Project Progress Report to the

Oregon Department of Fish and Wildlife
Marine Resources Program

Evaluation of Estimators for Rockfish Species Compositions.

David B. Sampson
Oregon State University
Coastal Oregon Marine Experiment Station

and

Yong-Woo Lee
University of Arkansas at Pine Bluff
Dept. of Aquaculture & Fisheries

Introduction

Rockfish (*Sebastes* spp.) are an important component of the US West Coast commercial fishery for groundfish. Although recent commercial landings of rockfish are greatly reduced compared to levels taken during the 1980s and early 1990s, rockfish continue to have a large impact on commercial and recreational fisheries along the West Coast because severe management restrictions were enacted to rebuild several rockfish species that were assessed as being severely depleted. Because most rockfish species have generally similar market characteristics, commercial fishermen generally do not sort rockfish to species and land them separately unless they are legally required to do so. In Oregon, for example, during the period 1985 to 1993 all landings of rockfish were reported in one of six so-called "market categories": Pacific ocean perch, widow rockfish, yellowtail rockfish, thornyheads, small rockfish, and large rockfish. In 1994 a separate category was established for black rockfish, and in 1995 three additional categories were established: for canary rockfish, longspine thornyhead, and shortspine thornyhead. While the species-specific categories (e.g., widow rockfish, yellowtail rockfish) are generally uncontaminated by other species, landings of these categories are often not entirely pure, and the species compositions of the general rockfish categories (e.g., small rockfish) are often quite variable (Crone 1995). In Oregon, agents from the Oregon Department of Fish and Wildlife (ODFW) routinely take samples (generally consisting of two 25-lb "clusters") from all rockfish market categories to derive estimates of the species composition of each category (Sampson et al. 1997). There are similar sampling programs in Washington (Tagart 1997) and California (Erwin et al. 1997).

Commercial landings of rockfish are weighed when the fish are off-loaded at the fish processing plants; the weights are recorded on official "fish tickets" that are assembled and tabulated by the state fisheries agencies, which regularly send landings statistics to the Pacific Fishery Information Network (PacFIN). The fish tickets provide nearly a complete census of the weight of commercial landings for all the rockfish market categories, but relatively few landings are sampled to measure the species composition (percentage by weight of each species) of the market categories. The state agencies provide PacFIN with estimates of rockfish species composition that the PacFIN system applies to the rockfish landings by market category to apportion the landings to individual rockfish species. In Oregon the ODFW provides PacFIN with species composition proportions for each rockfish market category for each combination of year, quarterly period, port, gear type, statistical area, and condition (dead or live) that produced any landings of that market category. With such a fine scale of "stratification" there are often strata cells for which no species composition samples were taken, in which case ODFW produces species composition percentages by "borrowing" estimates from adjacent ports or time periods for which species composition samples are available.

For those major rockfish species that are generally landed in relatively pure market categories (e.g., widow rockfish) the accuracy of the estimated landings may not be seriously affected by having limited species composition sampling. However, in the earliest years of the PacFIN system (1980 to 1984) all but two rockfish species were landed as either small or large rockfish and the accuracy of the landings estimates for all these other rockfish species is limited by the available species composition information. For minor rockfish species and ones that are not landed in relatively pure market categories, it seems likely that the PacFIN estimates of landings by species could be quite inaccurate.

An additional statistical issue, beyond the problem of limited sampling and the need to fill in data gaps, is how to estimate average species composition percentages when data from several samples are available. Because the species composition for a given species in a given market category is based on the weight rather than the numbers of fish, the species composition values are binomial-like but they are not binomial random variables. Standard methods for analyzing binomial data (e.g., logistic regression, McCullagh and Nelder, 1988) do not apply. Further, in samples from mixed rockfish market categories the distribution of percent composition values for any given rockfish species is often very over-dispersed, with the given rockfish species being absent from large proportions of the species composition samples.

The general goal of the project was to develop better procedures for estimating rockfish landings by species. The specific objectives for the project were: to evaluate procedures used by ODFW for generating rockfish species composition estimates for PacFIN, specifically the process of borrowing data to fill in un-sampled strata (year, quarter, port, gear, area, market category, condition); to evaluate alternative statistical models that could be used to describe the random sampling error in estimates of species composition; and to develop an alternative process for estimating rockfish species composition estimates that will provide more accurate estimates.

Most of the data manipulation and statistical analyses for the project were conducted during July and August of 2006 while Dr Lee was visiting the Hatfield Marine Science Center during his summer vacation. Dr Lee received no monetary compensation for the time and effort he contributed to the project, but the ODFW reimbursed his living expenses for the month he was in Newport and for his travel from Arkansas.

Materials and Methods

The system of rockfish market categories in Oregon has undergone considerable expansion in recent years with the addition of many new categories. For this project we focused on rockfish species composition data for the years 1997 to 1999. During this period there were nine rockfish market categories that had been part of the data system since 1995. Mr Mark Karnowski (ODFW, Marine Resources Program) provided us with two data files: one consisting of the fully developed species composition estimates that ODFW headquarters had uploaded to PacFIN and the other consisting of the rockfish species composition sample data that had been collected by ODFW port agents. Because there is limited documentation of the process used to derive the species composition estimates that are uploaded to PacFIN, we developed routines to recreate as closely as possible the uploaded species composition estimates from the species composition sample data and thereby verify that we understood how the species composition estimates are derived from sample data and the rules for borrowing species composition estimates when sample data are unavailable.

When we were satisfied that we had valid approaches for recreating the uploaded species composition estimates, we began developing a suite of statistical models using the Statistical Analysis System (SAS) to describe the observed species composition estimates for 1997 to 1999. If we could develop reasonably accurate statistical models, they could be used to provide estimates of species composition for strata for which no sample data were available, as an alternative to the current process of borrowing estimates from strata with samples.

Rules for borrowing species composition estimates

To compare the statistical models with the current process of borrowing estimates from strata with samples we took the data set with the observed species compositions and for each strata cell with actual estimates (as opposed to borrowed ones) we generated "borrowed" estimates using the following simplified rules:

(1) if there are no actual species composition estimates for landings of a particular market category, then comparable samples (with similar quarter, gear, and area fished characteristics) from the nearest major port are used;

(2) if after step (1) there are no applicable estimates available, then actual estimates are used from the preceding quarter for the nearest major port; and

(3) if after step (3) there are no applicable estimates available, then actual estimates are used from the succeeding quarter for the nearest major port.

To perform these reassignment tasks, Astoria, Newport, Coos Bay, and Brookings were defined as the major ports and the ports were matched at each step as shown in Table 1. We did not attempt to go beyond these three steps in developing borrowed estimates, but the system at ODFW headquarters has at least one additional rule that it applies in the event that the first three rules do not produce a set of species composition estimates.

Statistical models for species composition

The species composition sample data were analyzed using several statistical approaches to evaluate how well different statistical models could replicate the sample data. The statistical approaches differed in several aspects, one being how they weighted observations. In one of the approaches individual data observations were analyzed using the market category landing weights as a weighting variable, so that large landings contribute more than small landings to the overall average proportion for a species. In another approach the individual data observations were weighted by the number of species composition samples, and in a third approach the individual data observations were weighted equally (no weighting).

