

Planning programs to estimate salmon harvest with coded-wire tags

David R. Bernard, Robert P. Marshall, and John E. Clark

Abstract: Methods are presented for planning individual catch-sampling, tagging, and field-sampling programs to estimate salmon (*Oncorhynchus* spp.) harvest in recreational and commercial fisheries from several hatchery-produced and wild cohorts through recovery of coded-wire tags. We show how to determine sample sizes sufficiently large to detect harvest and link sample sizes to expenditures through linear and allometric cost functions to determine optimal tagging and catch-sampling rates. Sample sizes that will minimize bias and variance are charted for field-sampling programs designed to estimate the fraction of a cohort with tags. We describe sampling strategies that can be used to detect or to minimize bias in harvest estimates from tag loss, tag-induced mortality, tag-induced straying, and nonrandom sampling. Methods are demonstrated with data on cohorts of chinook (*O. tshawytscha*) and coho salmon (*O. kisutch*) from Alaska.

Résumé : Nous présentons des méthodes pour la planification de programmes d'échantillonnage des captures, de marquage et d'échantillonnage sur le terrain dans le but d'estimer la récolte par les pêches récréatives et commerciales de saumons (*Oncorhynchus* spp.) de diverses cohortes d'élevage et sauvages de par la récupération de fils codés. Nous montrons comment déterminer des effectifs d'échantillons suffisamment importants pour permettre de déceler la récolte et établissons une relation entre la taille de l'échantillon et les coûts par le moyen de fonctions de coûts linéaires et allométriques indiquant les taux optimaux de marquage et d'échantillonnage des captures. Les tailles des échantillons qui minimisent le biais et la variance sont présentées sous forme de tableaux pour les programmes d'échantillonnage sur le terrain conçus pour l'estimation de la fraction marquée d'une cohorte. Nous décrivons les stratégies d'échantillonnage qui peuvent être adoptées pour déceler ou minimiser le biais des estimations de récolte découlant de la perte d'étiquettes, de la mortalité due au marquage, de la fuite due au marquage et du caractère non aléatoire du marquage. Nous terminons par une démonstration des méthodes à l'aide de données sur des cohortes de saumons quinnat (*O. tshawytscha*) et coho (*O. kisutch*) de l'Alaska.

[Traduit par la Rédaction]

Introduction

Harvests from specific groups (cohorts) of Pacific salmon (*Oncorhynchus* spp.) exploited in fisheries along the north-west coast of North America are often estimated through the recovery of coded-wire tags (CWTs) etched with a binary code (Johnson 1990; Lapi et al. 1990). A sample from a cohort is doubly marked by inserting CWTs into the cartilaginous snouts of juvenile salmon and by excising their adipose fins. Adult salmon caught in commercial and recreational (sport) fisheries are subsequently sampled, and CWTs from recaptured fish in the sample are extracted and decoded. Harvest from a cohort is estimated by expanding the number

of CWTs recovered according to the fraction of the catch sampled, the fraction of the cohort carrying CWTs, the fraction of heads from recaptured fish that reach a laboratory, and the fraction of dissected heads from which a CWT is decoded. Clark and Bernard (1987) and Schnute (1992) have investigated statistical methods of estimating these harvests when stratified catch-sampling programs are the only source of measurement error. Bernard and Clark (1996) described methods to estimate harvest that incorporate additional measurement error from sampling to estimate catch of all salmon and from field-sampling programs to estimate the fraction of a cohort carrying CWTs.

How many juveniles should be tagged, how many adults should be inspected, and how much will these activities cost are common questions when planning programs to estimate salmon harvest. Enough salmon in a cohort must be tagged and enough catch inspected so that some CWTs are recovered when some of the cohort has been harvested. Beyond detecting harvest, maximizing precision of its estimate is a matter of minimizing variances by optimally allocating monies in sampling and tagging programs. Although simultaneous, optimal allocation of available funds among all sampling and tagging programs is conceivable (see Clark and Bernard 1987), it is often not possible. Individual programs are funded by different agencies and states and sometimes different nations in different fiscal years (see Johnson 1990). Planning is distributed among many program supervi-

Received October 16, 1996. Accepted March 31, 1998.
J13705

D.R. Bernard.¹ Division of Sport Fish, Alaska Department of Fish and Game, 333 Raspberry Road, Anchorage, AK 99518-1599, U.S.A.

R.P. Marshall. Division of Sport Fish, Alaska Department of Fish and Game, 802 3rd Street, Douglas, AK 99824-0020, U.S.A.

J.E. Clark. Division of Commercial Fisheries Management and Development, Alaska Department of Fish and Game, 802 3rd Street, Douglas, AK 99824-0020, U.S.A.

¹Author to whom all correspondence should be addressed.
e-mail: davidbe@fishgame.state.ak.us

Table 1. Definition of notation.

a_i	Cost of sampling all catch in stratum i
b_i	Exponent in allometric cost function for sampling in stratum i
β_j	Probability (risk) of recovering no CWTs from some harvest of cohort j
c_j	Per-unit cost of (1) tagging a single member of cohort j in a hatchery or (2) a unit of sampling effort to capture and tag members of a wild cohort
C	Available monies to fund catch-sampling or tagging programs
$\Delta(y)$	Function of variable y expressing anticipated changes in π from past cohorts
γ_{jk}	Ratio of θ_k to θ_j that produces minimal variance for estimated harvest
i, t	Subscript i represents a single stratum in a catch sampling program; t represents another stratum
j, k	Subscript j represent a single cohort of salmon; k represents another cohort
j'	Subscript j' represents a past cohort with background similar to cohort j , the cohort to be tagged
λ_i	Decoding rate of CWTs for marked fish in the sample from stratum i
m_{ij}	Number of decoded CWTs from cohort j in the sample from stratum i
M_j	Number of juvenile salmon in cohort j at their release
n_i	Size of sample taken from the catch in stratum i
N_i	Size of the catch in stratum i
ω_{it}	Ratio of ϕ_t to ϕ_i that produces minimal variance for estimated harvest
p_{ij}	Proportion of fish caught in stratum i having a CWT from cohort j
π_{ij}	Probability that a CWT will be recovered in stratum i from a tagged member of cohort j
ϕ_i	Fraction of catch inspected (sampled) in stratum i (the catch-sampling rate)
q_j	Catchability coefficient for wild juveniles in cohort j
r_{ij}	Number of fish harvested from cohort j in stratum i
R_j	Subset of M_j juvenile salmon released with CWTs
s_j	Size of sample used to estimate θ_j
S_j	Survival rate of juveniles in cohort j
T	Harvest from several cohorts over several strata
θ_j	Fraction of adult salmon in cohort j carrying a CWT (the tagging rate)
U_{ij}	Exploitation rate for cohort j in stratum i

sors, each responsible for tagging a few cohorts or for sampling a few fisheries. By default, questions of how many fish should be sampled or tagged are asked from the supervisor's point of view. To be useful, methods used to answer these questions should have the same perspective.

Regardless of sample size, tagged and inspected adult salmon must be representative of their cohort and the catch if estimates of harvest are to be accurate. Use of statistical models to estimate harvest implies that tagged and inspected salmon are representative because they have been randomly chosen. Unfortunately, salmon in CWT programs are rarely chosen randomly, but are tagged or sampled according to their accessibility. Random selection is impossible or is just

too costly. Yet samples can be representative without being randomly chosen, as is also true for more conventional mark-recapture experiments (see Seber 1982, section 11.1.4). The likelihood that tagged juveniles will represent the cohort as adults or that sampled salmon will be representative of caught salmon can be improved through planning.

In this paper, we present methods developed during a decade of planning individual tagging, catch-sampling, and field-sampling programs to estimate salmon harvest with CWTs (see Bernard et al. 1993). We begin by laying out the philosophy and procedures common to planning all tagging and sampling programs (Section 1). Methods to determine how many hatchery-produced or wild salmon to tag to detect most harvest or to cost-effectively estimate harvest with available monies are given in Section 2. Section 3 covers cost-effective sample sizes when sampling catches from commercial and sport fisheries, while sample size in a field-sampling program to estimate the marked fraction of a wild cohort is the topic in Section 4. Section 5 covers sampling protocols to minimize bias in estimates or to recognize failure to do so. Some nonmathematical aspects to planning are discussed in the final section.

1. Common elements

We presume that someone planning a tagging or a sampling program to estimate salmon harvest with CWTs wants to spend available monies in such a way as to detect all significant harvest and to maximize precision of its estimate. Detection and precision are functions of numbers tagged and sample sizes, and sample sizes and numbers tagged are functions of monies spent. Generally, the more money spent, the more fish that can be tagged or inspected, and the more harvest that can be detected and estimated with greater precision, although the marginal improvement declines with increases in spending. Yet regardless of the funds available, there is often an optimal way of spending money within a sampling or tagging program.

