercial fishery applications of alternative two-stage sampling designs. When there are many discrete subunits at the second stage and second stage sampling fractions are small, the relative performances of alternative two-stage sampling designs may be very similar; second stage variance may be large compared with first stage variance (see Cochran 1977, p. 310).

In the stream survey context, choice among alternative sampling designs rests almost entirely on the characteristics of the sampling universe (the correlation between Y_i and M_i and

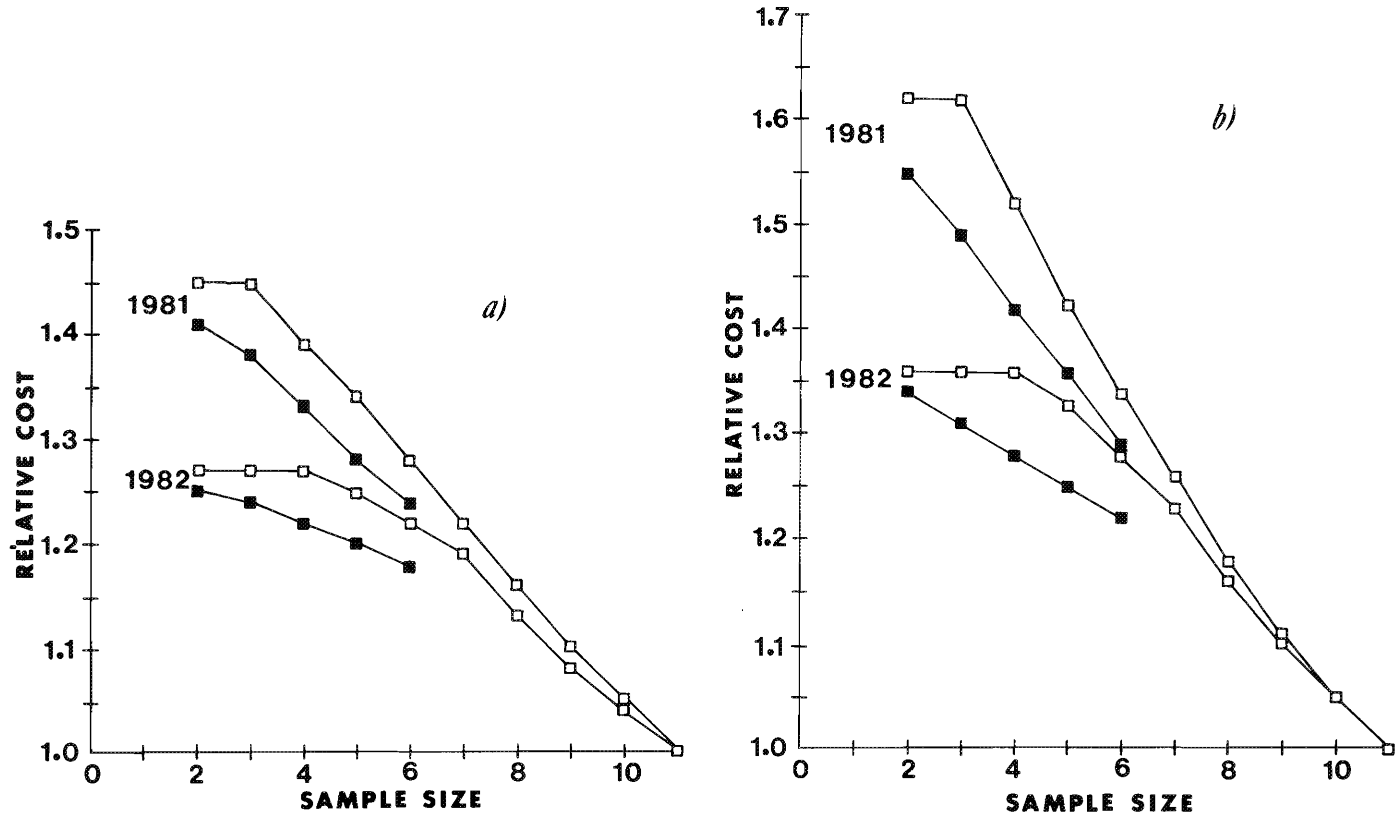

FIG. 8. Relative costs for the PPSWOR design, selection methods 1 (solid symbols) and 2 (open symbols), plotted against sample size for example 3 in 1981 and 1982. (a) Actual sizes of primary units; (b) effective sizes. Sample sizes for which one $\pi_{ij} = 0$ for the second selection method are listed in the caption to Fig. 7. Relative costs for PPSWR design in 1981 and 1982 were 1.45 and 1.27 for actual sizes and 1.88 and 1.49 for effective sizes.

the size of the sampling universe) and on sampling fraction. Examples provided calculations of net relative efficiencies of four alternative designs applied to three sets of small ($N \leq 15$) sampling universes, with a broad contrast of correlations between Y_i and M_i, and to one large ($N = 50$) sampling universe for which the correlation between Y_i and M_i was large (0.760) but realistic. Based on these examples, it is possible to make some fairly general, if crude, recommendations for the use of these four alternative designs when primary units are of unequal sizes.