The statistical approaches also differed in how they dealt with observed species proportions that were zero or one. Statistical models that have a dependent variable that is a proportion require special precautions to make certain they do not produce nonsensical predicted values, because proportions cannot be negative and cannot exceed one. Most of the statistical approaches considered in this project involved breaking the species proportions into three separate components: one for the proportions that were zero, a second for the proportions that were one, and a third for the proportions not zero and not one. The separate components were then combined to produce an overall average species proportion. Symbolically, the average proportion of a species in a market category is

$$p_0 * 0 + p_1 * 1 + p_Q * Q ,$$

where p_0 is the probability that the species proportion is zero, p_1 is the probability that the species proportion is one, and p_Q is the probability that the species proportion Q is great than zero and less than one. Note that $p_Q = (1 - p_0 - p_1)$. Predictions for p_0 and p_1 and Q can be combined into an overall predicted proportion using

$$p_1 + (1 - p_0 - p_1) * Q .$$

For example, consider the following set of 10 observed species proportions: (0, 0, 0, 0.3, 0.3, 0.4, 0.5, 0.6, 0.7, 1), where the fraction of zeros is 0.3, and the fraction of ones is 0.1. The average proportion in the set is 0.38, which is equivalent to 0.1 + (1-0.3-0.1)*0.46667. The number 0.46667 is the average of the values in the set that are neither zero nor one.

All the statistical approaches in the study used the same following general linear model (GLM) structure:

$$Y = Unsp\_Sp * Species * (Port + Yr + Qtr + Gear + Area + Condition) ,$$

where Y is the dependent variable, *Unsp_Sp* represents the rockfish market category (the unspecified species), *Species* represents the rockfish species, *Yr* represents the year, *Qtr* represents the quarter of the year, *Gear* represents the fishing gear category, *Area* represents the Pacific Marine Fisheries Commission statistical area from which the catch was thought to have been taken, and *Condition* represents whether the fish were dead or live. All of the independent variables were treated as categorical variables. We did not examine models with interactions among the variables *Port*, *Yr*, *Qtr*, *Gear*, *Area,* and *Condition* because we wanted to maintain reasonably simple model structures.

Most of the statistical approaches involved three separate models: one with a dependent variable for the fraction of species proportions that were zero ($Y_1 = p_0$), a second with a dependent variable for the fraction of species proportions that were one ($Y_2 = p_1$), and a third for the logit of the fraction of species proportions (Q) that were neither zero nor one,

$$Y_3 = \ln[Q / (1 - Q)] , \ 0 < Q < 1.$$

The logit transformation is often used in the analysis of proportions to make them behave more like a normally distributed random variable (Snedecor and Cochran 1967), but the logit transformation is undefined for proportions of zero or one. In one set of statistical approaches the observed species proportions were analyzed without using a transformation (i.e., $Y_3 = Q$).

We used four approaches for estimating the proportions of zeros (p_0), the proportions of ones (p_1), and the proportions that were neither zero nor one (p_Q). In one approach we used separate binomial (logistic) regressions (without weighting) for each of p_0, p_1, and p_Q. In a second approach were used separate binomial regressions with weighting by the pounds landed of the sampled market category (*Lbs_Landed*). The third approach was like the first, and the fourth approach was like the second, but the third and fourth approaches used multinomial logistic regression. The multinomial logistic model accounts for the dependence of p_0, p_1, and p_Q, which must sum to one for each sample. The binomial logistic models for p_0, p_1, and p_Q assume these proportions are independent.

In total we developed and evaluated 16 different statistical models for estimating species composition (Table 2). The four sets of models for estimating the Q, the proportion not zero and not one, were crossed with the four sets of models for estimating p_0, p_1, and p_Q.

Results

For the period considered in this study (1997-1999) ODFW headquarters uploaded species composition estimates to PacFIN for a total of 1,531 individual strata (Market Category, Year, Quarter, Port, Gear, Area, and Condition) (Table 3). For almost half of these strata (771) the species composition estimates were borrowed from other sampled ports or quarters (indicated in Table 3 by the column labeled 0 for the "number of samples per strata"). The species composition estimates for another 331 strata were based on single species composition samples. These numbers suggest that the rockfish landings derived from the Oregon species composition estimates may be quite inaccurate, although a full evaluation of this would require further analyses of the magnitude of the landings associated with the sampled and un-sampled strata.

The data used to develop the suite of statistical models (Table 4) consisted only of data from actual samples, as opposed to borrowed estimates. We did not examine the rockfish landings data (the fish tickets) associated with the species composition samples to determine whether the ODFW port agents had collected the species composition data evenly across strata, but Sampson et al. (1997) found that the agents had sampled more or less in proportion to the rockfish landings during 1991 and 1992. In that period, however, there were only six rockfish market categories. With additional market categories there are additional strata that require species composition estimates, and it becomes increasingly difficult to obtain adequate samples.

For the period examined in this study there were a total of 39 different rockfish species observed in the species composition samples (Table 5). For developing the statistical models we limited our analysis to 30 species, choosing those for which there were the largest number of samples during the period 1987 to 2005.

Data uploaded to PacFIN versus sample-level species composition data

Based on a previous investigation of the Oregon groundfish data processing system (Sampson et al. 1997) we assumed that sample estimates of species proportions for a given market category were combined into strata-level estimates based on the following equation,

$$\mathrm{Av}(\, p_i \,) = \mathrm{sum}(\, Lbs\_Landed_j * p_{i,j} \,) \,/\, \mathrm{sum}(\, Lbs\_Landed_j \,),\ \text{for } j = 1, 2, \ldots\ n_j,$$

where p_i denotes the proportion of species i in the market category and stratum of interest, Lbs_Landed_j is the pounds of the market category in sample j from the stratum, $p_{i,j}$ is the proportion by weight of species i in sample j from the stratum, and n_j, is the number of samples from the market category in the stratum. For each of the strata-by-species combinations in the species composition data file (SPCOMP) we calculated the species composition proportion and compared it with the corresponding value from the data file of values uploaded to PacFIN (UPLOAD).

Our strata-level estimates of species proportions calculated from the SPCOMP data should agree with the strata-level estimates of species proportions in the UPLOAD file, but the corresponding values for any given stratum would differ if the two files were based on different sample data. One-to-one matches of the strata data from the two files indicated 14 strata (out of 1,531) in the UPLOAD file that were not in the SPCOMP file, and six strata (out of 752) in the

SPCOMP file that were not in the UPLOAD file (Table 6). Examination of the data records indicated that the differences in the two data sets were all due to discrepancies in the Area field. We speculate that the data processing system in ODFW headquarters used additional information (e.g., from trawl logbooks) to modify the area of capture in one of the files, but the system did not carry the changes to the other file.

We also compared our strata-level calculations of species proportions derived from the SPCOMP data with the corresponding estimates in the UPLOAD file and found five instances (out of 2,716 comparisons) where the absolute value of the difference was greater than 0.0005, (Table 7). In four cases the SPCOMP data and the UPLOAD data were based on differing numbers of samples. One-to-one matches of the strata data from the two files indicated one additional instance of differences in the numbers of samples (but a minor difference in the species proportions).

