1999/2000 National Marine Recreational Fishing Survey: harvest estimates

R. O. Boyd[1]
J. L. Reilly[2]

[1]Kingett Mitchell Ltd

[2]Statistical Insights Ltd

Draft New Zealand Fisheries Assessment Report
September 2002

1. Executive Summary

Boyd, R. O.; Reilly, J. L. (2002). 1999/2000 National Marine Recreational Fishing Survey: recreational harvest estimates.
Draft New Zealand Fisheries Assessment Report 2002/XX

Recreational harvest estimates for 1999-2000 (1 December 1999 to 30 November 2000) are presented for a wide range of fish and shellfish species. The estimates are based on a similar, but enhanced, survey methodology that has been adopted for estimating recreational harvests in previous surveys. The survey methodology involves combining results from at three separate but related surveys. An estimate of fisher prevalence derived from a nation-wide face to face survey is combined with detailed diary data of recreational harvests recorded by a nation-wide sample of recreational fishers recruited by telephone, to estimate recreational harvests in numbers of fish or shellfish. Estimated harvests in numbers of fish or shellfish were converted to total harvest weight using the results of a boat ramp survey to estimate the mean weight of recreationally harvested fish and shellfish.

Key enhancements over previous surveys included the use of a face to face survey for measuring fisher prevalence, improved methods for weighting up diarists' harvests using extensive demographic data and a more appropriate method for estimating coefficients of variation.

Estimates for the 1999-2000 national marine recreational fishing survey are much higher than the estimates from previous surveys. The harvest estimate for SNA1 which has a c.v. of 11% is in excess of 6.9 million fish and 6 200 tonnes. Very few of the harvest estimates have c.v.s of less than 20%. Most of the fishstock recreational harvest estimates presented in the report are higher than previous estimates by a factor of two to three times. Coefficients of variation (c.v.s) for the harvest estimates are much larger than estimated for previous surveys but are more reflective of the complex nature of the survey design and the highly skewed nature of diarists' harvests. These factors were not taken into account in the method used for estimates of c.v.s in previous surveys.

Some of the 1999-2000 harvest estimates, particularly the estimates for a number of key fishstocks in QMA2 appear to be implausibly high. While the reasons for this are not known, the small sample size for this area may have resulted in a biased sample of diarists.

Results from pilot surveys undertaken as part of the 1999-2000 survey together with a review of the available literature strongly suggest that previous harvest estimates from the 1996 national survey and earlier regional surveys are highly unreliable and probably much too low. Therefore, caution should be exercised in comparing the estimates presented here with the estimates from previous surveys as such comparisons are likely to be misleading.

The improved survey methodology and estimation procedures adopted for the 1999-2000 national marine recreational fishing survey mean that the reported harvest estimates should be more accurate than the estimates from prior surveys. The much higher recreational harvest estimates have significant fisheries sustainability and management consequences. Future surveys to estimate recreational harvests will need to focus on making further improvements to the survey methodology and improving the precision of estimates.

2. Introduction

This report provides the estimated recreational harvests of fish and shellfish from the 1999-2000 national marine recreational fishing survey (NMRFS). The principal objective of the NMRFS is to estimate the harvest by recreational fishers of the main fishstocks.

The 1999-2000 NMRFS comprised a number of separate but related surveys. The overall methodology follows the same approach adopted for estimating marine recreational harvests in 1996 described by Bradford (1998a). A nation-wide telephone survey was used to obtain demographic data on marine recreational fishers and to recruit a sample of recreational fishers to keep fishing diaries (Ransom & Boyd in prep.). The random sample of recreational fishers recruited in the telephone survey kept detailed diaries of their fishing activity over the 12 month period from 1 December 1999 to 30 November 2000 (Boyd & Gowing in prep. A). A door to door 12 month recreational fishing prevalence survey (Reilly in prep. A) was used to estimate the numbers of recreational fishers. Reilly (in prep. B) describes the weighting method used to scale up the harvest data from diarists to the estimated total marine recreational harvest in numbers of fish or shellfish for each species or fishstock. A boat ramp survey to measure fish and shellfish has been used to estimate the mean weight of recreationally harvested fish and shellfish (Boyd & Gowing in prep. B).