For small sampling universes (e.g. $N < 20$) neither design B nor design C can be recommended regardless of the correlation between Y_i and M_i or of sampling fraction. PPS with replacement will perform poorly relative to PPS without replacement in small sampling universes, and the use of SRS/ratio estimation cannot be recommended because of possible serious underestimation in sample-based estimates of variance when sample sizes are small ($n < 12$; Cochran 1977, p. 162–164). One is thus left to choose between the SRS design (A) and the PPS without replacement design (D). When the correlation between Y_i and M_i exceeds about 0.5, PPS without replacement appears to be the design of choice, regardless of sample size. For small sample sizes (e.g. $n \leq 5$) the first PPS without replacement selection method can be used; for larger sample sizes the first selection method cannot be used, but the second selection method (Chao 1982) can be used (see Appendix B). For correlations between 0.3 and 0.5, choice of design appears to depend on sample size; for small sample sizes the SRS design is probably the design of choice, whereas for larger samples Chao's method of selecting PPS without replacement samples is recommended. When the correlation between Y_i and M_i is less than 0.3, the SRS design is recommended; the correlation is so low that one cannot take advantage of the potential benefits of PPS without replacement selection.

For large sampling universes ($N \approx 50$) choices are less clear-cut. When the correlation between Y_i and M_i exceeds 0.5, then design C (PPS with replacement) is recommended for small samples; the performances of PPS with and without replacement designs are so similar for small sample sizes that one cannot justify the additional computational complexities of the without replacement design. For larger sample sizes, Chao's method of selecting PPS without replacement samples is recommended and will outperform the with replacement design. For correlations less than 0.3, the SRS design is again the design of choice, and for correlations between 0.3 and 0.5 recommendations follow those for small sampling universes. Finally, for very large sampling universes (e.g. $N > 100$) with correlations between Y_i and M_i exceeding 0.5, PPS with replacement is recommended for small samples (e.g. $n \leq 12$); for larger samples, design B (SRS/ratio estimation) is recommended. Formulas for the SRS/ratio estimation design become approximately valid for large sample sizes selected from large sampling universes, and the PPS without replacement design is simply too unwieldly to use.

The author makes the above recommendations with some reluctance for two reasons. First, they are based on a comparison of the relative performances of alternative designs as applied to a limited and specific set of sampling universes.