To evaluate our simplified algorithm for borrowing species proportions we generated replacement values for the 2,540 borrowed species proportion estimates in the UPLOAD file. We applied our algorithm to the UPLOAD strata with missing data and borrowed estimates derived from UPLOAD strata that had sample data. Our algorithm was able to find replacement values for 1,952 species proportions, which we then compared with the corresponding species proportions in the UPLOAD file. Of these comparisons the absolute value of the difference was less than or equal to 0.0005 in 1,927 cases. In the other 25 cases the species proportions in the UPLOAD file appeared to have been borrowed using slightly different rules (Table 8). We did not have time to fully investigate the cause of the differences.

In general our comparisons of the UPLOAD data with corresponding data and derived estimates from the SPCOMP file indicate relatively few and minor discrepancies, thus validating our assertion that our algorithms correctly mimic the data processing routines used in ODFW headquarters to produce the species composition estimates provided to PacFIN.

Composite GLM estimates of species composition

The distributions of observed species proportions are quite variable from market category to market category (Table 9). Some market categories, such as longspine thornyhead, are almost entirely pure. Others, such as unspecified rockfish (URCK) have a wide diversity of species, with some species absent from some samples and dominating other samples (e.g., CHNA in URCK).

Given that we had 16 different statistical approaches to evaluate, it seemed important to develop simple criteria for identifying and eliminating approaches that produced infeasible results, such as proportions that were less than zero or greater than one. We produced simple tabulations and range calculations of the estimated values for p_0, p_1, p_Q, and Q for strata-by-species combinations that had observed species proportions that were between zero and one. We found that approaches (4), (8), (12), and (16), which did not apply the logit transformation to Q, sometimes produced estimated proportions less than zero or greater than one. Also, the approaches that used independent binomial models to derive estimates for p_0, p_1, and p_Q produced sets of estimates whose sum differed markedly from one for some strata, indicating that at least one of the estimates in the set was drastically inaccurate for those strata. Thus approaches (1) through (8) did not appear to be reasonable candidates for producing accurate estimates of species proportions.

To evaluate statistical approaches (9), (10), (11), (13), (14), and (15) we calculated logit-transformed residuals of the species proportions for the 5,228 strata-by-species combinations that had observed species proportions that were between zero and one,

$$\mathrm{L\_residual} = \ln[\ P\ /\ (\ 1 - P\ )\ ] - \ln[\ P'\ /\ (\ 1 - P'\ )\ ]\ ,$$

where P denotes the observed species proportion for a given stratum and P' denotes the predicted species proportion. The statistical approach that produced the most accurate estimates, as measured by the root mean squared error, equal to av(*L_residual*) ^{2} + StDev(*L_residual*) ^{2}, was approach (11), which used an unweighted GLM model for logit(Q) and an unweighted multinomial logistic model for p_Q and p_1 to estimate the species proportions. Approach (10) produced slightly less accurate estimates than approach (11). The other approaches produced considerably less accurate estimates than (11).

GLM estimates versus borrowed estimates

The species proportions from approach (11) were compared to estimates derived from the simplified borrowing algorithm (described previously) to evaluate which approach produced more accurate estimates. We generated replacement values for the 5,228 strata-by-species estimates that had observed species proportions between zero and one. We applied our algorithm to each strata-by-species combination and borrowed proportions as appropriate from other strata in the file. Our algorithm was able to find replacement values for 4,200 species proportions, which we then compared with the corresponding observed species proportions and with those estimated by approach (11). The estimated proportions from approach (11) were much more closely related to the original proportions than were the proportions derived from the borrowing algorithm (Fig. 1) and the root mean squared error from approach (11) was almost 1/10 the magnitude of the root mean squared error from the borrowed estimates. Even the statistical approach, however, produced some highly inaccurate estimates, as indicated in the figure.

Discussion

The evaluations of the different approaches were based on their predictions for those observed strata-by-species proportions that were between zero and one. A more complete evaluation would also consider the accuracy of the approaches for predicting observed species proportions at the limits, zero or one. The results of our analysis nonetheless strongly suggest that rockfish species composition estimates would be much more accurate if they were derived from a statistical approach such as described in this report. The current approach of borrowing estimates from other strata has no valid theoretical basis and appears to produce results that often are highly inaccurate.

The current system relies heavily on borrowed species proportions, with almost half of the strata left un-sampled during the study period. Short of hiring additional port agents to collect additional species composition samples, it may be possible to adjust the port sampling system to make operations more efficient and make better use of existing resources. For example, an analysis could be conducted that would determine whether sampling rates are even across strata so that resources could be switched from over-sampled strata to under-sampled strata. Also, if the port sampling system were more closely monitored, it might be possible to direct port agents to collect samples from particular strata so that there would be fewer instances where borrowing was necessary. Finally, it is probably unrealistic to expect the data processing staff at ODFW headquarters to begin using sophisticated statistical models to develop estimates of species compositions. We developed and explored a simplified algorithm for estimating species proportions based on simple factor-level averages, but have not yet fully evaluated the performance of the algorithm and do not report on it here. If the algorithm can be shown to be reasonably accurate, and more accurate than the current system, then it should be relatively straight-forward for the data processing staff to implement the algorithm using existing software tools, such as SQL.

References

Crone, P.R. 1995. Sampling design and statistical considerations for the commercial groundfish fishery of Oregon. Can. J. Fish. Aquat. Sci. 52:716-732.

Erwin, B.A., Thomas, D.H., Kobylinski, G.J., and Bence, J.R. 1997. Groundfish data collection in California. *In*: Sampson, D.B. and Crone, P.R. (ed.) Commercial fisheries data collection procedures for the U.S. Pacific coast groundfish. NOAA Tech. Memo. NMFS-NWFSC-31, p.105-140.

McCullagh, P. and Nelder, J.A. 1983. *Generalized Linear Models*. Chapman and Hall, London.

Sampson, D.B., Crone, P.R., and Saelens, M.R. 1997. Groundfish data collection in Oregon. *In*: Sampson, D.B. and Crone, P.R. (ed.) Commercial fisheries data collection procedures for the U.S. Pacific coast groundfish. NOAA Tech. Memo. NMFS-NWFSC-31, p.53-104.

Snedecor, G.W. and Cochran, W.G. 1967. *Statistical Methods*. Iowa State University Press, Ames, Iowa.

Tagart, J.V. 1997. Groundfish data collection in Washington. *In*: Sampson, D.B. and Crone, P.R. (ed.) Commercial fisheries data collection procedures for the U.S. Pacific coast groundfish. NOAA Tech. Memo. NMFS-NWFSC-31, p.11-52.

Table 1. Port codes and alternatives ports for borrowing species composition estimates.

			Alternative Ports	
ODFW	PacFIN	Name	Step (1)	Steps (2) and (3)
02	AST	Astoria	NEW	AST
10	TLL	Tillamook/Garibaldi	AST	AST
16	PCC	Pacific City	NEW	NEW
24	NEW	Newport	AST	NEW
32	WIN	Winchester Bay	COS	COS
34	COS	Coos Bay/Charleston	BRK	COS
36	BDN	Bandon	COS	COS
38	ORF	Port Orford	BRK	BRK
40	GLD	Gold Beach	BRK	BRK
42	BRK	Brookings	COS	BRK

Table 2. Statistical models developed for estimating species composition proportions.