Two important changes to the methods used in 1996 have been adopted for estimating harvests in the 1999-2000 NMRFS. One has been the use of a separate door-to-door survey to estimate fisher prevalence and fishery entry and exit. Ransom & Boyd describe the reasons for this in some detail but in summary the results of a number of pilot telephone surveys indicated that the use of a telephone survey may not provide a reliable estimate of the numbers of marine recreational fishers. In 1996, the telephone survey used to recruit diarists was also used to estimate fisher prevalence. The second change for the 1999-2000 NMRFS has been the methodology used for weighting diarists' harvests up to the estimated total harvests. Reilly (in prep. B) describes how the weighting process has been improved over that adopted in 1996 by incorporating behavioural and demographic weighting variables to combat non-response bias and how substantial assumptions regarding fisher entry and exit have also been removed, increasing the reliability of the harvest estimates.

The accuracy or reliability of the harvest estimates depends on a number of factors. One of the most important is the number of diarists reporting they fished for or harvested a particular species or fishstock. Another is the skewness of the annual harvest distributions. Bradford (2000) found that in the order of 450 or more fishers would be required to detect changes from –20% to +25% in the mean annual harvest of a particular fishstock and that the sample size required to detect changes increased with the skewness of the annual harvest distributions. Bradford (1998a) and Boyd and Gowing (in prep. A) showed that diarists' annual harvest are highly skewed, with a small proportion of diarists responsible for a significant proportion of the total diarists' harvest. It is therefore possible for a small number of diarists to have a significant impact on the total harvest estimate for a particular stock.

Reilly (in prep. B) notes that sampling methods that focus data collection effort more on more frequent fishers would enable better harvest estimates. The representativeness of diarists is also a factor in the accuracy of the estimates. Boyd & Gowing (in prep. A) show the high attrition of diarists over the course of the diary year. Reilly (in prep. B) describes the adjustments to account for non-response as well as other skews relative to population in estimating total diarists' harvests, but it is inevitable that some skews will remain.

Overall, the estimation of marine recreational harvests from the 1999-2000 NMRFS incorporate some significant advances over previous methods and the estimates should accordingly be more accurate. The most accurate estimates should be those for species or stocks such as snapper, kahawai and blue cod that were fished for or harvested by the largest number of diarists.

3. Methods

3.1. Estimated harvests

The total numbers of fish or shellfish harvested from each fishstock (except for the Chatham Islands, which is not included in the NMRFS) has been estimated by weighting up diarists' harvests from the 1999-2000 diary survey as described by Reilly (in prep. B). Briefly, the weighting methodology produced one weight for each diarist and each diary period for which the diarist provided data. These weights adjust for features of the survey design and help to correct for non-response biases. Total harvest estimates for each fishstock or species in numbers of fish or shellfish were then calculated by applying these diarist weights to their diary data.

In this report, numbers of fish or shellfish harvested have been rounded to the nearest 1 000. Where estimated harvests for a fishstock are greater than zero but less than 500, they are given as <500. These protocols follow Bradford (1998b).

3.2. Coefficients of variation

The formulae used for estimating coefficients of variation (c.v.s) for harvest estimates in previous recreational fishing surveys (see Bradford 1998b and Tierney *et al.* 1997) did not properly reflect the complex survey designs on which the harvest estimates were based. It was agreed that a more sophisticated re-sampling approach would be used to estimate c.v.s for the 1999/2000 harvest estimates, to remove the need for distributional assumptions and to better reflect the effect of the survey design.

Coefficients of variation for the 1999-2000 harvest estimates have been calculated using the extended delete-a-group jack-knife (EDAGJK) method (Kott 1999, Kott 2001). This is a re-sampling method (as are balanced repeated replication (BRR) and the bootstrap), so it involves repeatedly calculating harvest estimates on sub-samples carefully selected from the original data set to reflect the original sample design. These results are then combined to estimate the sampling errors (expressed in this report as c.v.s).

A full description of the EDAGJK method used to calculate c.v.s for the 1999-2000 harvest estimates is given in Appendix 1.