Risk of not detecting harvest, and precision of estimated harvest detected, can be determined by first linking cost to numbers tagged or sample sizes with cost functions and then by connecting sample sizes or numbers tagged to risk with risk functions or to precision with variance equations. For example, given a budget of C monies and cost c to tag a young salmon, $R = C/c$ fish can be tagged from a cohort of M salmon (see Table 1 for definitions of notation). If π is the proportion of R tagged salmon that ultimately yield a CWT to a catch-sampling program, $R\pi (= m)$ is the expected number of tags recovered from that cohort. Obviously, the larger the number marked R , the greater the probability of detecting harvest by recovering at least one tag in a catch-sampling program. Also, the larger R is relative to M (the greater the tagging rate θ) in a cohort, the smaller the variance will be for estimated harvest (Bernard and Clark 1996). Once costs are transformed into samples sizes or numbers tagged, monies needed to achieve an acceptable risk of not detecting harvest can be determined by inverting the risk function. The same procedure holds for determining precision of estimated harvest from money spent, except when a planner is concerned with simultaneously maximizing precision (minimizing variance) of estimated harvest from two or

more cohorts or from catch in two or more fisheries. In this situation, the cost function also acts as a constraint in a simultaneous solution of relative tagging or sampling rates that will minimize overall variance. With this solution and the cost function, a planner can anticipate the best precision for monies spent or determine the minimum amount of money to achieve estimates with acceptable precision.

While cost functions differ across tagging, catch-sampling, and field-sampling programs, the same risk function and variance equation can be used to plan all programs. When planning, all variables in the risk function or variance equation are known, either because they are “controlled” (sample sizes or numbers tagged) or because they are “fixed.” Fixed variables represent statistics that are outside the influence of the planner. For instance, variance of estimated catch from a creel survey is usually outside the control of a tagging program, yet variance of estimated catch is part of the overall variance of estimated harvest (Bernard and Clark 1996) and should be considered when determining how many young salmon in a cohort should be tagged. The kind of program being planned determines which variables will be controlled and which will be fixed. Historical information and insight are used to establish values for fixed variables, while values of controlled variables depend ultimately on available monies.

The risk function for detecting harvest contains two variables, R and π , the first a function of tagging young salmon and the second a function of sampling adults in the catch. For a cohort j , the number of tags recovered from a single stratum i follows the binomial probability distribution:

$$\text{Prob}(m_{ij} | R_j, \pi_{ij}) = \binom{R_j}{m_{ij}} \pi_{ij}^{m_{ij}} (1 - \pi_{ij})^{R_j - m_{ij}}$$

A stratum is usually a fishing period in a fishing district for a specific fishery, such as the first 2 weeks in June of the sport fishery centered around Ketchikan, Alaska. No special distinction is needed for strata from different fisheries because catch sampling is independent among strata regardless of fishery. From this distribution, $\text{Prob}(m_{ij} = 0) = (1 - \pi_{ij})^{R_j}$ and $\text{Prob}(m_{ij} > 0) = 1 - (1 - \pi_{ij})^{R_j}$. The product of the $\text{Prob}(m_{ij} > 0)$ across all strata is an approximate expression of the power $(1 - \beta_j)$ to simultaneously detect all harvest with β_j the risk of missing some harvest from cohort j , making the risk function

$$(1.1) \quad 1 - \beta_j \approx \prod_i [1 - (1 - \pi_{ij})^{R_j}]$$

Equation 1.1 would be exact if $\{m_{ij}\}$ are independent across strata, which, of course, they are not. However, because

$$R_j \gg \sum_i m_{ij} \text{ and } 1 \gg \sum_i \pi_{ij} \text{ in CWT programs, dependence among } \{m_{ij}\} \text{ is small and can be ignored with negligible consequence.}$$

Someone planning to tag a cohort would determine risk by controlling R_j and by fixing $\{\pi_{ij}\}$. In contrast, someone in charge of catch sampling would fix R_j and would control $\{\pi_{ij}\}$ through catch-sampling rates.

Each π_{ij} is a product of several independent, process-oriented probabilities relating to survival rates, migratory

patterns, and exploitation rates and to probabilities of a salmon being sampled and its tag decoded. If considerable recovery and ancillary data are available for past cohorts, $\{\pi_{ij}\}$ can be obtained by fitting log-linear models to data to estimate survival rates, distribution, and exploitation rates for past cohorts, as is done for the annual cohort analysis for chinook salmon (*O. tshawytscha*) caught in fisheries along the northeast edge of the Pacific Ocean (see Chinook Technical Committee 1988).

A simpler procedure is to adjust MLEs of $\{\pi_{ij}\}$ from past cohorts for anticipated circumstances. Numbers of fish in a population with different capture histories in a mark-recapture experiment follow a multinomial distribution (see Seber 1982, pp. 431–432). Consider a past cohort, cohort j' , with a pedigree similar to the current cohort j . In a mark-recapture experiment with CWTs, $m_{ij'}$ tags from cohort j' recovered in stratum i follow a binomial distribution conditioned on $\pi_{ij'}$ and $R_{j'}$, the number of fish tagged, and the MLE $\hat{\pi}_{ij'} = m_{ij'}/R_{j'}$. Now, if the current cohort (cohort j) is expected to encounter the same circumstances as its predecessor, cohort j' , then $\hat{\pi}_{ij'}$ can be used directly as π_{ij} . If change is anticipated in some process variable y associated with π , $\Delta(y) = y/\hat{y}$ is an adjustment for that change such that

$$(1.2) \quad \pi_{ij} = \hat{\pi}_{ij'} \Delta(S_j) \Delta(U_{ij}) \Delta(\phi_i) \Delta(\lambda_i)$$

where S_j is the juvenile survival rate for cohort j , U_{ij} is its exploitation rate in stratum i , ϕ_i is the catch-sampling rate, and λ_i is the decoding rate for tags from recovered salmon. For example, if the only anticipated change is a 10% increase in ϕ_i from added funding of catch-sampling programs, $\pi_{ij} = \hat{\pi}_{ij'} (1.10)$. Decoding rates $\{\lambda_i\}$ are products of two proportions: the proportion of sampled salmon with missing adipose fins whose heads reach a dissecting laboratory and the proportion of CWTs detected in heads that are successfully decoded. Historical precedent is often used to anticipate $\{\lambda_i\}$, so $\{\Delta(\lambda_i)\}$ usually equals $\{1\}$. Extreme annual variability in survival rates of smolts makes information from past cohorts of questionable relevance for a cohort yet to be tagged; however, arbitrarily adjusting $\{\pi_{ij}\}$ for the lowest observed survival rate over several past cohorts (the worst-case scenario) should produce conservatively high sampling or tagging rates. The $\{\Delta(U_{ij})\}$ are usually set to 1 when planning tagging programs because future changes in fisheries management are rarely known when fish are tagged. Such changes are sometimes known when planning catch-sampling programs. If fishing is intensely competitive among strata, as in many gauntlet fisheries, adjustments in $\hat{\pi}_{ij'}$ might require incorporating changes in exploitation rates in preceding strata:

$$(1.3) \quad \pi_{ij} = \hat{\pi}_{ij'} \Delta(S_j) \Delta(\phi_i) \Delta(\lambda_i) \Delta(U_{ij}) \frac{\prod_{t=1}^{i-1} [1 - \Delta(U_{tj'}) \hat{U}_{tj'}]}{\prod_{t=1}^{i-1} [1 - \hat{U}_{tj}]}$$

where fishing in stratum t precedes fishing in stratum i . When exploitation rates are low, eq. 1.2 remains a reasonable approximation for competitive fishing. Influence of annual mortality rates, maturity schedules, and migratory patterns on $\{\pi_{ij}\}$ are usually considered constant across

years when planning and are implicitly imbedded in past estimates $\{\hat{\pi}_{ij}\}$.

The variance equation for estimated harvest contains many variables, including π . An estimated total harvest \hat{T} and its variance for one or more cohorts of salmon exploited in one or more fisheries are sums of statistics across cohorts and fisheries:

$$(1.4) \quad \hat{T} = \sum_i \sum_j \hat{r}_{ij}$$

$$(1.5) \quad \text{Var}(\hat{T}) = \sum_i \sum_j \text{Var}(\hat{r}_{ij}) + 2 \sum_i \sum_j \sum_{k>j} \text{Cov}(\hat{r}_{ij}, \hat{r}_{ik})$$

where r_{ij} (or r_{ik}) is harvest in stratum i from cohort j (or k). Harvest from cohort j in stratum i can be expressed $r_{ij} = N_i p_{ij} \theta_j^{-1}$ (from Bernard and Clark 1996, eq. 7), and its variance

$$(1.6) \quad \text{Var}(\hat{r}_{ij}) = r_{ij}^2 [G(\hat{p}_{ij}) + G(\hat{N}_i) + G(\hat{\theta}_j^{-1}) + G(\hat{p}_{ij})G(\hat{N}_i) + G(\hat{p}_{ij})G(\hat{\theta}_j^{-1}) + G(\hat{N}_i)G(\hat{\theta}_j^{-1}) + G(\hat{p}_{ij})G(\hat{N}_i)G(\hat{\theta}_j^{-1})]$$

where p_{ij} is the fraction of catch in stratum i with CWTs representing cohort j , N_i is the catch in stratum i , and $G(\cdot)$ denotes the squared coefficient of variation for the specified variable. Equation 1.6 can be derived from methods in Bernard and Clark (1996) but differs somewhat from their development (their eq. 13) because the latter is for an estimate of variance. Because variables are treated as known when planning, variance of estimated harvest is also known, and its derivation takes the form expressed in Goodman (1960, eq. 15). For the same reason, covariance between estimated harvests from any two cohorts would also be considered known, and as transformed from the expression in Bernard and Clark (1996) would be

$$(1.7) \quad \text{Cov}(\hat{r}_{ij}, \hat{r}_{ik}) = r_{ij} r_{ik} G(\hat{N}_i)$$