Model structure shared by all models:

$$Y = Unsp\_Sp * Species * (Port + Yr + Qtr + Gear + Area + Condition)$$

	for Q, $0 < Q < 1$		for (p_0, p_1, p_Q)	
Model	Y variable	Weighting	Model type	Weighting
(1)	logit(Q)	Lbs_Landed	binomial	none
(2)	logit(Q)	N_Sample	binomial	none
(3)	logit(Q)	none	binomial	none
(4)	Q	none	binomial	none
(5)	logit(Q)	Lbs_Landed	binomial	Lbs_Landed
(6)	logit(Q)	N_Sample	binomial	Lbs_Landed
(7)	logit(Q)	none	binomial	Lbs_Landed
(8)	Q	none	binomial	Lbs_Landed
(9)	logit(Q)	Lbs_Landed	multinomial	none
(10)	logit(Q)	N_Sample	multinomial	none
(11)	logit(Q)	none	multinomial	none
(12)	Q	none	multinomial	none
(13)	logit(Q)	Lbs_Landed	multinomial	Lbs_Landed
(14)	logit(Q)	N_Sample	multinomial	Lbs_Landed
(15)	logit(Q)	none	multinomial	Lbs_Landed
(16)	Q	none	multinomial	Lbs_Landed

Table 3. Number of strata with species composition data uploaded to PacFIN. A stratum is a combination of Market Category, Year, Quarter, Port, Gear, Area, and Condition.

		- - - - - - - - - - - - - - Number of samples per strata - - - - - - - - - - - - - -										
Factor variable & level		0	1	2	3	4	5	6	7	8	9+	Total
Distribution by Market Category												
BLK1	Nominal black rf	18	14	5	3	2	1	2			1	46
CNR1	Nominal canary rf	110	40	25	13	3	7	2	2		5	207
LSP1	Nom. longspine thyhd	67	30	17	7	10	4	2	1			138
POP1	General shelf/slope rf	94	32	9	6	9	8	6	3	4	7	178
POP2	Nom. Pac. oc. perch	37	15	10	3	4	5	4				78
SSP1	Nom. shortspine thyhd	87	29	15	15	3	4	3	1			157
URCK	Unspecified rf	136	56	26	14	5	6	8	4	5	17	277
WDW1	Nominal widow rf	107	51	19	12	7	5	2	3	3	1	210
YTR1	Nominal yellowtail rf	115	63	25	11	9	5	7		3	2	240
Distribution by Year												
1997		322	114	43	29	13	17	17	7	8	14	584
1998		226	117	56	26	19	15	10	5	6	5	485
1999		223	99	52	29	20	13	9	2	1	14	462
Total		771	330	151	84	52	45	36	14	15	33	1531
Distribution by Quarter												
1		178	70	24	18	13	8	8	2	1		322
2		164	108	44	29	21	12	12	7	9	19	425
3		199	86	50	21	13	19	14	5	4	14	425
4		230	66	33	16	5	6	2		1		359
Distribution by Port												
AST	Astoria	162	74	43	27	14	12	8	4	2	7	353
BRK	Brookings	108	48	26	6	5	3			1		197
COS	Coos Bay	128	73	31	20	12	9	9	3	6	6	297
DPO	Depoe Bay	15										15
FLR	Florence	29										29
GLD	Gold Beach	11	6	4			1	1			2	25
NEW	Newport	210	89	35	20	13	14	10	5	3	6	405
ORF	Port Orford	5	34	11	11	6	6	8	2	3	12	98
PCC	Pacific City	23	1	1		1						26
TLL	Tillamook	61	5			1						67
WIN	Winchester Bay	19										19

Table 3. Number of strata with species composition data uploaded to PacFIN (continued).

		- - - - - - - - - - - - - - Number of samples per strata - - - - - - - - - - - - - -										
Factor variable & level		0	1	2	3	4	5	6	7	8	9+	Total
Distribution by Gear												
DST	Double shrimp trawl	37	27	2	3	2	1	2		2		76
FPT	Fish pot	1	2									3
GFT	Groundfish trawl	543	192	117	61	41	35	25	11	9	17	1051
LGL	Longline	30	40	6	6	6	3	5	2	3	1	102
MDT	Mid-water trawl	38	24	6	5	2	1		1			77
OHL	Other hook & line	122	45	20	9	1	5	4		1	15	222
Distribution by PMFC Area												
1B			3									3
1C		49	23	6	1	1						80
2A		100	80	45	16	12	10	9	2	4	14	292
2B		176	68	34	21	11	9	9	3	6	6	343
2E		106	26	2	1	1						136
2F		144	48	22	18	12	14	10	5	3	6	282
3A		153	46	27	24	9	9	6	4	2	5	285
3B		43	36	15	3	6	3	2			2	110
Distribution by Condition												
D	Dead	762	321	148	81	51	44	35	14	14	29	1499
L	Alive	9	9	3	3	1	1	1		1	4	32
Total		771	330	151	84	52	45	36	14	15	33	1531

Table 4. Summary of the species composition data examined in the statistical models.

Factor	1997	1998	1999	Total	Factor	1997	1998	1999	Total
Numbers of strata by Market Category and Yr					*Numbers of samples by Market Category and Yr*				
BLK1	7	11	9	27	BLK1	14	25	27	66
CNR1	36	32	30	98	CNR1	111	63	88	262
LSP1	27	23	21	71	LSP1	75	44	45	164
POP1	28	31	24	83	POP1	132	112	79	323
POP2	17	13	11	41	POP2	43	30	36	109
SSP1	21	25	22	68	SSP1	61	55	43	159
URCK	40	48	52	140	URCK	147	165	204	516
WDW1	37	34	29	100	WDW1	100	76	65	241
YTR1	50	37	37	124	YTR1	133	85	79	297
Total	263	254	235	752	Total	816	655	666	2137
Numbers of strata by Port and Yr					*Numbers of samples by Port and Yr*				
AST	71	63	58	192	AST	195	165	153	513
BRK	14	31	44	89	BRK	19	60	82	161
COS	71	66	31	168	COS	245	163	77	485
GLD			14	14	GLD			44	44
NEW	84	55	50	189	NEW	256	143	137	536
ORF	23	31	38	92	ORF	101	110	173	384
PCC		2		2	PCC		5		5
TLL		6		6	TLL		9		9
Numbers of strata by Qtr and Yr					*Numbers of samples by Qtr and Yr*				
1	59	43	42	144	1	138	91	104	333
2	87	83	86	256	2	323	245	269	837
3	75	73	76	224	3	276	200	242	718
4	42	55	31	128	4	79	119	51	249

Table 4. Summary of the species composition data examined in the statistical models (cont.).