3.3. Estimation of total harvest weights

Length data from the 1999-2000 boat ramp survey conducted as part of the 1999-2000 NMRFS is available for a wide range of species. This length data has been used to estimate the mean weight (i.e., mean green weight) for each fishstock or species using known length-weight relationships compiled from a range of published and unpublished sources (Boyd & Gowing, in prep. B.). The mean weights have been used to convert total numbers harvested to total harvest weight (green weight) by simply multiplying the estimated total numbers of fish or shellfish harvested from each fishstock by the appropriate mean weight. All total harvest weights have been converted to tonnes and rounded to one decimal place. For many species for which length data is available from the boat ramp survey, no length-weight relationships have been established so the total harvest weight can not be estimated. For these species, harvest estimates are given in numbers of fish or shellfish only.

Male and female rock lobsters show sexual dimorphism and have different tail width to weight relationships. The sex ratios obtained for rock lobster from the boat ramp survey have been used to estimate the proportion of each sex harvested recreationally from each rock lobster fishstock, except for CRA3 where the CRA2 sex ratio was adopted. The mean weight determined for each sex from the boat ramp survey was then used to estimate the total rock lobster harvest weight by calculating the harvest weight separately for each sex and summing these estimates. For some stocks, notably CRA3, 7, 8 and 9, sample sizes for estimating mean rock lobster weights were very small (Table 1). In spite of the small sample sizes for these fishstocks, these were used for estimating total harvest weights as the species mean weights were deemed inappropriate.

193. Estimated 1999-2000 Recreational Harvests

754974721.0. Mean fish and shellfish weights

Table 1 provides the mean fish and shellfish weights derived from the 1999-2000 boat ramp survey that have been used in estimating total recreational harvest weights. Where the sample sizes from the boat ramp survey were considered large enough to calculate a reliable mean weight for a particular fishstock (usually n = 50 or more) the mean weight for that fishstock was used. Where sample sizes for estimating mean weights for a particular fishstock were considered too small and potentially unreliable, the length data from all areas for that species from the boat ramp survey were combined to calculate a mean species weight. These mean species weights were used to estimate total harvest weights for those fishstocks. In a small number of cases, there were large differences in mean weights for sub-areas of a fishstock. For these estimates, mean weights from the boat ramp survey were also calculated for the sub-areas.

As noted in Boyd & Gowing (in prep. B), all of the length data collected to estimate mean weights has been derived from sampling the harvest from boats, whilst sampling from boat ramps. Shore fishing was not sampled. Therefore mean weights derived from the boat ramp survey may not be fully representative for those fishstocks, species or areas where shore fishing accounts for a significant proportion of the total harvest.

3.4. Harvest estimates

Unlike the harvest estimates from the 1996 recreational fishing surveys (Bradford, 1998a), no attempt has been made to estimate harvests for fishers under age 15 years, who were not diarists (Reilly, in prep B).

Table 2 provides recreational harvest estimates for 1999-2000 for finfish species by fishstock or quota management area, and in some cases by sub-area, both in numbers of fish and in tonnes. For comparative purposes, the commercial harvest for the 1999-2000 fishing year is also given for each fishstock, where available. Table 3 provides the same results for shellfish species. Table 4 provides recreational harvest estimates in numbers of fish and shellfish respectively, for those species or fishstocks where there is no mean weight available to convert the numbers harvested to tonnes. For a number of fishstocks in Table 2, alternative harvest estimates are provided for sub-areas using mean fish weights calculated separately for the sub-area from the boat ramp survey.

The c.v.s for the harvest estimates are also provided in Tables 2, 3 and 4. In addition, the number of diarists who reported they harvested the fishstock and the unweighted diarists' total harvest reported for each fishstock is given. These latter data are provided to assist in assessing the sample sizes available to estimate harvests. Along with the c.v.s, they assist in interpreting the reliability of the estimates. Many of the harvest estimates for minor species are based on a very small number of diarists.

Table 5 provides harvest estimates for so called 'generic species', where diarists reported their harvest using non-specific names such as cod or shark which could not be attributed to a particular fish species. These harvests will include some species for which harvest estimates are given in Tables 2, 3 and 4, although with the exception of 'mussels' and 'oysters' plus 'cod' in QMA7 the harvests using generic names are relatively small. The majority of 'mussels' in Table 5 are highly likely to be green-lipped mussels in all areas. The majority of 'oysters' in Table 5 are likely to be Pacific oysters in the North Island, dredge oysters in the southern half of the South Island and a mix of the two in the northern half of the South Island. Most of the 'cod' in QMA7 are likely to be blue cod.