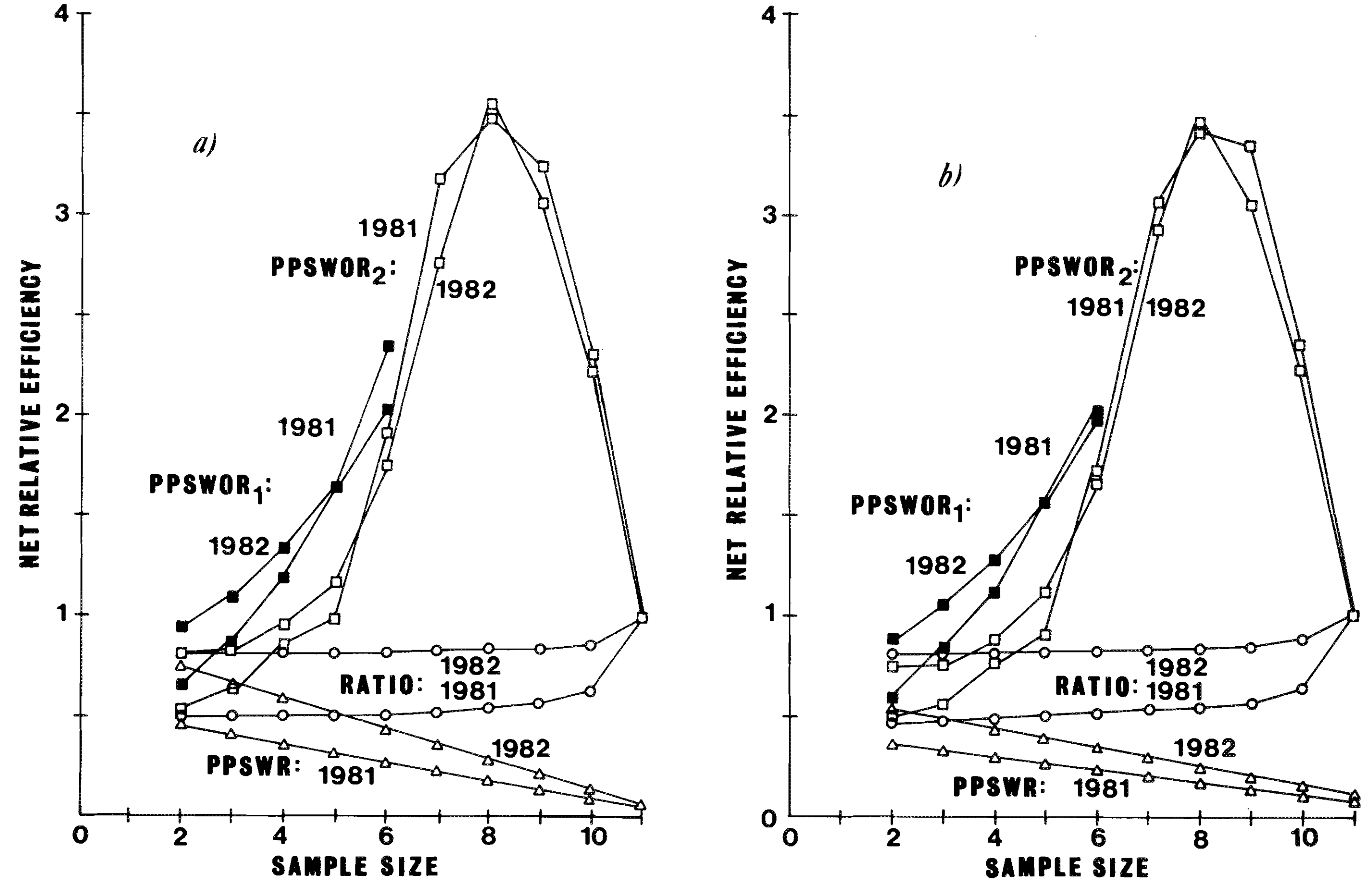

FIG. 9. Net relative efficiencies for RATIO, PPSWR, and PPSWOR designs plotted against sample size for example 3 in 1981 and 1982. (a) Actual sizes of primary units; (b) effective sizes. Subscripts on PPSWOR indicate selection methods 1 and 2. Sample sizes for which one $\pi_{ij} = 0$ for the second selection method are listed in the caption to Fig. 7.

Second, they strongly support the use of PPS without replacement designs in many contexts when the correlation between Y_i and M_i exceeds about 0.3. It is only recently that methods such as Chao's (1982) have been developed and have allowed extension of the PPS without replacement selection method to large sample sizes drawn from large sampling universes. Both PPS without replacement methods that were used in this paper require large numbers of computations (for $n > 2$). These computations required use of a computer; in the case of Chao's method one must use a computer to select a particular sample. It is clear, however, that for $N \approx 50$ and $n > 5$ Chao's method performs well and is feasible in a practical fisheries context. The author constructed his own computer programs (in APL) for implementing both methods of PPS without replacement selection that were used in this paper and is unaware of any packaged programs for these selection methods, although they may indeed exist.

Because of the above considerations, it is probably worthwhile to stress the generally solid performance of the SRS/ratio estimation design for moderate and large sample sizes ($n > 12$) drawn from large sampling universes for which the correlation between Y_i and M_i exceeds 0.5. There are two significant advantages of this design. First, computations are simple, can be done on a hand calculator if desired, and are easily extended to any sample size and any size sampling universe. Second, because primary units are selected by SRS (as in design A), one can always compare the performances of this design and the SRS design. That is, having selected n primary units by SRS, one may always incorporate a measure of the sizes of selected primary units (using equation 6) and determine whether there would be any gain in precision over results from design A (using equation 3). There would be no additional cost imposed for this comparison if the sizes of selected primary units were to be measured anyway for other purposes. However, absent knowledge of the total size of all primary units (M_0), one could not use SRS/ratio estimation to estimate the total over all primary units (see equation 4).

The fact that SRS/ratio estimation and PPS designs require knowledge of the total size of all primary units may seriously compromise the validity of net relative efficiency calculations that have been presented in this paper. As they have been presented, the PPS designs require a map of a stream that specifies the sizes of all primary units. The SRS/ratio estimation design requires knowledge of the location of all units, the total size of all units (M_0), and the sizes of particular units that are selected in a sample. In contrast, because the SRS design (A) does not incorporate a measure of the sizes of selected units, this design requires only a map of the location of all primary units. Thus, the relative costs for the SRS/ratio estimation and PPS designs might be far greater than those calculated in this paper. However, it is likely that very simple measures of the sizes of primary units may provide excellent measures for the assign-