Factor	1997	1998	1999	Total
Numbers of strata by Gear and Qtr				
DST	15	9	13	37
FPT	1	1		2
GFT	175	184	147	506
LGL	21	19	31	71
MDT	16	9	12	37
OHL	35	32	32	99
Numbers of strata by Area and Yr				
1B		3		3
1C	9	12	10	31
2A	51	65	76	192
2B	71	55	41	167
2E	13	4	3	20
2F	47	45	46	138
3A	45	50	37	132
3B	26	17	20	63
3C	1	3	2	6
Numbers of strata by Condition and Yr				
D	263	246	221	730
L		8	14	22
Total	263	254	235	752

Factor	1997	1998	1999	Total
Numbers of samples by Gear and Yr				
DST	42	19	18	79
FPT	1	1		2
GFT	557	480	395	1432
LGL	66	55	65	186
MDT	32	15	22	69
OHL	118	85	166	369
Numbers of samples by Area and Yr				
1B		3		3
1C	10	13	19	42
2A	148	182	261	591
2B	236	153	97	486
2E	18	4	3	25
2F	209	125	132	466
3A	132	135	113	380
3B	61	36	38	135
3C	2	4	3	9
Numbers of samples by Condition and Yr				
D	816	618	599	2033
L		37	67	104
Total	816	655	666	2137

Table 5. Species observed in rockfish species composition samples, 1997-99.

Numbers of samples (excluding borrows) for each species by Year						Numbers of strata (incl. borrows) for each species by Year					Used in GLMs ?
Species	Name	1997	1998	1999	Total	Species	1997	1998	1999	Total	
ARRA	Aurora	124	147	116	387	ARRA	57	55	61	173	1
BANK	Bank	68	67	48	183	BANK	15	12	24	51	1
BCAC	Bocaccio	166	119	98	383	BCAC	76	50	64	190	1
BLCK	Black	20	27	27	74	BLCK	18	17	12	47	1
BLGL	Blackgill	54	40	41	135	BLGL	17	12	19	48	1
BLUR	Blue	47	46	36	129	BLUR	5	11	8	24	1
CHNA	China	44	75	127	246	CHNA	11	21	24	56	1
CLPR	Chilipepper	49	70	46	165	CLPR	15	24	20	59	1
CNRY	Canary	232	148	149	529	CNRY	128	88	80	296	1
COPP	Copper	42	81	117	240	COPP	10	16	18	44	1
CWCD	Cowcod	39	58	31	128	CWCD	21	13	15	49	0
DBRK	Darkblotched	247	210	163	620	DBRK	124	94	107	325	1
GRAS	Grass			34	34	GRAS			4	4	0
GPRK	Greenspotted	115	111	58	284	GSPT	39	35	24	98	1
GSRK	Greenstriped	197	185	98	480	GSRK	79	85	51	215	1
LSPN	Longspine	131	89	73	293	LSPN	75	76	58	209	1
ORCK	Other Rockfish			10	10	ORCK			1	1	0
PGMY	Pygmy		46		46	PGMY		14		14	0
POP	POP	241	201	173	615	POP	111	99	95	305	1
QLBK	Quillback	52	75	126	253	QLBK	16	16	23	55	1
RDBD	Redbanded	196	185	142	523	RDBD	82	87	71	240	1
REDS	Redstripe	183	180	119	482	REDS	72	73	61	206	1
REYE	Rougheye	190	156	144	490	REYE	72	76	84	232	1
ROSY	Rosy	17	5		22	ROSY	5	2		7	0
RSTN	Rosethorn	175	150	141	466	RSTN	61	57	45	163	1

Table 5. Species observed in rockfish species composition samples, 1997-99 (continued).

Numbers of samples (excluding borrows) for each species by Year						Numbers of strata (incl. borrows) for each species by Year					Used in GLMs ?
Species	Name	1997	1998	1999	Total	Species	1997	1998	1999	Total	
SBLY	Shortbelly	94	48	12	154	SBLY	21	11	7	39	0
SHRP	Sharpchin	244	179	164	587	SHRP	107	69	80	256	1
SLGR	Silvergray	130	116	85	331	SLGR	49	46	45	140	1
SNOS	Splitnose	226	192	149	567	SNOS	98	80	95	273	1
SPKL	Speckled		4		4	SPKL		1		1	0
SRKR	Shortraker	118	73	52	243	SRKR	30	22	23	75	1
SSPN	Shortspine	214	150	134	498	SSPN	131	92	100	323	1
STRK	Stripetail	3	50	36	89	STRK	2	12	20	34	0
TIGR	Tiger	113	72	105	290	TGR1	38	14	18	70	0
VRML	Vermilion	44	67	117	228	VRM1	11	11	18	40	1
WDOW	Widow	224	177	143	544	WDOW	130	113	103	346	1
YEYE	Yelloweye	222	194	221	637	YEY1	104	78	69	251	1
YMTH	Yellowmouth	147	105	78	330	YMTH	54	31	40	125	1
YTRK	Yellowtail	189	133	150	472	YTRK	121	102	105	328	1
Total		4597	4031	3563	12191	Total	2005	1715	1692	5412	30

Table 6. Strata not in both the UPLOAD and SPCOMP files.

Port	Yr	Qtr	Gear	Area	Condition	Unsp_sp	Number Samples	Notes
Strata in UPLOAD that were missing from SPCOMP								
AST	1998	1	GFT	3B	D	CNR1	1	SPCOMP has 2 samples like this from 3A and 1 from 3C
AST	1998	3	GFT	3B	D	CNR1	1	SPCOMP has 3 samples like this from 3A and 1 from 3C
AST	1998	3	MDT	2E	D	WDW1	1	SPCOMP has 1 sample like this from 3A and 1 from 3B
AST	1999	1	GFT	3B	D	CNR1	2	SPCOMP has 2 samples like this from 3C
AST	1999	3	GFT	3B	D	CNR1	1	SPCOMP has 2 samples like this from 3A and 1 from 3C
COS	1998	3	GFT	2E	D	CNR1	1	SPCOMP has 5 samples like this from 2B
NEW	1998	2	GFT	2E	D	POP1	1	SPCOMP has 7 samples like this from 2F and 1 from 2B
NEW	1998	2	GFT	2E	D	SSP1	1	SPCOMP has 3 samples like this from 2F
NEW	1999	2	DST	2E	D	WDW1	1	SPCOMP has 1 sample like this from 2F
NEW	1999	2	DST	2E	D	YTR1	1	SPCOMP has 1 sample like this from 2F
NEW	1999	2	GFT	2E	D	SSP1	1	SPCOMP has 5 samples like this from 2F
NEW	1999	2	MDT	2E	D	WDW1	1	SPCOMP has 1 sample like this from 2F
ORF	1998	4	LGL	2E	L	URCK	1	SPCOMP has 6 samples like this from 2A
PCC	1998	2	OHL	2E	D	BLK1	2	SPCOMP has 4 samples like this from 2F
Strata in SPCOMP that were missing from UPLOAD								
AST	1997	2	GFT	3C	D	CNR1	2	UPLOAD has 3 samples like this from 3B
AST	1998	1	GFT	3C	D	CNR1	1	UPLOAD has 2 samples like this from 3A and 1 from 3B
AST	1998	2	GFT	3C	D	CNR1	2	UPLOAD has 2 samples like this from 3A and 3 from 3B
AST	1998	3	GFT	3C	D	CNR1	1	UPLOAD has 2 samples like this from 3A and 1 from 3B
AST	1999	1	GFT	3C	D	CNR1	2	UPLOAD has 2 samples like this from 3B
AST	1999	3	GFT	3C	D	CNR1	1	UPLOAD has 2 samples like this from 3A and 1 from 3B

Table 7. Discrepancies in strata-level estimates of species proportions, UPLOAD vs SPCOMP.