3.4.1. Blue cod

The BCO3 harvest estimate (Table 2) is just over 1 million fish with a c.v. of 29%, or 752 tonnes. This is more than four times larger than the reported commercial harvest of 168 tonnes in 1999-2000. The BCO7 harvest estimate is 542 000 fish with a c.v. of 20%, or 288 tonnes, more than ten times the reported commercial harvest. There were also significant recreational harvests of BCO2, BCO5 and BCO8.

3.4.2. Kahawai

The KAH1 harvest estimate (Table 2) is 1.86 million fish with a c.v. of 13%, or 2 195 tonnes. There is little difference (4.1%) between the combined sub-area harvest estimates using sub-area mean weights and the overall estimated harvest weight.

The harvest estimate for KAH2 of 1.81 million fish is nearly as large as the estimate for KAH1 but the c.v. is substantially higher at 74%, indicating the KAH2 harvest estimate is not reliable.

Substantial kahawai harvests were also estimated for KAH3 (413,000 fish) and KAH9 (337,000 fish).

3.4.3. Snapper

The SNA1 harvest estimate (Table 2) is in excess of 6.9 million fish with a c.v. of 11%, or 6,200 tonnes, approximately 1.4 times larger than the reported commercial harvest. For SNA1, there is little difference (< 1%) in the overall estimated harvest weight using the alternative sub-area mean weight estimates.

The SNA2 harvest estimate is 268 000 fish with a c.v. of 41%, or 718 tonnes. However, the estimated total harvest weight is based on a mean fish weight of more than 2.68 kg estimated from the 1999-2000 boat ramp survey. This mean fish weight is more than double the estimated mean weight reported of 1.28 kg for the 1993 boat ramp survey (see Bradford 1998b) and has a strong influence on the high tonnage estimate.

The SNA7 harvest estimate is 63 000 fish with a c.v. of 21%, or 134.4 tonnes, compared to a commercial harvest of 174 tonnes.

The SNA8 harvest estimate is 648 000 fish with a c.v. of 21%, or 660 tonnes. However, the estimated total harvest weight using sub-area mean weights is nearly 10% lower at 597 tonnes. This is a reflection of a large recreational snapper harvest in SNA8 N (QMA9) with a lower mean weight, and a much smaller harvest of SNA8 S (QMA8) with a much higher mean weight.

3.4.4. Tarakihi

The estimated harvest of TAR1 (Table 2) is more than 1 million fish with a c.v. of 19%, or 636 tonnes, nearly half as large as the reported commercial harvest of 1,387 tonnes in 1999-2000. A high proportion of the TAR1 harvest was taken in the Bay of Plenty. The TAR2 harvest estimate is 310 000 fish with a c.v. of 27%, or 191 tonnes. Much smaller recreational harvests of tarakihi were taken from other areas.

3.4.5. Trevally

The estimated harvest of TRE1 (Table 2) is 701 000 fish with a c.v. of 13%, or 677 tonnes, approximately half as large as the reported commercial harvest. The TRE2 harvest of 153 000 fish is also large, but has a c.v. of 60% indicating the estimate is not reliable. The TRE7 harvest estimate of 69 000 fish has a c.v. of 27%.

3.4.6. Rock lobster

Rock lobster harvest estimates are given in Table 3, by sex. The availability of sex ratios and mean weight data by sex calculated from the boat ramp survey database has enabled harvest estimates to be calculated in tonnes that take into account the sexual dimorphism of rock lobsters and the sex ratio encountered in each recreational fishery.

The CRA2 harvest estimate of 235 tonnes, is equal in size to the reported CRA2 commercial harvest. In CRA 1, 3, 4, 5 and 9, the recreational harvest estimates are also large in relation to the size of the commercial harvest, ranging from a third to 80% of the commercial harvests for each of those fishstocks. However, except for CRA 2 and CRA4 with c.v.s of 26% and 24% respectively, the c.v.s for the other CRA estimates suggest they are not very reliable, especially CRA1, CRA7 and CRA9 where the c.v.s are very large.

3.4.7. Paua

Paua harvests (Table 3) in both numbers and total weight have been estimated for all fishstocks in 1999-2000. Mean weight estimates were calculated from the boat ramp survey database for all fishstocks except PAU6. The mean paua weight for PAU5 was used for estimating the total recreational harvest weight in PAU6.