The value of $G(\hat{N}_i)$ is usually considered zero for commercial fisheries or a function of sample sizes and design of creels surveys of sport fisheries (Pollock et al. 1994). The variable $G(\hat{\theta}_j^{-1})$ is usually considered zero for most hatchery-reared cohorts or a function of sample size s in a field-sampling program when wild cohorts are involved. The statistic $G(\hat{p}_{ij})$ is a mix of fixed and controlled variables and is the link to cost functions. From Bernard and Clark (1996, eq. 12), $\text{Var}(\hat{p}_{ij}) = p_{ij} (1 - \lambda_i \phi_i \theta_j) (\lambda_i n_i)^{-1}$ where n_i is the number of adults inspected for marks in stratum i in a catch-sampling program and ϕ_i is the sampling rate in stratum i ($= n_i / N_i$). As a function of a sampling rate ϕ_i , a tagging rate θ_j , and a harvest r_{ij} ($= N_i p_{ij} \theta_j^{-1}$):

$$(1.8) \quad G(\hat{p}_{ij}) = \frac{\text{Var}(\hat{p}_{ij})}{p_{ij}^2} = \frac{1 - \lambda_i \phi_i \theta_j}{\lambda_i \phi_i \theta_j r_{ij}}$$

Assuming that tagged salmon are representative of untagged salmon in a cohort, harvest r_{ij} can also be expressed as function of the initial size of a cohort M_j and the fraction π_{ij} of

tagged members that survived to yield tags discounted for decoding and catch-sampling rates:

$$(1.9) \quad r_{ij} = \frac{M_j \pi_{ij}}{\lambda_i \phi_i}$$

In planning CWT programs, $\{r_{ij}\}$ and $\{\lambda_i\}$ are considered fixed variables, while $\{\phi_i\}$ and $\{\theta_j\}$ can be either fixed or controlled.

2. Tagging

We begin development and demonstration of planning procedures specific to tagging programs with the situation of tagging a single cohort. For a hatchery-produced cohort, $C = c_j R_j = c_j M_j \theta_j$ where C is the amount of money available for tagging and c_j is the per-unit cost of tagging one member of cohort j . Substituting this cost function into eq. 1.8 gives

$$(2.1) \quad G(\hat{p}_{ij}) = \frac{c_j M_j - \lambda_i \phi_i C}{\lambda_i \phi_i r_{ij} C}$$

Variance $\text{Var}(\hat{T})$ is now a function of available monies as per eqs. 1.5–1.7. Because wild juveniles are captured, not corralled, the usual cost function for tagging salmon from a wild cohort is $C = c_j q_j^{-1} \theta_j$ where q_j is the standard catchability coefficient and c_j is the cost of a single unit of sampling effort. So long as θ_j is relatively small, which unfortunately is usually so for wild cohorts, this linear equation is a reasonable cost function. Rearranging the function to $\theta_j = q_j c_j^{-1} C$ and substituting the result into eq. 1.8:

$$(2.2) \quad G(\hat{p}_{ij}) = \frac{c_j - \lambda_i \phi_i q_j C}{\lambda_i \phi_i r_{ij} q_j C}$$

For wild cohorts, r_{ij} must be calculated by estimating the size of a cohort when tagged, i.e., \hat{M}_j instead of M_j in eq. 1.9. Subsequent calculations proceed as with a hatchery-produced cohort.

If several hatchery-produced cohorts are to be tagged with the same monies and estimates of their harvests summed, the linear cost function becomes $C = \sum_j c_j M_j \theta_j = \sum_j c_j R_j$, and

there is an optimal set of numbers tagged $\{R_{j(\text{opt})}\}$ and tagging rates $\{\theta_{j(\text{opt})}\}$ such that $\text{Var}(\hat{T}|C)$ is minimized. The relationship between tagging rates θ_j and θ_k for any two cohorts j and k that minimizes variance is

$$(2.3) \quad \frac{\theta_k}{\theta_j} = \frac{\sqrt{c_j M_j \sum_i \frac{r_{ik}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]}}{\sqrt{c_k M_k \sum_i \frac{r_{ij}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]}} = \gamma_{jk}$$

Derivation of eq. 2.3 is described in the Appendix. Inspection of eq. 2.3 shows that variance is reduced when smaller cohorts that are cheaper to tag and contribute relatively more harvest are tagged at a higher rate. The optimal tagging rate can be obtained by substituting the relationship $\theta_k = \gamma_{jk} \theta_j$ into the cost function and solving:

Table 2. Statistics for two cohorts of chinook salmon caught as 5- and 6-year-olds in troll and sport fisheries near Ketchikan, Alaska.

Stratum	Year	\hat{N}_i	$CV(\hat{N}_i)$	n_i	$\hat{\phi}_i$	λ_i	Whitman Lake		Neets Bay	
							m_{ij}	$\hat{\pi}_{ij} (\times 10^{-4})$	m_{ik}	$\hat{\pi}_{ik} (\times 10^{-4})$
Troll fishery										
1 Oct. – 31 Dec.	1988	4764	—	2333	0.49	0.90	22	4.118	3	0.640
1 Jan. – 14 Apr.	1989	2487	—	888	0.36	0.97	4	0.749	1	0.213
4–17 June	1989	3495	—	1423	0.41	1.00	8	1.497	8	1.708
18 June – 2 July	1989	4063	—	2024	0.50	0.98	7	1.310	3	0.640
3 – 23 July	1989	9513	—	3862	0.41	0.99	6	1.123	5	1.067
1 Oct. – 31 Dec.	1989	4303	—	1601	0.37	0.86	0	0	0	0
1 Jan. – 22 Apr.	1990	5610	—	2504	0.45	1.00	0	0	0	0
3–16 June	1990	2984	—	1652	0.55	1.00	1	0.187	1	0.213
17 June – 1 July	1990	7205	—	3842	0.53	0.92	5	0.936	1	0.213
Sport fishery										
24 Apr. – 4 June	1989	1011	0.12	329	0.33	1.00	5	0.936	2	0.427
5–18 June	1989	1329	0.20	193	0.15	1.00	3	0.562	2	0.427
19 June – 2 July	1989	935	0.20	88	0.09	0.87	3	0.562	2	0.427
3 July – 30 Sept.	1989	2220	0.15	238	0.11	1.00	2	0.374	0	0
7 May – 3 June	1990	1704	0.24	352	0.21	1.00	1	0.187	2	0.427
4–17 June	1990	2039	0.18	547	0.27	1.00	0	0	1	0.213
18 June – 2 July	1990	4566	0.17	488	0.11	1.00	1	0.187	0	0

$$(2.4) \quad \theta_{j(\text{opt})} = \frac{C}{\sum_k c_k M_k \gamma_{jk}}$$

Note that when $j = k$ in eq. 2.4, $\gamma_{jk} = 1$ and $\gamma_{kj} = \gamma_{jk}^{-1}$. Optimal numbers to tag are thus $R_{j(\text{opt})} = M_j \theta_{j(\text{opt})}$. Similar equations to calculate optimal tagging rates for wild cohorts were not derived because there has been little demand for sums of estimated harvests across wild cohorts.

Methods in this section and the next are illustrated with harvests from 1988 to 1990 from two cohorts of chinook salmon caught in troll and sport fisheries near Ketchikan, Alaska. Five- and 6-year-olds represent almost all harvest of chinook salmon in these fisheries. In 1986, 119 100 chinook salmon smolts, 53 428 (44.86%) of which carried CWTs, were released from Whitman Lake Hatchery near Ketchikan; in the same year, 930 072 chinook salmon smolts were released from nearby Neets Bay Hatchery; 46 845 (5.04%) had CWTs. From October 1988 through September 1989, 75 tags were recovered from both cohorts (Table 2) as 5-year-olds in troll and sport fisheries (by convention, chinook salmon become 1 year older on 1 October). Exploitation and catch-sampling rates in 1989–1990 were similar to rates a year earlier, yet only 13 tags were recovered from 6-year-olds. The same maturity schedule is indicated from tags recovered inside hatcheries. Parentage of these cohorts comes from stocks in the nearby Unuk and Chickamin rivers, fish that rear mainly within the inside waters of southeast Alaska.

We begin by assessing risk of not detecting significant harvest from a hypothetical cohort from Whitman Lake Hatchery. Considering the low abundance of 6-year-olds relative to younger fish for our past cohorts, harvest is likely negligible in strata from October 1989 onward; the relatively few tags recovered after October 1989 bear this out (Table 2). Thus, these strata were ignored when assessing risk. If $c_j = \$0.20$ and $C = \$20\,000$, $R_j = 100\,000$ from the in-

verted cost function. If no changes are anticipated from past survival, exploitation, sampling, or decoding rates experienced by the Whitman Lake cohort released in 1986, risk of not detecting some significant harvest in a new cohort from that hatchery would be 3% with 100 000 smolts tagged (Table 3A). If \$10 000 is spent to tag fish, 50 000 smolt will be tagged, and risk increases to 28% with some harvest in the sport fishery most in jeopardy of going undetected. Estimated survival rate \hat{S}_j for smolts released from Whitman Lake Hatchery in 1986 was 0.12, twice the lowest estimate of 0.06 for cohorts from local hatcheries released during 1980–1989 (as calculated by the Pacific Salmon Commission with methods described in Chinook Technical Committee (1988)). If this lowest rate is anticipated for a new cohort (a conservative assumption) with 100 000 tagged members (\$20 000 spent tagging fish), $\Delta(S_j) = 0.5$, and β_j would balloon to 28%. Adding an additional \$10 000 to tag smolts would shrink risk back to 9%, but if the same \$10 000 would double catch-sampling rates in the sport fishery instead ($\Delta(\phi_j) = 2.0$), risk would be 6% of missing some significant harvest (Table 3B).