								Number of samples		Est. proportion	
Port	Yr	Qtr	Gear	Area	Condition	Unsp_sp	Species	UPLOAD	SPCOMP	UPLOAD	SPCOMP
AST	1997	2	GFT	3B	D	CNR1	CNRY	3	1	0.9983	0.9951
AST	1997	2	GFT	3B	D	CNR1	REDS	3	1	0.0017	0.0049
AST	1998	2	GFT	3B	D	CNR1	CNRY	3	1	0.5778	0.0182
AST	1998	2	GFT	3B	D	CNR1	YEY1	3	1	0.4222	0.9818
BRK	1999	3	DST	2B	D	URCK	DBRK	1	1	0.5854	0.5860

Table 8. Discrepancies in borrowed estimates of species proportions, UPLOAD vs simplified borrowing rules.

								Est. proportion				
Port	Yr	Qtr	Gear	Area	Cond.	Unsp_sp	Species	Original	New	New stratum key	Original stratum key used in UPLOAD	Notes
AST	1997	3	GFT	2E	D	POP1	ARRA	0.06140	0.01980	AST19974GFT2EDPOP1	NEW19971GFT2EDPOP1	jumped port and 2 qtrs
AST	1997	3	GFT	2E	D	POP1	DBRK	0.68480	0.58950	AST19974GFT2EDPOP1	NEW19971GFT2EDPOP1	jumped port and 2 qtrs
AST	1997	3	GFT	2E	D	POP1	POP	0.05240	0.07050	AST19974GFT2EDPOP1	NEW19971GFT2EDPOP1	jumped port and 2 qtrs
AST	1997	3	GFT	2E	D	POP1	REYE	0.07500	0.00560	AST19974GFT2EDPOP1	NEW19971GFT2EDPOP1	jumped port and 2 qtrs
AST	1997	3	GFT	2E	D	POP1	SHRP	0.00190	0.04010	AST19974GFT2EDPOP1	NEW19971GFT2EDPOP1	jumped port and 2 qtrs
AST	1997	3	GFT	2E	D	POP1	SNOS	0.00600	0.25190	AST19974GFT2EDPOP1	NEW19971GFT2EDPOP1	jumped port and 2 qtrs
BRK	1997	2	GFT	1C	D	YTR1	YTRK	0.99100	1.00000	BRK19973GFT1CDYTR1	COS19971GFT1CDYTR1	jumped port and 1 qtr
COS	1997	1	GFT	2A	D	POP1	DBRK	0.06400	0.03500	COS19972GFT2ADPOP1	BRK19972GFT2ADPOP1	jumped port and 1 qtr
COS	1997	1	GFT	2A	D	POP1	GSPT	0.01380	0.00940	COS19972GFT2ADPOP1	BRK19972GFT2ADPOP1	jumped port and 1 qtr
COS	1997	1	GFT	2A	D	POP1	GSRK	0.43000	0.09160	COS19972GFT2ADPOP1	BRK19972GFT2ADPOP1	jumped port and 1 qtr
COS	1997	1	GFT	2A	D	POP1	RDBD	0.00090	0.00150	COS19972GFT2ADPOP1	BRK19972GFT2ADPOP1	jumped port and 1 qtr
COS	1997	1	GFT	2A	D	POP1	RSTN	0.03570	0.05580	COS19972GFT2ADPOP1	BRK19972GFT2ADPOP1	jumped port and 1 qtr
COS	1997	1	GFT	2A	D	POP1	SHRP	0.42390	0.28660	COS19972GFT2ADPOP1	BRK19972GFT2ADPOP1	jumped port and 1 qtr
COS	1997	1	GFT	2A	D	POP1	SNOS	0.03150	0.09370	COS19972GFT2ADPOP1	BRK19972GFT2ADPOP1	jumped port and 1 qtr
COS	1998	1	GFT	2A	D	YTR1	YTRK	0.96440	1.00000	COS19982GFT2ADYTR1	BRK19982GFT2ADYTR1	jumped port and 1 qtr
COS	1999	4	GFT	2A	D	POP1	DBRK	0.01040	0.63200	COS19993GFT2ADPOP1	BRK19993GFT2ADPOP1	jumped port and 1 qtr
COS	1999	4	GFT	2A	D	POP1	POP	0.00100	0.31550	COS19993GFT2ADPOP1	BRK19993GFT2ADPOP1	jumped port and 1 qtr
COS	1999	4	GFT	2A	D	POP1	RDBD	0.00250	0.00580	COS19993GFT2ADPOP1	BRK19993GFT2ADPOP1	jumped port and 1 qtr
COS	1999	4	GFT	2A	D	POP1	SHRP	0.00200	0.01070	COS19993GFT2ADPOP1	BRK19993GFT2ADPOP1	jumped port and 1 qtr
COS	1999	4	GFT	2A	D	POP1	SNOS	0.01980	0.03600	COS19993GFT2ADPOP1	BRK19993GFT2ADPOP1	jumped port and 1 qtr
NEW	1997	2	GFT	2E	D	LSP1	SSPN	1.00000	0.00400	NEW19971GFT2EDLSP1	AST19971GFT2EDLSP1	jumped port and 1 qtr
NEW	1997	2	GFT	3A	D	CNR1	CNRY	0.98650	0.98500	NEW19973GFT3ADCNR1	AST19971GFT3ADCNR1	jumped port and 1 qtr
NEW	1999	1	GFT	3A	D	POP1	RDBD	0.01430	0.02800	NEW19992GFT3ADPOP1	AST19992GFT3ADPOP1	jumped port and 1 qtr
NEW	1999	1	GFT	3A	D	POP1	REYE	0.01530	0.97200	NEW19992GFT3ADPOP1	AST19992GFT3ADPOP1	jumped port and 1 qtr
NEW	1999	4	GFT	3A	D	YTR1	YTRK	1.00000	0.98300	NEW19993GFT3ADYTR1	AST19993GFT3ADYTR1	jumped port and 1 qtr

Table 9. Distributions of species proportions in rockfish market categories, 1997 to 1999.