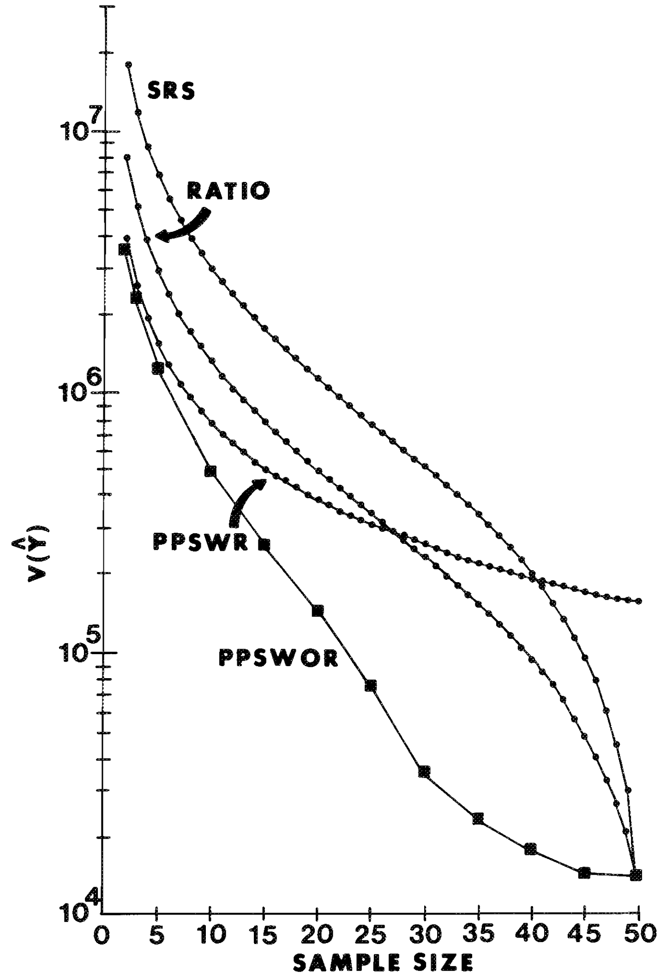

FIG. 10. Sampling variances, $V(\hat{Y})$, for SRS, RATIO, PPSWR, and PPSWOR designs plotted against sample size for example 4. For the PPSWOR designs, $V(\hat{Y})$ is computed following selection method 1 for $n = 2, 3$ and for other sample sizes following selection method 2. No $\pi_{ij} = 0$ for any sample size for the PPSWOR design.

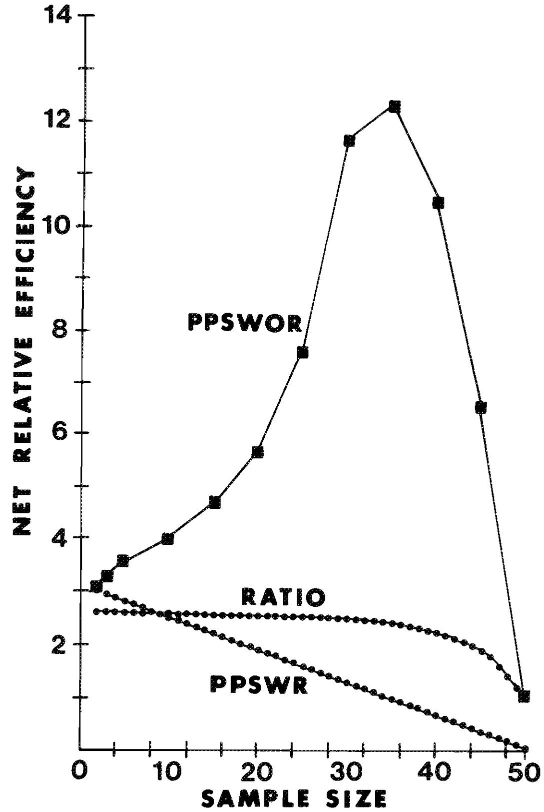

FIG. 11. Net relative efficiencies for RATIO, PPSWR, and PPSWOR designs plotted against sample size for example 4. For the PPSWOR design, $V(\hat{Y})$ is computed following selection method 1 for $n = 2, 3$ and for other sample sizes following selection method 2. No $\pi_{ij} = 0$ for any sample size for the PPSWOR design.

ment of PPS selection probabilities with little loss of efficiency. For example, if a particular reach of stream were of fairly uniform width, then the length of a pool could be a measure of size that would be highly correlated with pool area. This much simpler and less costly to obtain measure of primary unit sizes could then be used to assign PPS selection probabilities. The marginal cost of obtaining such simple measures of primary unit size would probably be small compared with the cost of locating all primary units; all primary units must be located for all designs. The author hopes to investigate this possibility in the future because of its obvious practical relevance for possible use of PPS designs.

For preliminary field research one usually has no existing data from which one could construct plausible sampling universes and conduct comparisons like those performed in this paper. Given such circumstances it is probably best to first sample primary units by SRS and then to compare the precisions of estimates based on designs A and B (assuming that sample sizes are large enough to warrant confidence in variance estimates from design B). If the sample correlation between Y_i and M_i is high, then future sampling might take advantage of this correlation by using one of the PPS designs following those general recommendations made previously. One should definitely not attempt immediate use of PPS designs when variation in primary unit sizes is small or when the correlation between Y_i and M_i is unknown or small. Preliminary field results can also give a great deal of insight into formation of strata within which alternative sampling designs could be used. Because habitat quality may vary markedly with location within a stream, the correlation between Y_i and M_i will probably be strongest within reaches of a stream that are most similar. Pools of the same size in a downstream, low-gradient region of a stream may hold very different numbers of fish than may pools of the same size in upstream, high-gradient regions. One might therefore choose to construct numerous small strata on the basis of general habitat quality and location within a stream, e.g. headwater pools for which stream gradient exceeds some specific amount, and to sample independently from each of these strata. It is very unlikely that there would be a strong correlation betwen Y_i and M_i over the entire length of any substantial stream. Only an intimate biological understanding of the stream to be sampled can lead to the most intelligent and effective construction of such strata; statistical advice alone is insufficient.