Our example on linking precision of estimated harvest to tagging rates is again based on a hypothetical cohort from Whitman Lake ($M_j = 200\,000$) with anticipated circumstances as reflected in Table 3B: a doubling of catch-sampling rates ($\Delta(\phi_j) = 2$) in upcoming sport fisheries over those in 1989–1990 ($\Delta(\phi_j) = 1$ for troll fisheries), a conservative adjustment to survival rate ($\Delta(S_j) = 0.5$), and no changes in exploitation ($\Delta(U_{ij}) = 1$) or decoding rates ($\Delta(\lambda_i) = 1$) from previous years. Increased catch sampling in onsite creel surveys can (but not always) shrink $\text{Var}(\hat{N}_i)$. Here, doubling catch-sampling rates in the sport fishery has reduced $CV(\hat{N}_i)$ by a third. If \$20 000 is available to tag the Whitman Lake cohort, and per-unit cost is \$0.20, tagging 100 000 smolt ($= C/c_j$) produces a variance $\text{Var}(\hat{T}|C)$ of 2557 (Table 4) with relative precision $CV(\hat{T}|C)$ of 0.12. Calculations

Table 3. Three printouts of a computer spreadsheet used to calculate risk β_j of not detecting significant harvest in troll and sport fisheries near Ketchikan, Alaska, from a hypothetical cohort of chinook salmon based on information from a past cohort released from Whitman Lake Hatchery.

Stratum	Year	$\hat{\pi}_{ij} (\times 10^{-4})$	$\Delta(S_j)$	$\Delta(U_{ij})$	$\Delta(\phi_i)$	$\Delta(\lambda_i)$	$\pi_{ij} (\times 10^{-4})$	$1 - (1 - \pi_{ij})^{R_j}$
(A) Statistics: $R_j = 100\ 000$; $1 - \beta_j = 0.969$; $\beta_j = 0.031$								
Troll 1	First	4.118	1.0	1.0	1.0	1.0	4.118	>0.999
Troll 2	Second	0.749	1.0	1.0	1.0	1.0	0.749	0.999
Troll 3	Second	1.497	1.0	1.0	1.0	1.0	1.497	>0.999
Troll 4	Second	1.310	1.0	1.0	1.0	1.0	1.310	>0.999
Troll 5	Second	1.123	1.0	1.0	1.0	1.0	1.123	>0.998
Sport 1	Second	0.936	1.0	1.0	1.0	1.0	0.936	0.996
Sport 2	Second	0.562	1.0	1.0	1.0	1.0	0.562	0.996
Sport 3	Second	0.562	1.0	1.0	1.0	1.0	0.562	0.976
Sport 4	Second	0.374	1.0	1.0	1.0	1.0	0.374	0.976
Product (eq. 1.1)								0.969
(B) Statistics: $R_j = 100\ 000$; $1 - \beta_j = 0.941$; $\beta_j = 0.059$								
Troll 1	First	4.118	0.5	1.0	1.0	1.0	2.059	>0.999
Troll 2	Second	0.749	0.5	1.0	1.0	1.0	0.374	0.976
Troll 3	Second	1.497	0.5	1.0	1.0	1.0	0.749	0.999
Troll 4	Second	1.310	0.5	1.0	1.0	1.0	0.655	0.999
Troll 5	Second	1.123	0.5	1.0	1.0	1.0	0.562	0.996
Sport 1	Second	0.936	0.5	1.0	2.0	1.0	0.936	>0.999
Sport 2	Second	0.562	0.5	1.0	2.0	1.0	0.562	0.996
Sport 3	Second	0.562	0.5	1.0	2.0	1.0	0.562	0.996
Sport 4	Second	0.374	0.5	1.0	2.0	1.0	0.374	0.976
Product (eq. 1.1)								0.941
(C) Statistics: $R_j = 12\ 000$; $1 - \beta_j = 0.084$; $\beta_j = 0.916$								
Troll 1	First	4.118	1.0	1.0	1.0	1.0	4.118	0.993
Troll 2	Second	0.749	1.0	1.0	1.0	1.0	0.749	0.593
Troll 3	Second	1.497	1.0	1.0	1.0	1.0	1.497	0.834
Troll 4	Second	1.310	1.0	1.0	1.0	1.0	1.310	0.792
Troll 5	Second	1.123	1.0	1.0	1.0	1.0	1.123	0.740
Sport 1	Second	0.936	1.0	1.0	2.0	1.0	1.872	0.894
Sport 2	Second	0.562	1.0	1.0	2.0	1.0	1.123	0.740
Sport 3	Second	0.562	1.0	1.0	2.0	1.0	1.123	0.740
Sport 4	Second	0.374	1.0	1.0	2.0	1.0	0.749	0.593
Product (eq. 1.1)								0.084

cover all strata where 5- and 6-year-old chinook salmon could be harvested, even though harvests of 6-year-olds are anticipated to be negligible.

Instead of \$20 000 to tag a hatchery-produced cohort, let funding for nine person-months of labor ($C = 9$) be available to tag a wild cohort. Say a crew of two can capture and tag an anticipated 2% of smolt emigrating over a 6-week period ($q_j = 0.02$), and this single unit of labor costs 3 person-months ($c_j = 3$). From eq. 2.2, $G(\hat{p}_{ij}) = 0.396 = [3 - \lambda_i \phi_i (0.02) 9] [\lambda_i \phi_i 93 (0.02) 9]^{-1}$ for the first troll stratum, provided $\{\pi_{ij}\}$ for the hypothetical cohort from Whitman Lake Hatchery (Table 3B) are germane to this hypothetical wild cohort. In a wild cohort, the 200 000 smolts would be estimated, not known with negligible error as in a hatchery. Variance $\text{Var}(\hat{T}|C) = 21\ 724$ for this "wild" cohort, making $\text{CV}(\hat{T}|C) = 0.35$. As imprecise as estimated harvest would be for this wild cohort, risk of undetected harvest would be worse. Using the inverted cost function, 6% ($= \theta_j(100) =$

$q_j c_j^{-1} C(100)$) of 200 000 smolts, or only 12 000 ($= \theta_j \hat{M}_j$) smolts, would be tagged. Considerable harvest would most likely go undetected ($\beta_j = 0.92$), even if survival rates are average and sampling rates in the sport fishery are doubled (Table 3C).

A hypothetical cohort from Neets Bay Hatchery ($M_k = 1\ 000\ 000$) is now added to the example to demonstrate finding optimal tagging rates across cohorts. If \$20 000 is available to tag both cohorts and per-unit cost is \$0.10 for both, tagging 21% of the Whitman Lake cohort and 16% of the Neets Bay cohort should minimize variance $\text{Var}(\hat{T}|C)$. In this instance, 11 929 is the denominator in eq. 2.3 ($= [c_k M_k (1423.0)]^{1/2}$ with 1423.0 calculated as per Table 4), 8914 ($= [c_j M_j (3973.1)]^{1/2}$) the numerator, and $\gamma_{jk} = 0.747$. From eq. 2.4, $\theta_{j(\text{opt})} = C[c_j M_j (1) + c_k M_k (0.747)]^{-1} = 0.211$ and $\theta_{k(\text{opt})} = C[c_j M_j (1.338) + c_k M_k (1)]^{-1} = 0.158$. Variance $\text{Var}(\hat{T}|C)$ can be found by inserting optimal tagging rates into eq. 1.8 (or optimal allocation of monies into eq. 2.1) for

Table 4. Printout of a computer spreadsheet used to anticipate harvest in troll and sport fisheries near Ketchikan, Alaska, from a hypothetical cohort of chinook salmon yet to be released from Whitman Lake Hatchery when $M_j = 200\,000$, $c_j = \$0.20$, and $C = \$20\,000$.

Stratum	Year	$G(\hat{N}_i)$	ϕ_i	λ_i	$\pi_{ij} (\times 10^{-4})$	r_{ij}	$G(\hat{p}_{ij})$	$\text{Var}(\hat{r}_{ij})$	$\frac{r_{ij}[1 + G(\hat{N}_i)]}{\lambda_i \phi_i}$
Troll 1	First	—	0.49	0.90	2.059	93	0.038	330	211.7
Troll 2	Second	—	0.36	0.97	0.374	21	0.221	101	61.3
Troll 3	Second	—	0.41	1.00	0.749	37	0.106	142	89.1
Troll 4	Second	—	0.50	0.98	0.655	27	0.115	82	54.6
Troll 5	Second	—	0.41	0.99	0.562	28	0.142	109	68.2
Troll 6	Second	—	0.37	0.86	0	0	0	0	0
Troll 7	Third	—	0.45	1.00	0	0	0	0	0
Troll 8	Third	—	0.55	1.00	0.093	3	0.780	9	6.1
Troll 9	Third	—	0.53	0.92	0.468	19	0.162	60	39.4
Sport 1	Second	0.0064	0.66	1.00	0.936	28	0.072	63	42.7
Sport 2	Second	0.0178	0.30	1.00	0.562	37	0.151	241	125.5
Sport 3	Second	0.0178	0.18	0.87	0.562	72	0.164	952	468.0
Sport 4	Second	0.0100	0.22	1.00	0.374	34	0.238	289	156.1
Sport 5	Third	0.0256	0.42	1.00	0.187	9	0.422	36	22.0
Sport 6	Third	0.0144	0.54	1.00	0	0	0	0	0
Sport 7	Third	0.0128	0.22	1.00	0.187	17	0.476	143	78.3
Sum						425		2557	1423.0

Table 5. Optimal tagging rates and related statistics under three scenarios of per-unit tagging costs c_j with \$20 000 available to tag two hypothetical cohorts of chinook salmon from Whitman Lake ($M_j = 200\,000$) and Neets Bay ($M_j = 1\,000\,000$) hatcheries caught as 5- and 6-year-olds in troll and sport fisheries near Ketchikan, Alaska (total harvest $T = 1596$).