		Number of samples with given species proportion												
Market			- - - - - - - - - - - - - Species proportion mid-point - - - - - - - - - - - - -											
Category	Species	0	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	1	Total
BLK1	BLCK	0	0	0	1	1	0	1	2	3	5	15	38	66
BLK1	BLGL	65	0	1	0	0	0	0	0	0	0	0	0	66
BLK1	BLUR	40	14	4	3	2	1	0	1	1	0	0	0	66
BLK1	QLBK	65	1	0	0	0	0	0	0	0	0	0	0	66
BLK1	YTRK	65	1	0	0	0	0	0	0	0	0	0	0	66
BLK1 Total		1885	16	5	4	3	1	1	3	4	5	15	38	1980
CNR1	BCAC	261	1	0	0	0	0	0	0	0	0	0	0	262
CNR1	BLGL	261	1	0	0	0	0	0	0	0	0	0	0	262
CNR1	CLPR	261	1	0	0	0	0	0	0	0	0	0	0	262
CNR1	CNRY	0	1	0	0	0	0	0	0	0	0	38	223	262
CNR1	DBRK	259	3	0	0	0	0	0	0	0	0	0	0	262
CNR1	GSPT	252	10	0	0	0	0	0	0	0	0	0	0	262
CNR1	GSRK	260	2	0	0	0	0	0	0	0	0	0	0	262
CNR1	POP	260	2	0	0	0	0	0	0	0	0	0	0	262
CNR1	RDBD	260	2	0	0	0	0	0	0	0	0	0	0	262
CNR1	REDS	257	5	0	0	0	0	0	0	0	0	0	0	262
CNR1	REYE	261	1	0	0	0	0	0	0	0	0	0	0	262
CNR1	SHRP	254	8	0	0	0	0	0	0	0	0	0	0	262
CNR1	SNOS	261	1	0	0	0	0	0	0	0	0	0	0	262
CNR1	WDOW	261	1	0	0	0	0	0	0	0	0	0	0	262
CNR1	YEY1	256	5	0	0	0	0	0	0	0	0	1	0	262
CNR1	YMTH	261	1	0	0	0	0	0	0	0	0	0	0	262
CNR1	YTRK	260	2	0	0	0	0	0	0	0	0	0	0	262
CNR1 Total		7551	47	0	0	0	0	0	0	0	0	39	223	7860
LSP1	LSPN	2	0	0	0	0	0	1	1	7	18	62	73	164
LSP1	SSPN	73	62	18	7	1	1	0	0	0	0	0	2	164
LSP1 Total		4667	62	18	7	1	1	1	1	7	18	62	75	4920
POP1	ARRA	240	67	9	2	1	0	0	0	1	1	2	0	323
POP1	BANK	293	25	5	0	0	0	0	0	0	0	0	0	323
POP1	BCAC	278	37	6	0	2	0	0	0	0	0	0	0	323
POP1	BLGL	315	7	0	1	0	0	0	0	0	0	0	0	323
POP1	CLPR	284	34	3	0	1	0	0	1	0	0	0	0	323
POP1	CNRY	290	33	0	0	0	0	0	0	0	0	0	0	323
POP1	COPP	322	0	0	0	0	0	1	0	0	0	0	0	323
POP1	DBRK	73	72	33	27	20	22	19	12	16	16	10	3	323
POP1	GSPT	286	34	1	1	1	0	0	0	0	0	0	0	323
POP1	GSRK	120	81	31	20	24	11	11	10	8	7	0	0	323
POP1	LSPN	322	1	0	0	0	0	0	0	0	0	0	0	323
POP1	POP	154	150	15	3	1	0	0	0	0	0	0	0	323
POP1	QLBK	321	1	1	0	0	0	0	0	0	0	0	0	323
POP1	RDBD	174	140	5	2	0	1	0	0	0	1	0	0	323
POP1	REDS	154	78	30	11	10	11	7	6	4	3	6	3	323
POP1	REYE	229	59	16	10	0	1	1	1	1	0	3	2	323
POP1	RSTN	157	145	11	1	3	1	0	0	2	1	1	1	323
POP1	SHRP	104	113	38	20	16	13	10	3	4	2	0	0	323

Table 9. Distributions of species proportions in rockfish market categories, 1997 to 1999 (cont.).

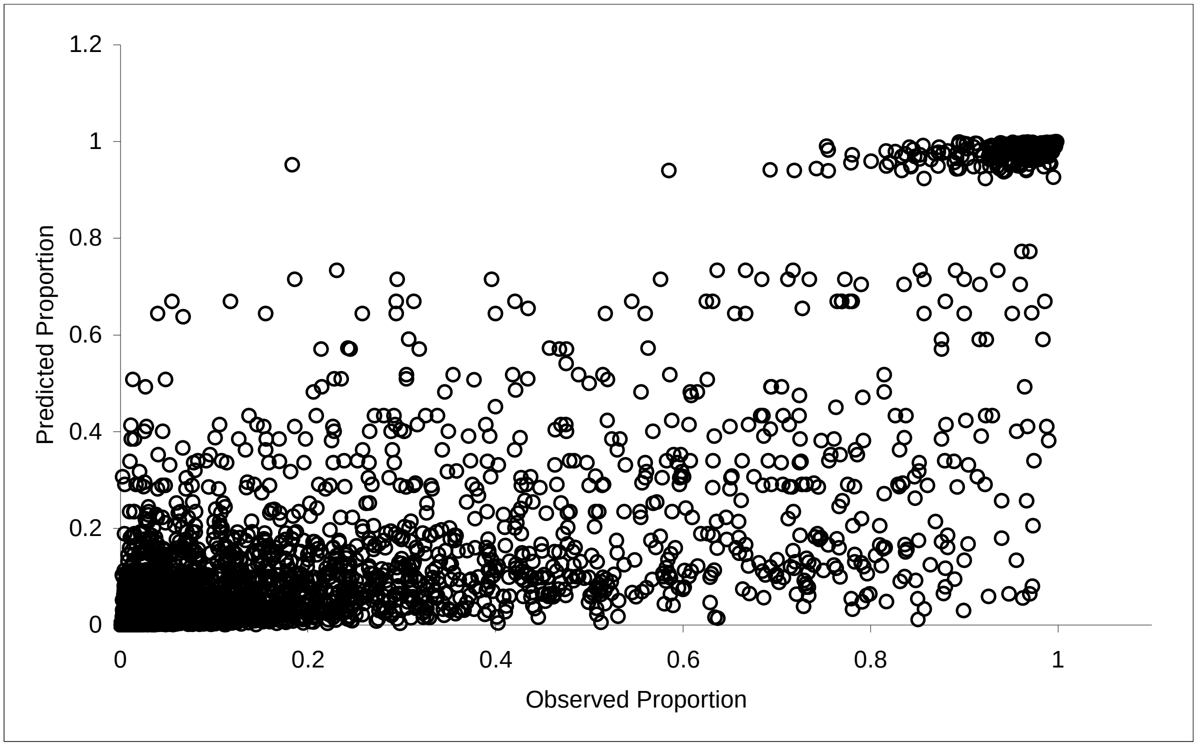
		Number of samples with given species proportion												
			- - - - - - - - - - - - - Species proportion mid-point - - - - - - - - - - - - -											
Market Category	Species	0	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	1	Total
POP1	SLGR	284	28	6	4	1	0	0	0	0	0	0	0	323
POP1	SNOS	144	113	23	14	11	6	5	1	3	1	2	0	323
POP1	SRKR	311	4	3	3	1	0	0	0	0	0	0	1	323
POP1	SSPN	291	32	0	0	0	0	0	0	0	0	0	0	323
POP1	VRM1	322	1	0	0	0	0	0	0	0	0	0	0	323
POP1	WDOW	306	17	0	0	0	0	0	0	0	0	0	0	323
POP1	YEY1	271	49	3	0	0	0	0	0	0	0	0	0	323
POP1	YMTH	260	31	11	5	3	4	0	3	4	2	0	0	323
POP1	YTRK	318	5	0	0	0	0	0	0	0	0	0	0	323
POP1 Total		7592	1357	250	124	95	70	54	37	43	34	24	10	9690
POP2	ARRA	104	5	0	0	0	0	0	0	0	0	0	0	109
POP2	BANK	108	1	0	0	0	0	0	0	0	0	0	0	109
POP2	CNRY	108	1	0	0	0	0	0	0	0	0	0	0	109
POP2	DBRK	87	21	1	0	0	0	0	0	0	0	0	0	109
POP2	GSRK	106	3	0	0	0	0	0	0	0	0	0	0	109
POP2	LSPN	105	4	0	0	0	0	0	0	0	0	0	0	109
POP2	POP	0	0	0	0	0	0	0	0	0	8	58	43	109
POP2	RDBD	105	4	0	0	0	0	0	0	0	0	0	0	109
POP2	REDS	106	3	0	0	0	0	0	0	0	0	0	0	109
POP2	REYE	93	15	1	0	0	0	0	0	0	0	0	0	109
POP2	RSTN	107	2	0	0	0	0	0	0	0	0	0	0	109
POP2	SHRP	72	35	2	0	0	0	0	0	0	0	0	0	109
POP2	SNOS	86	23	0	0	0	0	0	0	0	0	0	0	109
POP2	SSPN	106	3	0	0	0	0	0	0	0	0	0	0	109
POP2	WDOW	106	3	0	0	0	0	0	0	0	0	0	0	109
POP2	YMTH	100	9	0	0	0	0	0	0	0	0	0	0	109
POP2	YTRK	108	1	0	0	0	0	0	0	0	0	0	0	109
POP2 Total		3024	133	4	0	0	0	0	0	0	8	58	43	3270
SSP1	LSPN	105	48	5	0	0	0	0	0	0	1	0	0	159
SSP1	REYE	158	1	0	0	0	0	0	0	0	0	0	0	159
SSP1	SNOS	158	1	0	0	0	0	0	0	0	0	0	0	159
SSP1	SSPN	0	0	1	0	0	0	0	0	0	5	48	105	159
SSP1 Total		4555	50	6	0	0	0	0	0	0	6	48	105	4770
URCK	ARRA	447	61	7	0	0	0	0	0	1	0	0	0	516
URCK	BANK	499	15	1	1	0	0	0	0	0	0	0	0	516
URCK	BCAC	399	45	26	11	10	7	5	3	5	2	0	3	516
URCK	BLCK	515	0	0	0	0	0	0	0	0	0	0	1	516
URCK	BLGL	481	30	4	0	0	0	1	0	0	0	0	0	516
URCK	BLUR	506	4	1	2	1	0	0	0	1	1	0	0	516
URCK	CHNA	318	14	13	33	15	13	16	19	18	13	9	35	516
URCK	CLPR	496	13	4	1	1	0	0	0	0	0	0	1	516
URCK	CNRY	490	22	2	0	0	0	0	0	0	0	1	1	516
URCK	COPP	425	27	26	21	7	4	3	0	1	0	0	2	516
URCK	DBRK	328	56	24	16	20	9	15	13	14	10	10	1	516
URCK	GSPT	474	33	1	2	1	3	0	1	0	0	0	1	516