The estimated PAU1 and PAU2 recreational harvests stand out as very large in relation to the commercial fisheries in these areas. There is a negligible commercial fishery in PAU1. The mean length of PAU1 from the boat ramp survey (Table 1) is smaller than the minimum legal size, indicating the small size of paua harvested. The PAU2 harvest of more than 400 tonnes appears implausibly large, and has a c.v. of 46% which suggests it is not reliable. The c.v.s for all of the paua fishstock harvest estimates indicate that they have a low to moderate degree of reliability.

3.4.8. Scallop

Scallop harvests (Table 3) have been estimated for QMAs and for sub-areas representing each scallop stock, where diary data is available. The estimated SCA harvest in QMA7 (which incorporates the Challenger Scallop Enhancement Area) was 3.4 million scallops with a c.v. of 20%, or 339 tonnes. The SCA harvest in QMA1 was 634 000, with a c.v. of 34%, or 70 tonnes. The separate harvests estimated for the Northland and Coromandel scallop stocks within QMA1 were calculated using the harvest data reported from the appropriate diary zones corresponding approximately to the boundaries of each fishery.

3.4.9. Other Species

Significant recreational harvests were also estimated for a number of other finfish and shellfish species. In Table 2, these include flatfish in all areas, John dory in JDO1, grey mullet in GMU1, red gurnard in GUR1 and GUR2, hapuku/bass in all areas, kingfish in KIN1 and KIN2, blue moki in MOK1, and red cod in RCO3. The c.v.s for most of these other estimates are well over 20%, with the exception of those for GUR1 and KIN1 which are 16% and 18% respectively. Tables 4 and 5 show significant harvests for shellfish species, especially cockle, kina, mussels and pipi. Most of these shellfish estimates have large c.v.s.

~~5.~~4. Discussion

4.1. Comparison with earlier recreational harvest estimates

Recreational harvest estimates from the 1999-2000 NMRFS are substantially greater than estimated from either the 1996 NMRFS (Bradford 1998) or the earlier regional surveys undertaken by the Ministry of Fisheries from 1991-92 to 1993-94 (Tierney et al. 1997). In most cases, harvest estimates for key fishstocks such as snapper in SNA1 are two to three times higher than previously estimated. This difference is almost entirely due to the much higher fisher prevalence estimated for the 1999-2000 NMRFS (38.9% of households) compared to 1996 (13.9% of households) as the diary results for the two surveys are very similar.

However, there are problems in attempting to make any comparisons between the 1999-2000 harvest estimates and the results of the earlier recreational fishing surveys reported by Bradford (1998) and Tierney *et al.* (1997). This is because the work undertaken for the 1999-2000 NMRFS indicates that the 1996 telephone recruitment survey reported by Bell (1996) is likely to have seriously underestimated fisher prevalence and therefore the associated harvest estimates.

Bell (1996) provides no contact rate, response rate or co-operation rate data for the 1996 NMRFS telephone survey. This lack of key quality assurance information is unusual for a survey of this size and creates very real problems when attempting to evaluate the quality and accuracy of the 1996 telephone survey. However, it is possible to evaluate the accuracy of the 1996 survey indirectly. Two other New Zealand telephone surveys on recreational fishing (NRB 1991, Ackroyd Walshe Ltd 1999) provided detailed contact and response rate data. Both document large numbers of unobtainable or not-in-service telephone numbers, non-contacts for valid numbers (i.e. not answered), business numbers and direct refusals. The ratio of successful interviews to direct refusals in NRB (1991) and Ackroyd Walshe Ltd (1999) is in the order of 2:1. The ratio of total telephone numbers selected to the number of completed interviews in the same surveys is in the order of 2.5:1. These published results conflict with Bell (pers. comm.) who indicates there were very few hard refusals and a low number of non-contacts, but is unable to provide documentation. Bell (1996) states that a 35,000 randomly selected telephone numbers were selected. Given the contact rate data reported by NRB (1991) and Ackroyd Walshe (1999), one would expect over 80,000 randomly selected telephone numbers would need to be selected to be able to obtain the reported 35,038 successful telephone interviews reported for the 1996 telephone survey. This estimate takes into account the proportion of unobtainable numbers, non-contacts, business numbers and refusals recorded in the other surveys. Bell (1996) makes no mention of any unsuccessful contacts in the survey report. This is inconsistent with the large number (perhaps in the order of 80,000 telephone numbers) that should be expected based on the fully documented NRB (1991) and Ackroyd Walshe (1999) surveys. The uncertainty associated with contact rates, non-documentation of important response rate results in Bell (1996), together with the inconsistencies with response rates compared to surveys which have well documented contact rate information mean that the 1996 telephone survey results reported by Bell (1996) must be considered as unreliable.