	Whitman Lake	Neets Bay	Whitman Lake	Neets Bay	Whitman Lake	Neets Bay
c_j	\$0.10	\$0.10	\$0.20	\$0.10	\$0.10	\$0.20
γ	0.747	1.338	1.057	0.946	0.528	1.893
$\theta_{j(\text{opt})}$	0.211	0.158	0.137	0.145	0.159	0.084
$R_{j(\text{opt})}$	42 250	157 750	27 500	145 000	31 850	84 100
Cost	\$4 225	\$15 775	\$5 500	\$14 500	\$3 185	\$16 815
$\text{Var}(\hat{T} C)$		32 182		37 889		55 884
$\text{CV}(\hat{T} C)$		0.112		0.122		0.148

both cohorts and calculating with eqs. 1.5–1.7. Results from this and two other scenarios are listed in Table 5. Precision profiles for pairs of cohorts can be calculated with spreadsheets to graphically show how sensitive precision is to allocations of monies among cohorts (Fig. 1). In the three cost-per-unit scenarios in Table 5, allocating 10–40% of monies to tagging the hypothetical Whitman Lake cohort (the balance would be used to tag the Neets Bay cohort) produces minimum or near-minimum variance. This same technique can be used to investigate sensitivity of optimal tagging rates to choices for $\{\pi_{ij}\}$.

3. Catch sampling

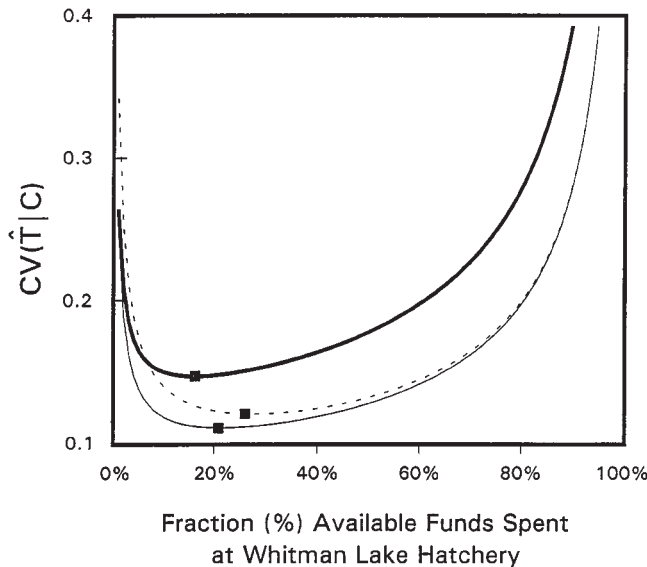
The cost C_i of catch sampling during a stratum can be modeled as a nonlinear, allometric function of sampling rate $C_i = a_i \phi_i^{b_i}$ where a_i is cost of sampling the entire catch ($n_i = N_i$) and $b_i \geq 1$. Catch sampling is usually concentrated at locations within a stratum where most salmon can be inspected; additional samples would come from locations with fewer salmon and higher per-unit costs of sampling. The

more divergent the per-unit costs of sampling among locations, the greater the value of b_i . For instance, if sampling half the catch in stratum i ($\phi_i = 0.5$) costs half the maximum ($= a_i/2$), $b_i = 1$ and the allometric cost function becomes linear; if sampling half the catch costs a quarter of the maximum, $b_i = 2$, and if sampling a half costs an eighth, $b_i = 3$. This cost function can be inserted into eq. 1.5 to link expenditures (sampling effort) to variance $\text{Var}(\hat{T})$:

$$(3.1) \quad G(\hat{p}_{ij}) = \frac{1 - \lambda_i \theta_j \sqrt[b_i]{C_i/a_i}}{\lambda_i \theta_j r_{ij} \sqrt[b_i]{C_i/a_i}}$$

When sampling occurs at the same locations in each fishing period (stratum), the cost function can be simplified to $C_i = a_i \phi_i^{b_i}$. Differences in maximum cost a_i among exclusively time-defined strata are only related to the length of each stratum in days. If the same sampling effort is expended each day in all such time-defined strata, sampling tends to be self-weighted (proportional), the cost function becomes $C_i = a_i \phi_i^{b_i}$, and for the season, $\sum_i C_i = \phi^b \sum_i a_i = C = a \phi^b$.

Fig. 1. Precision profiles for estimated harvest \hat{T} as functions of how funds to tag smolts are allocated between two hypothetical cohorts of chinook salmon at Whitman Lake and Neets Bay hatcheries. With \$20 000 to tag smolts, profiles were calculated for three scenarios of per-unit costs: \$0.10 in both hatcheries (thin line), \$0.20 and \$0.10 (broken line), and \$0.10 and \$0.20 (thick line). Optimal allocation is marked on each profile.



Sample sizes in catch-sampling programs are usually proportional or nearly proportional to catch because the same technicians work from beginning to end of each season in each program sampling mostly at the same sites. If catch sampling is self-weighted, the seasonal cost function can be used to link expenditures to variance of estimated harvest when only a single, seasonal estimate of catch will be available. Plugging the seasonal cost function into eq. 1.8 and summing anticipated harvests r_{ij} across strata produces a seasonal statistic $G(\bar{p}_j)$ for each cohort:

$$(3.2) \quad G(\bar{p}_j) = \frac{\sum_i \frac{r_{ij}}{\lambda_i} (1 - \lambda_i \theta_j \sqrt{C/a})}{\theta_j \sqrt{C/a} \left(\sum_i r_{ij} \right)^2}$$

where \bar{p}_j is a pooled, scaled statistic $\left(\sum_i m_{ij} \lambda_i^{-1} / \sum_i n_i \right)$ used here because sampling is self-weighted. Note that if all $\lambda_i = 1$, \bar{p}_j and eq. 3.2 represent a simple pooling of data from catch sampling across strata. In eqs. 1.5–1.7, terms $G(\hat{N}) \rightarrow G(\hat{N}_i)$, $G(\bar{p}_j) \rightarrow G(\hat{p}_{ij})$, and $\sum_i r_{ij} \rightarrow r_{ij}$ to complete the link between sampling effort and variance.

When catch sampling can be concentrated among strata, there is an optimal allocation of sampling effort that minimizes variance $\text{Var}(\hat{T}|C)$. Given the cost function $C = \sum_i a_i \phi_i^{b_i}$, the relationship between sampling rates for any two strata i and t that minimizes variance is

Table 6. Statistics used to determine an optimal allocation of catch sampling effort between troll and sport fisheries in the same year near Ketchikan, Alaska, exploiting two hypothetical cohorts of 5- and 6-year-old chinook salmon from Whitman Lake Hatchery where 45 and 20%, respectively, had been tagged.

	Troll fishery	Sport fishery
a_i	17	21
b_i	1.5	1.4
λ_i	1.0	1.0
Summed over strata within fishery		
$\sum r_{ij}$ (5-year-olds)	206	171
$\sum r_{ij}$ (6-year-olds)	22	26
Redefined stratum corresponds to fishery		
$G(\hat{N}_i)$	0	0.0098
$\sum_j \frac{r_{ij}}{\theta_j} (1 + G(\hat{N}_i))$	567.8	515.0

$$(3.3) \quad \frac{\phi_t^{b_t+1}}{\phi_i^{b_i+1}} =$$

$$\frac{a_i b_i \lambda_i \sum_j r_{ij} \theta_j^{-1} (1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i) G(\theta_j^{-1}))}{a_t b_t \lambda_t \sum_j r_{ij} \theta_j^{-1} (1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i) G(\theta_j^{-1}))} = \omega_{it}$$

Derivation of eq. 3.3 is described in the Appendix. Inspection of this equation shows that there should be (in rough order of importance) more sampling in strata with more harvest, lower per-unit sampling costs, and less precise knowledge of catch. From eq. 3.3, $\phi_t = \frac{b_t+1}{\sqrt{\omega_{it} \phi_i^{b_i+1}}}$, and if this result is inserted back into the cost function:

$$(3.4) \quad C = \sum_i a_i \left(\frac{b_i+1}{\sqrt{\omega_{it} \phi_i^{b_i+1}}} \right)^{b_i}$$

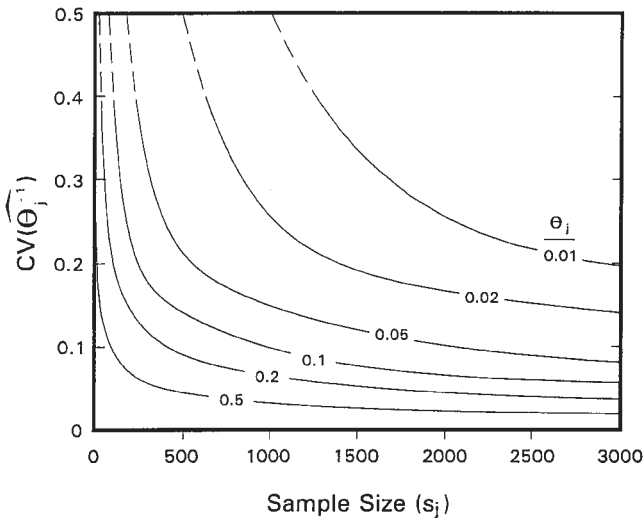
Note that when $i = t$ in eq. 3.4, $\omega_{it} = 1$. Optimal solutions $\{\phi_{i(\text{opt})}\}$ to eq. 3.4 can be found quickly through trial and error with a spreadsheet and then inserted back into the cost function to provide an optimal allocation of sampling effort (expenditures):

$$(3.5) \quad C_{i(\text{opt})} = a_i \phi_{i(\text{opt})}^{b_i}$$

which in turn can be inserted into eq. 3.1 to calculate variance through eqs. 1.5–1.8.