Finally, the author wishes to once more stress the importance of allowing primary unit sizes to vary according to natural habitat unit sizes. Regardless of the choice (or choices) that one makes among alternative two-stage sampling designs for small streams, one should minimize nonsampling errors that result from placement of block nets when primary units are essentially equivalent to natural habitat units. Also, one can learn far more

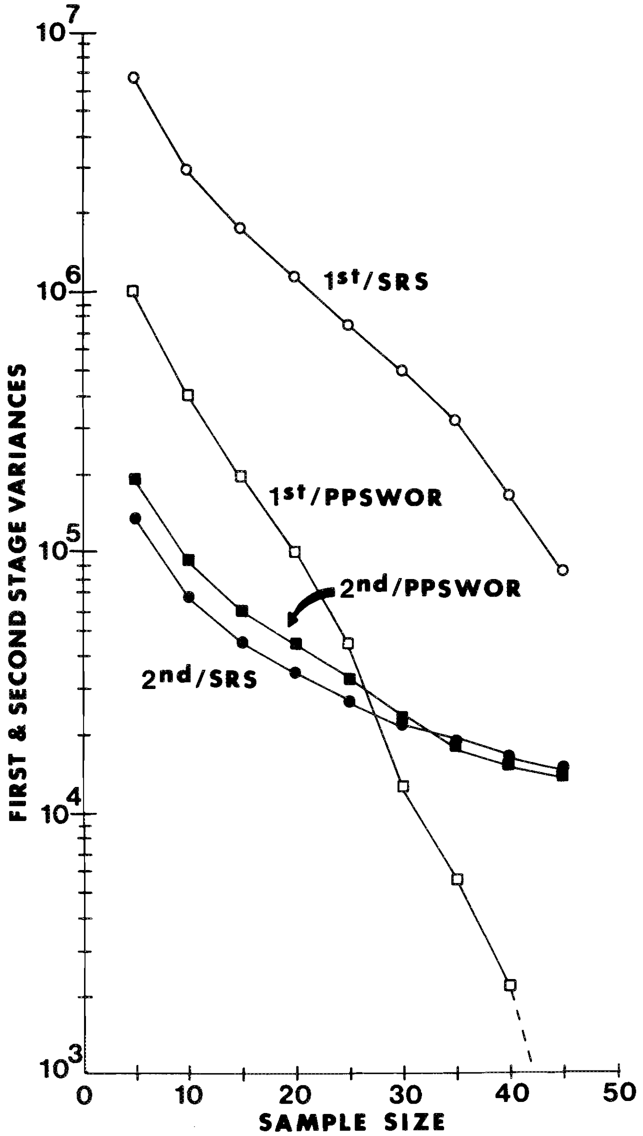

FIG. 12. First and second stage variances for SRS (1st/SRS and 2nd/SRS) and PPS without replacement (1st/PPSWOR and 2nd/PPSWOR) designs for example 4. All PPS calculations were based on selection method 2 and sample sizes are in increments of five.

about the relationships between fish abundance and distribution and habitat unit size and quality when one adopts the practice of allowing primary units to vary in size. After all, fish live in natural habitat units that we choose to call pools, runs, or riffles. We should sample them in their homes.

Acknowledgments

I extend sincere thanks to Fred Everest for his encouragement throughout the development of this paper, for permission to present the Knowles Creek data (without which a realistic appraisal of the relative performances of alternative sampling designs would have been impossible), and for his practical insights, which are reflected especially in Appendices A and C. I also extend thanks to A. R. Sen and D. S. Robson for their reviews of drafts of this paper; the discussion was substantially improved as a result. Support for publication was provided by the Humboldt State University Foundation and by the Pacific Northwest Forest and Range Experiment Station, U.S. Department of Agriculture.

References

ABRAMSON, N. J. 1968. A probability sea survey plan for estimating relative abundance of ocean shrimp. Calif. Fish Game 54(4): 257–269.

BOHLIN, T. 1981. Methods of estimating total stock, smolt output and survival of salmonids using electrofishing. Inst. Freshw. Res. Drottingholm Rep. 59: 5–14.

1982. The validity of the removal method for small populations — consequences for electrofishing practice. Inst. Freshw. Res. Drottingholm Rep. 60: 15–18.

CHAO, M. T. 1982. A general purpose unequal probability sampling plan. Biometrika 69(3): 653–656.

COCHRAN, W. G. 1977. Sampling techniques. Wiley, New York, NY. 428 p.

CUFF, W., AND N. COLEMAN. 1979. Optimal survey design: lessons from a stratified random sample of macrobenthos. J. Fish. Res. Board Can. 36: 351–361.