Pilot survey results (Ransom & Boyd in prep.) for the 1999-2000 NMRFS demonstrated that using the 1996 telephone survey questionnaire provided an opportunity for soft refusals at the start of the interview. Soft refusals are a known problem in the survey industry and occur when a question posed at the start of the survey enables a respondent to respond with a 'no' as a means of quickly exiting the interview, without having to personally refuse to be interviewed. Depending on the circumstances, soft refusals may invalidate quantitative results and it appears that this did occur in 1996. The 1996 telephone survey questionnaire (Bell 1996) provided an opportunity for soft refusals so it is inevitable that they did occur. The pilot survey results for the 1999-2000 telephone survey (Ransom & Boyd, in prep) confirm this. Therefore, the household fisher prevalence estimate of 13.9% reported for the 1996 NMRFS (Bell 1996) must be considered both unreliable and an underestimate. However, soft refusals may be responsible for only part of the under-estimation of fisher prevalence in 1996. The remainder is accounted for a failure to record hard refusals. Bell (pers. comm.) reports very few hard refusals, but as discussed earlier, this is demonstrably inconsistent with the published response patterns in NRB (1991) and Ackroyd Walshe Ltd (1999), both of which report significant numbers of hard refusals.

The 1996 harvest estimates effectively assume that soft refusals did not occur, or that if they did, that all potential diarists who were soft refusals would have recorded zero harvests. The 1996 harvest estimates also effectively assume the same for all other forms of non-response that may have occurred in the total number of households contacted (35,308), (Bell 1996). These assumptions will result in underestimation of the true harvests by an amount proportional to the harvest of these non-respondents. Any other non-respondents (e.g. due to non-contacts, unrecorded contacts, or perhaps even refusals not recorded as contacts) will have been effectively assumed by the 1996 survey methodology to have similar fishing behaviour to responding diarists.

It is possible to produce numerical illustrations of the likely underestimates in 1996 under specific assumptions. If we assume that 75% of the 35,038 telephone contacts reported by Bell (1996) were soft or hard refusals and that these contacts exhibited the same fishing behaviour and success on average as the 25% who were interviewed, then the true fishing prevalence in 1996 would have been 55.5% (= 4 x 13.87%). The correct 1996 harvest estimates under this assumption would be four times the reported figures. The assumption of 75% soft and hard refusals is not unreasonable given the magnitude of the discrepancy between the published contact rate data provided in NRB (1991) and Ackroyd Walshe Ltd (1999) and Bell's (pers. comm.) undocumented results.

In summary, the 1996 harvest estimates have to be considered as unreliable and too low as they are substantially dependent on what appear to be serious methodological errors. They are also inconsistent with well-documented contact and response rate patterns in other surveys. For identical reasons, the earlier regional survey harvest estimates (1991-92 to 1993-94) also have to be rejected as unreliable. Therefore, comparisons of the 1999-2000 NMRFS harvest estimates with those of the earlier surveys should be avoided as such comparisons are likely to be very misleading.

4.2. Coefficients of variation

Using a re-sampling method to calculate harvest c.v.s has eliminated the previous dependence on questionable distributional assumptions. It has also reflected the complex survey designs underlying the harvest estimates. Due to the former factor in particular, the c.v.s for the 1999-2000 NMRFS are also markedly larger than the c.v.s reported for recreational harvest estimates from previous surveys (Bradford 1998a, Tierney *et al.* 1997). These factors have outweighed improvements in the harvest estimation process that are believed to have substantially reduced the true c.v.s relative to previous surveys. (This reduction in the c.v.s was observed when preliminary c.v.s for the 1999-2000 harvest estimates were calculated using the Poisson distributional assumption adopted for earlier surveys.)