Optimally allocating catch-sampling effort among strata can be demonstrated by treating sampling programs for troll and sport fisheries in our examples as competing for funds. Five technicians can inspect all catch sold to two processors in Ketchikan representing about 45% of the entire troll catch (from Table 2); sampling 90% of the entire catch would require 14 technicians, making $a_i = 17$ and $b_i = 1.5$. Seven harbors provide access to the sport fishery, and 2 person-days of sampling effort would cover a fishing day at a harbor. Because sampling all catch landed in the sport fishery requires 14 technicians during the week and seven more for the

Fig. 2. Estimated precision $\hat{CV}(\hat{\theta}_j^{-1})$ as a function of sample size s_j in a field-sampling program. For each sample size, $\hat{CV}(\hat{\theta}_j^{-1})$ was calculated from 5000 Monte Carlo simulations with each simulation producing an estimate $\hat{\theta}_j^{-1}$ drawn from $\text{binom}(y|s_j, \theta_j)$. Contours were generated by smoothing simulations through fitting the model $\hat{CV}(\hat{\theta}_j^{-1}) = g_1 \exp^{-g_2 s_j} + g_3 \exp^{-g_4 s_j} + g_5 \exp^{-g_6 s_j} + g_7$ to estimates and then generating contours using all estimated parameters $\{\hat{g}\}$ significantly different from zero ($\alpha = 0.05$). Broken sections of contours correspond to sample sizes that produce estimates $\hat{\theta}_j^{-1}$ with bias $>10\%$ where bias is estimated as the difference between $\hat{\theta}_j^{-1}$ and its average over each set of 5000 simulations.



weekend as per union rules, $a_i = 21$. Two technicians have on average sampled 18% across strata (from Table 2), making $2 = 21 (0.18)^{b_i}$ and $b_i = 1.4$. Allometric constants are the same across strata within each fishery because the same sampling opportunities will be present as time passes. The sampling fraction ϕ_i is also time-invariant because a constant sampling effort will be applied in each fishery. Under these common circumstances, each fishery can be treated as a single, combined "stratum" when determining the optimal allocation of sampling effort between them. Statistics in Table 2 for the past cohort from Whitman Lake Hatchery were separated by age to create two new hypothetical cohorts exposed to fishing in the same fiscal year as 5- and 6-year-olds (Table 6). Tagging rates for these two new hypothetical cohorts are 45 and 20%, respectively, $\Delta(S_j) = 0.5$, all $\lambda_i \leftarrow 1$, and all harvests (Table 4) were summed across strata within each fishery for each age group. The $G(\hat{N}_i)$ listed for the sport fishery in Table 6 is the average of two statistics, each published for onsite creel surveys in 1989 and 1990. From eqs. 3.3–3.5 with $C = 7$ technicians, $\omega_{it} = 0.79$, $\{\phi_{i(\text{opt})}\} = (0.34, 0.31)$, and $\{C_{i(\text{opt})}\} = (3, 4)$ for the troll and sport fishery, respectively. Increased expenditure indicated for the sport fishery in the optimal allocation would also improve precision of estimated catch from the creel survey, thereby recursively affecting the optimal allocation; however, the affect here would be negligible.

4. Field sampling

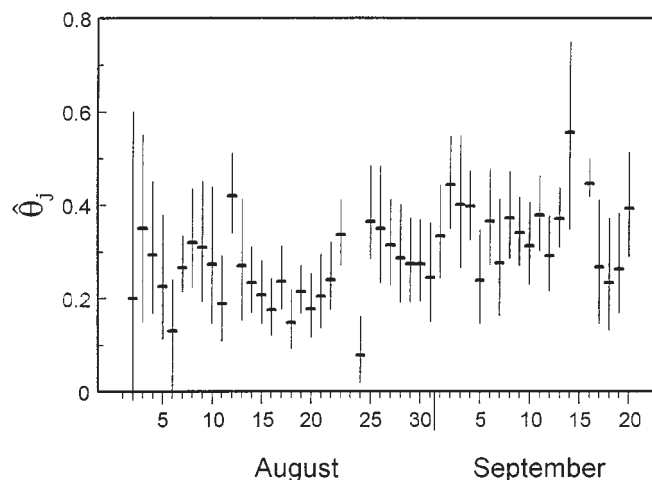
Relationships between sample sizes s_j in field-sampling

programs and precision of $\hat{\theta}_j^{-1}$ can be found by simulating $\hat{\theta}_j^{-1}$ where $\hat{\theta}_j^{-1} = s_j/y$ and $y \sim \text{binom}(s_j, \theta_j)$ as per Geiger (1990). Smoothed isopleths for precision as a function of sample size are presented in Fig. 2 for several simulations. Note that at small sample sizes, $\hat{\theta}_j^{-1}$ can be considerably biased (Bernard and Clark 1996). As pointed out by Geiger (1990), the contribution of $\text{Var}(\hat{\theta}_j)$ (and hence, $\text{Var}(\hat{\theta}_j^{-1})$) to $\text{Var}(\hat{r}_{ij})$ becomes negligible as sample sizes increase, with the effect accelerated at larger θ_j . For this reason, $G(\hat{\theta}_j^{-1})$ is most often considered zero for cohorts produced in hatcheries where sampling large numbers of fish is inexpensive. Although capturing wild, adult salmon is not cheap, the most expensive part of field sampling is often getting to the sampling site with equipment. Once there, the marginal cost of capturing fish is often sufficiently low to realistically consider sample sizes large enough to make $G(\hat{\theta}_j^{-1})$ negligible.

5. Minimizing and detecting bias

The fraction of a cohort tagged as juveniles can be significantly higher than the fraction of returning adults with tags (see Schnute et al. 1990); the correct fraction for estimating harvest is the fraction for adults. Natural loss of adipose fins (Blankenship 1990; Kaill et al. 1990), tag-induced mortality (Blankenship and Hanratty 1990; Peltz and Miller 1990; Vincent-Lang 1993), tag loss (Blankenship 1990; Kaill et al. 1990; Peltz and Miller 1990), size-selective tag loss (Blankenship 1990), and tag-induced straying (Morrison and Zajac 1987; Morrison et al. 1990) have been identified or postulated as potential causes of a drop in the tagged fraction over time. Collectively, these authors suggest procedures for tagging programs to mitigate these problems: (i) juveniles should be inspected for naturally missing adipose fins, (ii) fish to be tagged should be graded by size and appropriately sized head molds used to insert tags, (iii) some tagged fish should be held to determine tag-induced mortality and tag loss, and (iv) some tagged fish should be periodically dissected and histologically examined to ascertain proper tag size and placement. As prudent as these suggestions are, the only sure way to minimize bias from tag-induced mortality or tag loss is to estimate θ_j from mature adults on a hatchery rack or adults caught in a field-sampling program. Unfortunately, there is no way to confidently correct harvest estimates for tag-induced straying because tag-induced confusion could start before, during, or after affected fish had passed through fishing grounds. Morrison and Zajac (1987) and Morrison et al. (1990) found a high incidence of damaged olfactory organs in juvenile salmon with improperly placed tags and postulated straying as a consequence. Damage would be less likely for larger juveniles, such as those of coho (*O. kisutch*) and chinook salmon. There is little evidence of any straying, tag-induced or otherwise, by chinook salmon serving as examples in this paper. Only nine hatchery-produced chinook salmon were found in the Chickamin and Unuk rivers in 9 years (1985–1993) of field sampling (Pahlke 1995), even though these rivers were the source of brood stock for local hatcheries, including Whitman Lake and Neets Bay hatcheries. During the

Fig. 3. Daily estimates of θ_j and their estimated 95% confidence intervals for coho salmon migrating up the Kenai River, Alaska, in 1995 (from Carlon and Hasbrouck 1997).



same years, no wild chinook salmon tagged in the Unuk or Chickamin River were recovered in rivers other than their natal watershed.

Harvest estimates will be unbiased only when θ_j is time-invariant as the cohort passes through the fishing grounds, a feat much harder to achieve for wild cohorts. A wild cohort can be composed of several populations, each with its own migratory timing as smolts and as adults. Capturing smolts on the main river below rearing areas will produce a representative sample only if probability of a smolt being tagged is unrelated to when it was captured. To keep that probability time-invariant, the mix of gear used to capture smolts should not be changed from day to day and should be fished daily with equal effort on both banks throughout the migration. This strategy will not guarantee tagging a representative sample, but will increase its likelihood, especially if there is some overlap in migratory timing of populations. McPherson and Bernard (1996) described tagging chinook salmon smolts on a large river just above tidewater using this strategy. Another strategy is to tag juveniles or presmolts in rearing areas where they are more easily captured. Although this strategy often produces more tagged fish, these tagged fish will only be representative of all juveniles (and bias minimized) if populations within the watershed are present at the tagging site in proportion to their abundance throughout the watershed. Success in this instance requires a peculiar set of circumstances, including luck in finding the right sampling spot. Carlon and Hasbrouck (1997) presented an example of successfully applying this strategy to coho salmon in a large river.