Table 9. Distributions of species proportions in rockfish market categories, 1997 to 1999 (cont.).

		Number of samples with given species proportion												
Market			- - - - - - - - - - - - - Species proportion mid-point - - - - - - - - - - - - -											
Category	Species	0	0.05	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.95	1	Total
URCK	GSRK	452	56	7	1	0	0	0	0	0	0	0	0	516
URCK	LSPN	515	1	0	0	0	0	0	0	0	0	0	0	516
URCK	POP	371	119	11	7	5	2	1	0	0	0	0	0	516
URCK	QLBK	376	55	36	24	10	7	3	1	1	2	0	1	516
URCK	RDBD	373	127	12	2	0	1	1	0	0	0	0	0	516
URCK	REDS	455	53	5	1	2	0	0	0	0	0	0	0	516
URCK	REYE	367	71	26	11	8	6	5	4	6	4	4	4	516
URCK	RSTN	479	34	2	1	0	0	0	0	0	0	0	0	516
URCK	SHRP	441	67	4	3	0	1	0	0	0	0	0	0	516
URCK	SLGR	386	59	24	18	4	9	3	2	7	1	3	0	516
URCK	SNOS	394	105	10	3	3	1	0	0	0	0	0	0	516
URCK	SRKR	436	33	17	9	5	5	4	3	2	1	0	1	516
URCK	SSPN	502	14	0	0	0	0	0	0	0	0	0	0	516
URCK	VRM1	418	22	16	11	9	8	8	7	6	2	0	9	516
URCK	WDOW	485	31	0	0	0	0	0	0	0	0	0	0	516
URCK	YEY1	287	71	34	26	16	15	13	13	14	5	3	19	516
URCK	YMTH	457	44	5	4	3	1	1	1	0	0	0	0	516
URCK	YTRK	497	17	0	0	0	0	0	0	0	0	0	2	516
URCK Total		13069	1299	318	208	120	92	79	67	76	41	30	81	15480
WDW1	DBRK	240	1	0	0	0	0	0	0	0	0	0	0	241
WDW1	REDS	240	1	0	0	0	0	0	0	0	0	0	0	241
WDW1	SHRP	240	1	0	0	0	0	0	0	0	0	0	0	241
WDW1	WDOW	0	0	0	0	0	0	0	0	0	5	22	214	241
WDW1	YTRK	217	20	4	0	0	0	0	0	0	0	0	0	241
WDW1 Total		6962	23	4	0	0	0	0	0	0	5	22	214	7230
YTR1	ARRA	296	1	0	0	0	0	0	0	0	0	0	0	297
YTR1	CNRY	290	5	1	0	0	0	0	0	0	0	0	1	297
YTR1	DBRK	296	1	0	0	0	0	0	0	0	0	0	0	297
YTR1	GSRK	294	3	0	0	0	0	0	0	0	0	0	0	297
YTR1	POP	296	1	0	0	0	0	0	0	0	0	0	0	297
YTR1	REDS	295	2	0	0	0	0	0	0	0	0	0	0	297
YTR1	RSTN	296	1	0	0	0	0	0	0	0	0	0	0	297
YTR1	SLGR	293	4	0	0	0	0	0	0	0	0	0	0	297
YTR1	WDOW	252	44	1	0	0	0	0	0	0	0	0	0	297
YTR1	YTRK	1	0	0	0	0	0	0	0	0	2	56	238	297
YTR1 Total		8549	62	2	0	0	0	0	0	0	2	56	239	8910
Grand Total		57854	3049	607	343	219	164	135	108	130	119	354	1028	64110

Figure 1. Predicted versus observed species proportions from statistical approach (11) and from the current process of borrowing estimates from similar strata.

From statistical approach (11)



From the borrowing algorithm