HANIF, M., AND K. R. W. BREWER. 1980. Sampling with unequal probabilities without replacement: a review. Int. Stat. Rev. 48: 317–335.

JESSEN, R. J. 1978. Statistical survey techniques. Wiley, New York, NY. 520 p.

KENDALL, M., AND A. STUART. 1983. The advanced theory of statistics. Vol. 3. Design and analysis, and time-series. 4th ed. MacMillan, New York, NY. 780 p.

LENARZ, W. H., AND P. B. ADAMS. 1980. Some statistical considerations of the design of trawl surveys for rockfish (Scorpaenidae). Fish. Bull. 78(3): 659–674.

PETERSON, N. P., AND C. J. CEDERHOLM. 1984. A comparison of the removal and mark–recapture methods of population estimation for juvenile coho salmon in a small stream. N. Am. J. Fish. Manage. 4: 99–102.

RAJ, D. 1968. Sampling theory. McGraw-Hill, New York, NY. 302 p.

SCHNUTE, J. 1983. A new approach to estimating populations by the removal method. Can. J. Fish. Aquat. Sci. 40: 2153–2169.

SCHWEIGERT, J. F., AND J. R. SIBERT. 1983. Optimizing survey design for determining age structure of fish stocks: an example from British Columbia Pacific herring (*Clupea harengus pallasi*). Can. J. Fish. Aquat. Sci. 40: 588–597.

SEBER, G. A. F. 1982. The estimation of animal abundance. Macmillan, New York, NY. 654 p.

SEN, A. R. 1954. On the estimate of the variance in sampling with varying probabilities. J. Indian Soc. Agric. Stat. 4: 119–127.

SOUTHWARD, G. M. 1976. Sampling landings of halibut for age composition. Int. Pac. Halibut Comm. Sci. Rep. 58: 31 p.

TOMLINSON, P. K. 1971. Some sampling problems in fishery work. Biometrics 27: 631–641.

Appendix A. Calculation of Second Stage Variances, σ_i^2, for the Removal Method Estimator

Calculation of the σ_i^2 assumed that electrofishing was used to estimate a primary unit total based on two "passes" of equal effort. Prior to electrofishing, block nets are set to delimit primary unit boundaries. On the first pass, Y_i fish are available for capture, and of these, C_1 fish are captured by electrofishing and removed from the primary unit. On the second pass, $Y_i - C_1$ fish are available for capture and C_2 fish are captured. Then, if (a) $C_1 > C_2$ and (b) capture probability, q (the probability that a fish present at the time of a given pass will be captured) is assumed to be constant, one may estimate the primary unit total as (Seber 1982, Sect. 7.2)

(13) $\hat{Y}_i = C_1^2/(C_1 - C_2)$

(14) $\sigma_i^2 = V(\hat{Y}_i) \approx Y_i(1 - q)^2(2 - q)q^{-3}$.

$\hat{Y}_i$ is a maximum likelihood estimator and is therefore asymptotically unbiased. However, $\text{BIAS}(\hat{Y}_i) \approx (1 - q)(2 - q)q^{-3}$, which may be significant for small Y_i and small q. $V(\hat{Y}_i)$ is based on the delta method (Seber 1982, Sect. 1.3.3.) and is a large sample result.

For small Y_i (e.g. <50) and small q (<0.5), bias of equation 13 may be substantial relative to Y_i. For example, for $Y_i = 50$ and $q = 0.25$, then BIAS($\hat{Y}_i$)/Y_i is nearly 50%. In contrast, for large $q(q \geq 0.8)$, BIAS($\hat{Y}_i$) < 0.5 and squared bias is almost always insignificant compared with $V(\hat{Y}_i)$. These kinds of considerations have been recently examined in detail by Bohlin (1982). Peterson and Cederholm (1984) showed that the assumption that q is constant may not be valid for small streams, and Schnute (1983) developed alternative removal method estimators for situations when this assumption is violated.

In the example computations presented in this paper, q was set equal to 0.5, a fairly low figure, so as not to minimize the possible importance of second stage variance in the stream survey context. The author recognizes that if q were as low as 0.5, then many estimates of Y_i and σ_i^2 would be tenuous for Y_i < 100. In addition, q was assumed to be independent of primary unit size. However, the usual finding in small streams is that, instead, q is inversely related to primary unit size. One generally has large q (e.g. ≥0.8) for small (usually shallow) pools and small q (e.g. ≤0.5) for large (usually deep) pools. For the Knowles Creek data presented in Table 1, $0.8 \leq \hat{q} \leq 1$ for all pools < 140 m^2; for larger pools, $\hat{q}$ decreased with increasing pool size.

If q varies inversely with primary unit size, then σ_i^2 may dramatically increase with primary unit size when Y_i and M_i are highly correlated (see equation 14). However, this kind of dependence of q on primary unit size did not result in adverse impacts on relative performances of the PPS designs that are presented in this paper, even though these designs assigned higher selection probabilities to primary units of greater sizes and greater σ_i^2. In fact, when q was allowed to vary with primary unit size and to mimic results from the Knowles Creek data, numerical computations (not presented in this paper) showed that the relative performances of the PPS designs could actually be improved.