Regardless of the strategy used to tag wild juveniles, a temporal trend in θ_j can be detected with field sampling. If sampling occurs upstream in Z tributaries on spawning populations with dissimilar migratory timing, a $Z \times 2$ contingency table can be used to detect statistically significant differences in θ_j among populations. If sampling occurs downstream of spawning grounds, periodic estimates of θ_j (Fig. 3) can be regressed against time. Failure to reject the null hypotheses in either of these tests confirms the absence

of a temporal trend in θ_j , so long as sample size is sufficiently large to detect a meaningful trend. A trend is meaningful if it implies meaningful bias in estimated harvest. Sample sizes in field-sampling programs can be quite large, giving hypothesis tests more power than needed; therefore, rejecting the null hypothesis because of statistical significance does not necessarily imply meaningful bias in estimated harvests. For example, $\{\hat{\theta}_j\}$ in Fig. 3 have a statistically significant linear trend (logistic regression; $s_j = 4819$, $P < 0.0001$); however, the shallowness of this trend and variation about it indicate that resulting bias in \hat{T} should be relatively small.

Bias in \hat{T} from catch sampling can result from strata being too big, from concentrating sampling effort at certain locations within strata, and from undetected harvest. How well strata have been defined can be determined by fitting log-linear models with Poisson error to recovery data (Cormack and Skalski 1992). A scaled residual error (deviance) greater than 1 is evidence for strata being too physically large. However, the tendency for catch sampling to be proportional should blunt much of any bias from overdispersed recoveries in oversized strata. Concentrating sampling at a few locations within a temporal stratum will also bias estimated harvest if the pattern of tags in catch landed at other locations is dissimilar. Periodically sampling all (or nearly all) catch in a stratum can reveal these differences. If these differences imply meaningful bias, the solution is to split strata in future catch-sampling programs. Suspicions of undetected harvest arise when catch sampling recovers only one tag in each of a few strata and no tags elsewhere. Undetected harvest is certain when one or more singleton recoveries expand to a harvest estimate for a single cohort impossibly larger than the catch of all salmon in that stratum. Without this clear signal, suspicions of undetected harvest can not be confirmed without independent information, such as voluntarily returned CWTs, on the presence of that cohort or its predecessors over the fishing grounds at the appropriate time.

6. Some perspectives

There is an asymmetry in CWT programs that puts most of the responsibility for success on those in charge of tagging programs. Tagging salmon is the initial step in estimating harvest a year or several years later. If estimates with negligible bias and acceptable precision are to be attained, persons in charge of tagging must anticipate future fisheries and catch-sampling programs. In contrast, tagging rates are known when planning catch-sampling programs (at least with hatchery-produced cohorts), and these programs are often planned (although not budgeted) only a few weeks before fisheries begin. Failure from tagging too few young salmon can sometimes be avoided by inspecting more adults in the catch, provided persons in control of catch-sampling programs are willing and able to do so. Catch sampling is often concentrated on simultaneously estimating harvest from many cohorts and is limited by marginal improvements in sample size (nonlinear cost functions). Even if catch-sampling staff are sympathetic, there are limits to how far catch sampling can compensate for inadequate tagging. In

every situation, there is a tagging rate below which sampling 100% of catch would still be insufficient to salvage harvest estimates. This pitfall should be of special concern for persons responsible for tagging wild cohorts. As well as how many salmon are tagged, tagging programs are responsible for how young salmon are selected for tagging. No catch-sampling program of any size or design can salvage estimates if an unrepresentative sample of young salmon has been tagged.

Risk of undetected harvest should always be considered along with precision of estimated harvest when planning tagging programs. After all, how worthwhile is a harvest estimate and an estimate of its variance that are biased low from undetected harvest? However, excluding some harvest when determining risk can be a practical approach to planning. A few members of a cohort or cohorts will be harvested in strata on the margins of migrations or be harvested in odd years at odd ages. If these marginal strata or minor age groups are considered when planning, costs of tagging programs can be prohibitive, so there is a financial reason for ignoring some strata or age groups when planning. Only strata with harvests of 5-year-old chinook salmon near Ketchikan were considered important in our demonstration on determining risk of not detecting harvest. In contrast, strata with harvests of these other fish (represented by 6-year-old chinook salmon) were included when linking costs to precision of estimated harvest. Risk of undetected harvest is largely a function of small harvests, while precision of estimates is mostly a function of large ones. Methods described in Sections 1–3 can be exercised in spreadsheets to weigh monies spent against accuracy and precision to judge what strata or age groups represent important harvest.

Although our planning methods were largely demonstrated with chinook salmon migrating through fishing grounds near Ketchikan, Alaska, applying these methods to other species and other fisheries is only a matter of scale. Considering more fisheries when planning tagging and catch-sampling programs means adding more catch strata to analyses. More cohorts require more of the same kind of calculations. Planning programs for species that only remain a single year at sea means that analyses will concentrate on strata representing only a single year of returns. For those species that mature over several years, information from field-sampling programs can be accumulated over several years for each cohort when estimating the fraction of a cohort tagged. In short, scope of planning can be expanded or contracted as fits circumstances with no substantive change in the mathematics presented here.

Methods presented here on anticipating probabilities of recovering CWTs are deterministic. These methods are easily understood, require only a single year of historical data to use, readily incorporate professional judgement, and are quick to employ. Obviously, more rigorous stochastic methods, such as simulation, can also be used to establish the $\{\pi_{ij}\}$ when a wealth of data and ample time for planning are available. However, planning CWT programs must be complete before juvenile salmon reach a certain size or adults reach fishing grounds. Under press of these deadlines, quick, easily understood methods are often the planning methods of choice.

Acknowledgements

We thank Sandy Sonnichsen, David Evans, Jim Blick, and Steve Fleischman for their reviews of earlier drafts. We are grateful to the referees whose comments greatly improved the paper, especially Jon Schnute whose diligence greatly improved the clarity of our presentation. We also thank Allen Bingham, Paul Suchanek, John Carlile, Jim Hasbrouck, and Pat Hansen for their assistance with data in our examples. We are also grateful to Whitman Lake and Neets Bay hatcheries for providing information for our examples; our use of their statistics was designed to demonstrate our planning methods and is not necessarily indicative of their operation. This work was partially supported with funding authorized under the U. S. Federal Aid to Sport Fish Restoration Act.

References

- Bernard, D.R., and Clark, J.E. 1996. Estimating salmon harvest with coded-wire tags. *Can. J. Fish. Aquat. Sci.* **53**: 2323–2332.
- Bernard, D.R., Arvey, D.W., and Holmes, R.A. 1993. Operational planning: the Dall River and rescue of its sport fishery. *Fisheries (Bethesda)*, **18**(2): 6–12.
- Blankenship, H.L. 1990. Effects of time and size on coded wire tag loss from chinook and coho salmon. *Am. Fish. Soc. Symp.* **7**: 237–243.
- Blankenship, H.L., and Hanratty, P.R. 1990. Effects on survival of trapping and coded wire tagging coho salmon smolts. *Am. Fish. Soc. Symp.* **7**: 259–261.
- Carlson, J.A., and Hasbrouck, J.J. 1997. Stock assessment of coho salmon from the Kenai River, Alaska, 1995. *Fish. Data Ser. No. 97-7*, Alaska Department of Fish and Game, Anchorage, Alaska.
- Chinook Technical Committee. 1988. 1987 annual report. *Rep. TCCHINOOK (88)-2*, Pacific Salmon Commission, Vancouver, B.C.
- Clark, J.E., and Bernard, D.R. 1987. A compound multivariate binomial hypergeometric distribution describing coded wire tag recovery from commercial salmon catches in southeastern Alaska. *Inf. Leaflet No. 261*, Alaska Department of Fish and Game, Juneau, Alaska.
- Cormack, R.M., and Skalski, J.R. 1992. Analysis of coded wire tag returns from commercial catches. *Can. J. Fish. Aquat. Sci.* **49**: 1816–1825.
- Geiger, H.J. 1990. Parametric bootstrap confidence intervals for estimating contributions to fisheries from marked salmon populations. *Am. Fish. Soc. Symp.* **7**: 667–676.
- Goodman, L.A. 1960. On the exact variance of products. *J. Am. Stat. Assoc.* **55**: 708–713.
- Johnson, J.K. 1990. Regional overview of coded wire tagging of anadromous salmon and steelhead in northwest America. *Am. Fish. Soc. Symp.* **7**: 782–816.
- Kaill, W.M., Rawson, K., and Joyce, T. 1990. Retention rates of half-length coded wire tags implanted in emergent pink salmon. *Am. Fish. Soc. Symp.* **7**: 253–258.
- Lapi, L., Hamer, M., and Johnson, B. 1990. Data organization and coding for a coastwide mark-recovery data system. *Am. Fish. Soc. Symp.* **7**: 720–724.
- McPherson, S.A., and Bernard, D.R. 1996. Production of coho salmon from the Taku River, 1994–1995. *Fish. Data Ser. No. 96-25*, Alaska Department of Fish and Game, Anchorage, Alaska.