The bulk of the increase in c.v.s over those reported for the 1996 NMRFS is due to the skewed catch distribution among diarists, a feature of both surveys. For example, the c.v. of the harvest estimate for the BCO3 fishstock is 29%. Over half of the variability among the jack-knife replicates for this fishstock came from just one replicate, and was due to one diarist catching 467 blue cod over 33 fishing trips. This person accounted for 14% of the total harvest across all diarists. Whenever one person has such a substantial impact on the estimated harvest, this will result in a large jack-knife c.v.

The much larger c.v.s suggest that recreational harvest estimates appear to be considerably less reliable than may have been suggested in the reports on previous recreational fishing surveys. The higher c.v.s will need to be taken into account when interpreting the results of the 1999-2000 NMRFS. Very few (< 10) of the harvest estimates have c.v.s of less than 20%.

Looking beyond the current harvest estimates, alternative approaches to harvest estimation based on the existing data should be considered in the future. For example, model-assisted approaches that allow for highly skewed catch distributions and many zero catches may be productive. Alternative approaches to data collection may also be needed in the future. For example, recruiting larger numbers of frequent fishers would effectively decrease the skewness of the catch distribution, and potentially increase the reliability of the harvest estimates.

4.3. The 1999-2000 harvest estimates

The harvest estimates from the 1999-2000 NMRFS have been generated using a much more sophisticated approach than used for previous surveys by taking extensive demographic data into account. By properly accounting for demographic skews, much better harvest estimates have been generated which remove the need for many assumptions about diarist behaviour. While the 1999-2000 harvest estimates are also largely dependent on the prevalence result, soft refusals have been eliminated. The elimination of soft refusals appears to be the main reason for the marked increase in the estimated fisher prevalence. This in turn has resulted in much higher harvest estimates than suggested from previous recreational fishing surveys. The use of a face to face survey to estimate fisher prevalence has produced a higher response rate that should be more accurate than an estimate derived in a telephone survey. The face to face survey method provides more complete coverage of the New Zealand population for estimating fisher prevalence, as it is well documented that telephone penetration is low amongst some socio-economic groups.

Because of the higher fisher prevalence result, the harvest estimates for many of the most recreationally important fish stocks are higher than previously estimated even though the 1996 and 1999-2000 raw diary data is very similar. For some important fishstocks such as SNA1 and BCO7, the new recreational harvest estimates are now considerably larger than the reported commercial harvest. However, although the new harvest estimates are on average 2-3 times higher form most fish stocks than estimates from previous surveys, they should be evaluated on their own merits rather than by comparison with the prior estimates. Perceptions of what may be plausible harvests are strongly influenced by the harvest estimates from the 1996 NMRFS. As discussed earlier, there is very strong circumstantial evidence to suggest that the estimates from the earlier fishing surveys may be unreliable and too low. C.v.s for previous estimates were also optimistic as they were calculated using incorrect assumptions about diarists' harvest distributions.

For some fishstocks, such as the estimate for SNA1 (which has the lowest c.v. of any of the estimates at 11%), the harvest estimates are large. However, the SNA1 estimate of 6.9 million fish is not inconsistent with knowledge of factors that might influence harvests. These include the large population in the northern half of the North Island, year-round fishing opportunities, safe waters for small vessels, the large number of privately owned boats, the high abundance of snapper in inshore waters and the dominance of snapper as a desirable recreational fish to target. In other areas, especially in QMA2, the harvest estimates for a number of stocks seem implausibly large, although some caution must be expressed in adopting this view given that perceptions can be strongly influenced by previous estimates. However, given the relatively small diarist sample size in QMA2 and the skews apparent in the distribution of diarists' reported harvests, it is possible that the diarist sample in QMA2 is not representative. For example, it may have contained a higher proportion of frequent fishers who caught more fish on average than occurred in the population at large. Many of the QMA2 harvest estimates (e.g., GUR2, HPB2, KAH2, KIN2, SNA2, TAR2, TRE2, CRA3, CRA4, PAU2, SUR2) appear higher than might be expected when compared to harvests from adjacent areas with a larger population. The c.v.s for most of the QMA2 harvest estimates are very high indicating they are not very reliable. For example, the c.v. for the very large KAH2 harvest estimate is 74%.