Lack of adverse impact on the relative performances of the PPS designs appears to be related to weightings for the second stage variance terms for the four alternative two-stage designs. Weightings for the σ_i^2 were N/n for SRS and ratio estimation (equations 2 and 6), $1/np_i$ for PPS with replacement (equation 8), and $1/\pi_i$ for PPS without replacement (equation 11). For small n relative to N, N/n is large and the two SRS designs give this same large weighting to all primary units without regard to primary unit size or σ_i^2. In contrast, for the larger primary units, for which the σ_i^2 would also be far larger, the PPS designs will have large p_i or π_i and the weightings for these larger units are often less than for the SRS designs. For the large sampling universe for which $N = 50$ and with $n = 2$, the SRS weightings are all 25, while both PPS weightings are about 5 for the largest primary unit. Of course, for the smaller primary units, weightings are reversed and the PPS weightings will be far larger than the SRS weightings. However, for these smaller units the primary unit total (Y_i) will be small and capture probability will be large ($q \geq 0.8$) so that the σ_i^2 for these smaller units are effectively negligible when compared with σ_i^2 for the larger primary units. As n approaches N, all SRS and PPS without replacement weightings are approximately the same and approach 1. Weightings will not stabilize with increasing sample size for the PPS with replacement design (because the p_i are independent of sample size), but this design is not recommended for large sample sizes (see Discussion).

Appendix B. PPS Without Replacement Designs: Selection Methods and Variance Estimators

Two methods of selecting PPS without replacement samples were used for this paper. For the first method, the first unit is selected with probability $p_i = M_i/M_0$. The second and subsequent units are drawn with conditional probabilities proportional to the sizes of the remaining units. Computation of these conditional selection probabilities requires only manipulation of the original p_i. Let $z_i, z_j, z_k, \ldots, z_n$ denote the conditional probabilities of selecting primary unit i on the first draw, j on the second given i drawn on the first, k on the third given i and j drawn on the first two, etc. Then:

$$z_i = p_i;\ z_j = p_j/(1 - p_i);\ z_k = p_k/(1 - p_i - p_j);\ \ldots;$$
$$z_n = p_n/(1 - p_i - p_j - p_k - \ldots - p_{n-1});$$
$$i \neq j \neq k \neq \ldots \neq n.$$

For $n = 2$, this selection method gives the explicit results (Raj 1968, p. 51):

$$\pi_i = p_i + \sum_{j \neq i}^{N} p_j p_i/(1 - p_j)$$

and

$$\pi_{ij} = p_i p_j/(1 - p_i) + p_j p_i/(1 - p_j).$$

Thus, the probability that a sample of size 2 contains the units i and j (π_{ij}) is computed as the sum of the products of (a) the probability of selecting the ith (or jth) unit on the first draw and (b) the conditional probability of selecting the jth (or ith) unit on the second draw given that the ith (or jth) unit was selected on the first draw. This amounts to summing the probabilities of the two ordered selections in which both primary units i and j appear. For $n > 2$, the probability of selecting any ordered sample of size n that contains the specified units $i, j, k, \ldots, n$ can be computed as $z_i z_j z_k \ldots z_n$. Then, the probability that unit i appears in the sample, π_i, is computed as the sum of the probabilities of all those ordered selections that contain unit i. Similarly, π_{ij} may be computed as the sum of the probabilities of all those ordered selections in which both the units i and j appear. Because calculation of the π_i and π_{ij} by this selection method requires construction of all possible ordered samples, it has been termed an "all possible samples" selection method (Hanif and Brewer 1980). Note that there are $N!/(N - n)!$ possible ordered samples.

Selection of primary units with probabilities proportional to the sizes of remaining units (method one) has several disadvantages: (1) the sheer magnitude of computations necessary to calculate the π_i and π_{ij} strictly limits practical sample size, (2) there is distortion of inclusion probabilities in that the π_i are not proportional to M_i, and (3) the method does not guarantee that all $\pi_{ij} < \pi_i \pi_j$ (see below). However, for $n = 2$, calculations are rapid and explicit, and for small n and N this selection method may prove to be practical and may perform well (see Results).