- Morrison, J., and Zajac, D. 1987. Histologic effect of coded wire tagging in chum salmon. *N. Am. J. Fish. Manage.* **7**: 439–441.
- Morrison, J.K., Coyle, C.L., and Berton, S.E. 1990. Histological effect of tagging chum and coho salmon fry with coded wire tags. *Prog. Fish-Cult.* **52**: 117–119.
- Pahlke, K.A. 1995. Coded wire tagging studies of chinook salmon of the Unuk and Chickamin rivers, Alaska, 1983–1993. *Alaska Fish. Res. Bull.* **2**: 93–113.
- Peltz, L., and Miller, J. 1990. Performance of half-length coded wire tags in a pink salmon hatchery marking program. *Am. Fish. Soc. Symp.* **7**: 244–252.
- Pollock, K.H., Jones, C.M., and Brown, T.L. 1994. Angler survey methods and their applications in fisheries management. *Am. Fish. Soc. Spec. Publ.* No. 25.
- Schnute, J.T. 1992. Statistical analysis of embedded replicates in mark–recovery experiments. *Can. J. Fish. Aquat. Sci.* **49**: 432–442.
- Schnute, J.T., Mulligan, T.J., and Kuhn, B.R. 1990. An errors-in-variables model with an application to salmon hatchery data. *Can. J. Fish. Aquat. Sci.* **47**: 1453–1467.
- Seber, G.A.F. 1982. The estimation of animal abundance and related parameters. 2nd ed. Macmillan Publishing Co., New York.
- Vincent-Lang, D. 1993. Relative survival of unmarked and fin-clipped coho salmon from Bear Lake, Alaska. *Prog. Fish-Cult.* **55**: 141–148.

Appendix: Proofs for eqs. 2.3 and 3.3

The relationship between tagging rates for two or more hatchery-produced cohorts or between sampling rates in two or more catch strata that will minimize $\text{Var}(\hat{T})$ when costs are fixed can both be found with Lagrange's method (all notation is defined in Table 1). Combining eqs. 1.5–1.7:

$$(A.1) \quad \text{Var}(\bar{T}) = \sum_i \sum_j r_{ij}^2 [G(\hat{N}_i) + G(\hat{p}_{ij}) + G(\hat{\theta}_j^{-1})] \\ + G(\hat{N}_i)G(\hat{p}_{ij}) + G(\hat{N}_i)G(\hat{\theta}_j^{-1}) + G(\hat{p}_{ij})G(\hat{\theta}_j^{-1}) \\ + G(\hat{N}_i)G(\hat{p}_{ij})G(\hat{\theta}_j^{-1}) + 2 \sum_i \sum_j \sum_{k>j} r_{ij} r_{ik} G(\hat{N}_i)$$

To constrain this variance for fixed cost of tagging operations, both sides of the cost function $C = \sum_j c_j M_j \theta_j$ are multiplied by a constant L with the right-hand result added and the left-hand result subtracted from eq. A.1:

$$(A.2) \quad \text{Var}(\hat{T}|C) = \sum_i \sum_j r_{ij}^2 [G(\hat{N}_i) + G(\hat{p}_{ij}) + G(\hat{N}_i)G(\hat{p}_{ij})] \\ + 2 \sum_i \sum_j \sum_{k>j} r_{ij} r_{ik} G(\hat{N}_i) + L \sum_j c_j M_j \theta_j - LC$$

Note that $G(\hat{\theta}_j^{-1})$ is presumed equal to zero for hatchery-produced cohorts. When eq. A.2 is differentiated with respect to the tagging rate θ_j :

$$(A.3) \quad \frac{\partial \text{Var}(\hat{T}|C)}{\partial \theta_j} = Lc_j M_j + \sum_i r_{ij}^2 [1 + G(\hat{N}_i)] \frac{\partial G(\hat{p}_{ij})}{\partial \theta_j}$$

Note that r_{ij} and $G(\hat{N}_i)$ are not functions of θ_j . From eq. 1.8:

$$G(\hat{p}_{ij}) = \frac{1 - \lambda_i \phi_i \theta_j}{\lambda_i \phi_i \theta_j r_{ij}}$$

$$\frac{\partial G(\hat{p}_{ij})}{\partial \theta_j} = -\frac{1}{\lambda_i \phi_i r_{ij} \theta_j^2}$$

Substituting this result into eq. A.3 and simplifying:

$$\frac{\partial \text{Var}(\hat{T}|C)}{\partial \theta_j} = Lc_j M_j - \theta_j^2 \sum_i \frac{r_{ij}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]$$

As indicated by the positive second derivative of eq. A.2 with respect to θ_j , setting the constrained first derivatives to zero and solving will produce a minimum variance as a function of θ_j . Solving first for the constant L :

$$L = \frac{\sum_i \frac{r_{ij}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]}{c_j M_j \theta_j^2}$$

Because this solution holds regardless of which cohort is involved:

$$L = \frac{\sum_i \frac{r_{ij}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]}{c_j M_j \theta_j^2} = \frac{\sum_i \frac{r_{ik}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]}{c_k M_k \theta_k^2}$$

where k represents a second cohort. This equation can be rearranged to express the optimal relationship between θ_j and θ_k that will minimize variance:

$$\frac{\theta_k}{\theta_j} = \frac{\sqrt{c_j M_j \sum_i \frac{r_{ik}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]}}{\sqrt{c_k M_k \sum_i \frac{r_{ij}}{\lambda_i \phi_i} [1 + G(\hat{N}_i)]}}$$

This completes the proof for eq. 2.3.

Finding the optimal relationship for sampling fractions for any two strata in a catch-sampling program follows the same procedure, only with a different cost function. The cost function $C = \sum_i a_i \phi_i^{b_i}$ is multiplied by a constant L and the products added and subtracted from eq. A.1:

$$(A.4) \quad \text{Var}(\hat{T}|C) = \sum_i \sum_j r_{ij}^2 [G(\hat{N}_i) + G(\hat{p}_{ij}) + G(\hat{\theta}_j^{-1})] \\ + G(\hat{N}_i)G(\hat{p}_{ij}) + G(\hat{N}_i)G(\hat{\theta}_j^{-1}) + G(\hat{p}_{ij})G(\hat{\theta}_j^{-1}) \\ + G(\hat{N}_i)G(\hat{p}_{ij})G(\hat{\theta}_j^{-1}) + 2 \sum_i \sum_j \sum_{k>j} r_{ij} r_{ik} G(\hat{N}_i) \\ + L \sum_i a_i \phi_i^{b_i} - LC$$

When eq. A.4 is differentiated with respect to ϕ_i :

$$(A.5) \quad \frac{\partial \text{Var}(\hat{T}|C)}{\partial \phi_i} = La_i b_i \phi_i^{b_i-1} + \sum_j r_{ij}^2 [1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i)G(\theta_j^{-1})] \frac{\partial G(\hat{p}_{ij})}{\partial \phi_i}$$

Note that r_{ij} , $G(\hat{N}_i)$, and $G(\theta_j^{-1})$ are not functions of ϕ_i . Differentiating $G(\hat{p}_{ij})$ with respect to ϕ_i :

$$\frac{\partial G(\hat{p}_{ij})}{\partial \phi_i} = -\frac{1}{\lambda_i r_{ij} \theta_j \phi_i^2}$$

Substituting this result into eq. A.5 and simplifying:

$$\frac{\partial \text{Var}(\hat{T}|C)}{\partial \phi_i} = La_i b_i \phi_i^{b_i-1} - \phi_i^{-2} \sum_j \frac{r_{ij}}{\lambda_i \theta_j} [1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i)G(\theta_j^{-1})]$$

As indicated by the positive second derivative of eq. A.4 with respect to ϕ_i (remember $b_i \geq 1$), setting the constrained first derivatives to zero and solving will produce a minimum variance as a function of the $\{\phi_i\}$. Solving first for the constant L :

$$L = \frac{\sum_j \frac{r_{ij}}{\lambda_i \theta_j} [1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i)G(\theta_j^{-1})]}{a_i b_i \phi_i^{b_i+1}}$$

Because this solution holds regardless of which stratum is involved:

$$\begin{aligned} & \frac{\sum_j \frac{r_{ij}}{\lambda_i \theta_j} [1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i)G(\theta_j^{-1})]}{a_i b_i \phi_i^{b_i+1}} \\ &= \frac{\sum_j \frac{r_{ij}}{\lambda_t \theta_j} [1 + G(\hat{N}_t) + G(\theta_j^{-1}) + G(\hat{N}_t)G(\theta_j^{-1})]}{a_t b_t \phi_t^{b_t+1}} \end{aligned}$$

where t represents a second stratum. This equation can be rearranged to express the optimal relationship between ϕ_i and ϕ_t that will minimize estimated variance:

$$\begin{aligned} \frac{\phi_t^{b_t+1}}{\phi_i^{b_i+1}} &= \\ & \frac{a_i b_i \lambda_i \sum_j r_{ij} \theta_j^{-1} [1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i)G(\theta_j^{-1})]}{a_t b_t \lambda_t \sum_j r_{ij} \theta_j^{-1} [1 + G(\hat{N}_i) + G(\theta_j^{-1}) + G(\hat{N}_i)G(\theta_j^{-1})]} \end{aligned}$$

This completes the proof for eq. 3.3.