The second selection method used in this paper was recently proposed by Chao (1982) and is of an entirely different character. Chao's method avoids the need to construct all possible samples and ensures that π_i are proportional to M_i for most units. One first orders the primary units and then selects the first n units. Successive units from $n + 1$ through N are then selected according to a probability scheme such that if the $n + i$th unit is selected, it replaces one of those units that was previously in the sample; sample size remains fixed. The π_i may be read directly from the last column of an easily constructed "normalized" probability matrix, and calculation of the π_{ij} follows explicit formulas (for any n) that readily lend themselves to computer algorithms. The advantages of Chao's method are (1) the number of computations necessary to calculate the π_i and π_{ij} is dramatically reduced as compared with all possible sample methods, thus allowing one to draw large samples from large sampling universes (e.g. $N \approx 50$), (2) for all $\pi_i < 1$, the π_i are strictly proportional to M_i, and (3) the method ensures that all $\pi_{ij} < \pi_i\pi_j$. However, Chao's method has the following disadvantages: (1) ordering of primary units will affect the π_{ij}, (2) the method does not guarantee that all $\pi_{ij} > 0$, and (3) certain π_i may equal 1: a unit may be selected with certainty. The author's experience has been that some π_{ij} may equal zero for very small ($n < 4$) or very large ($n \approx N$) sample sizes when primary units are ordered by decreasing size; for intermediate sample sizes usually all $\pi_{ij} > 0$. Thus, the two selection methods appear to complement one another. For very small samples drawn from moderate sized sampling universes (e.g. $N < 25$), it is practical to use the first selection method; for small sample sizes, Chao's method exhibits its disadvantages. For larger samples drawn from large sampling universes ($N \approx 50$), Chao's method will perform well whereas the first method will be impossible to apply due to the sheer magnitude of necessary computations (except for $n = 2$).

These formal details of the two PPS without replacement selection methods used in this paper are relevant for two reasons: (a) they serve to show that neither selection method can be used for all sample sizes and for all sampling universes; and (b) the formal properties described above may have serious consequences for the validity of using equations 11 and 12. First, equations 11 and 12 both require that all $\pi_{ij} > 0$; for some sample sizes, Chao's method would not meet this requirement. Second, although Sen (1954) proved that, for $n = 2$, the first selection method (or any other PPS without replacement selection method) will ensure that all $\pi_{ij} < \pi_i\pi_j$ so that equations 11 and 12 will always be positive, for $n > 2$ one has no such assurance. This means that if $\pi_{ij} > \pi_i\pi_j$ (for some i, j), certain pairs of primary units may actually make a negative contribution to variance (see equations 11 and 12). In fact, it is theoretically possible that equation 11 (and 12) may result in negative variance (estimates)!

For those example computations performed for this paper, it was always true that $\pi_{ij} < \pi_i\pi_j$ for the first selection method so that there were no possibilities of negative contributions to variance. For Chao's method of selection, primary units were ordered by decreasing sizes, and data points plotted on figures are such that no more than one $\pi_{ij} = 0$. Thus, on some figures there are no points plotted for Chao's selection method for very small n. Figure captions list those points for which a single $\pi_{ij} = 0$.

Appendix C. Calculation of Relative Costs

Let C denote total survey costs for sampling n primary units of unequal sizes that are selected by SRS, and assume that exactly half of these costs are attributable to time spent electrofishing the selected units. Then:

$$C = 0.5C + \gamma X_{SRS}$$

where γ = cost per unit of (actual or effective) primary unit size and X_{SRS} = expected total (actual or effective) size of n primary units selected by SRS. Normalize the total cost of sampling these n units by setting $C = 1$. Then for a specified n:

$$\gamma = 0.5/X_{SRS}.$$

Note that γ is a function of sample size, n. Relative costs for alternative sampling designs may then be calculated as $0.5 + \gamma X_{alt}$, where X_{alt} denotes the expected total (actual or effective) size of n primary units selected by some alternative method.

Expected total (actual or effective) sizes of n selected primary units for the four alternative designs are

$$nM_0/N(\text{actual}) \text{ and } n \sum^{N} M_i^{1.5}/N(\text{effective})$$

for SRS and SRS/ratio estimation,

$$n \sum^{N} M_i p_i(\text{actual}) \text{ and } n \sum^{N} M_i^{1.5} p_i(\text{effective})$$

for PPS with replacement, and

$$\sum^{N} M_i \pi_i(\text{actual}) \text{ and } \sum^{N} M_i^{1.5} \pi_i(\text{effective})$$

for PPS without replacement.

Recall that π_i is a function of sample size. Let X_{WR} and X_{WOR} denote the expected total (actual or effective) sizes for the PPS with and without replacement designs. Then the relative costs for these designs with respect to the SRS design(s) would be

$$\text{RC(PPSWR/SRS)} = 0.5 + \gamma X_{WR}$$

and

$$\text{RC(PPSWOR/SRS)} = 0.5 + \gamma X_{WOR}.$$

Because the π_i for the two PPS without replacement methods will differ, expected total (actual or effective) sizes of n selected units were computed for both of these selection methods.

Note that the expected total sizes of selected primary units are strictly proportional to sample size for both SRS designs and for PPS with replacement. This means that the relative cost of PPS with replacement is constant with respect to the SRS designs and independent of sample size. Relative cost for the SRS/ratio estimation design is always 1 because selection is by SRS. In contrast, the expected total size of n selected primary units will not be strictly proportional to sample size when units are selected by PPS without replacement. Because all π_i are bounded by 1, relative cost for the PPS without replacement design is a decreasing function of sample size.