However, overall, the revised and improved survey methods adopted mean that the 1999-2000 NMRFS harvest estimates must be considered as more accurate and much closer to the true harvests than previous estimates. However, they are estimates and may be subject to error from a range of possible sources. The improved methods for calculating c.v.s indicate that the reliability or precision of many of the estimates is lower than may have been assumed in the past, with very few of the harvest estimates having c.v.s of less than 20%. A significant contributor to the estimated precision is the highly skewed catch distribution demonstrated by diarists, where a small proportion of diarists are responsible for a high proportion of the total harvest.

The high estimates of recreational harvest for many key species in 1999-2000 will have significant sustainability and fisheries management consequences. Under these circumstances, and given the high c.v.s of the estimates, there is a need for ongoing research to affirm in detail the robustness of each component of the survey methodology as well as the reliability of the harvest estimates that are generated using the current methodology.

5. Acknowledgements

This project has been carried out by Kingett Mitchell Ltd under contract REC9803 funded by the Ministry of Fisheries. ACNielsen Ltd undertook the telephone recruitment, face to face prevalence and diary surveys under subcontract to Kingett Mitchell Ltd. The commitment of all personnel at ACNielsen Ltd who have contributed their skills and experience to the project is gratefully acknowledged. All survey respondents who willingly gave their time deserve particular recognition as without their efforts the survey would not be possible.

6. References

Ackroyd Walshe Ltd, 1999: Motivations and perceptions of fishers. Report for Ministry of Fisheries Project REC9802. 134 p. + Appendices. (Unpublished report held in Ministry of Fisheries, Wellington)

Bell, J.D. & Associates 1996: Ministry of Fisheries and NIWA National Marine Recreational Fishing Survey: Findings from the telephone survey. NIWA Internal Report (Fisheries) No. 256 (Draft report held in NIWA library, Wellington). 21 p.

Boyd, R. O,; Gowing, L. (in preparation A): 1999-2000 National Marine Recreational Fishing Survey: results from the diary survey. *Draft Fisheries Assessment Report 2002/XX.* 71 p.

Boyd, R. O,; Gowing, L. (in preparation B): 1999-2000 National Marine Recreational Fishing Survey: results from the boat ramp survey. *Draft Fisheries Assessment Report 2002/XX.* 67 p.

Bradford, E. 1998a: National marine recreational fishing survey 1996: scaling the diary survey results to give the total recreational harvest. NIWA Technical Report 17. 33 p.

Bradford, E. 1998b: Harvest estimates from the 1996 national marine recreational fishing survey. New Zealand Fisheries Assessment Research Document 98/16. 26 p. (Unpublished report held in NIWA library, Wellington).

Bradford, E. 2000: Sample sizes needed for reliable estimates in marine recreational fishing surveys. *New Zealand Fisheries Assessment Report 2000/36.* 37p

Kott, P.S. 1999: The Extended Delete-A-Group Jack-knife. *Bulletin of the International Statistical Institute*, 52nd Session, Contributed Papers.

Kott, P.S. 2001: The Delete-A-Group Jack-knife. *Journal of Official Statistics,* ***17***, 521-526.

National Research Bureau 1991: The economic worth of recreational fishing in New Zealand. National Research Bureau Ltd. Auckland. 37 p.

Reilly, J. L., (in preparation A) 1999/2000 National Marine Recreational Fishing Survey: report on fishing prevalence measurement. *Draft Fisheries Assessment Report 2002/XX.* 105p.

Reilly, J. L., (in preparation B) 1999/2000 National Marine Recreational Fishing Survey: weighting methodology for harvest estimates. *Draft Fisheries Assessment Report 2002/XX.* 24p.

Smith, H. 2001: Investigation of the delete-a-group variance estimator for the HLFS. Statistics New Zealand, Wellington, New Zealand.

Teirney, L. D.; Kilner, A.R.; Millar, R.B.; Bradford, E.; Bell, J.D. 1997: Estimation of recreational harvests from 1991-92 to 1993-94. New Zealand Fisheries Assessment Research Document 97/15. (draft report held in NIWA library, Wellington)

~~Annala, J. H., Sullivan, K. J., O'Brien, C. J., Smith, N. W. McL., Varian, S. J. A. (Comps) 2002: Report from the Fishery Assessment Plenary, May 2002: stock assessments and yield estimates. 640 p. (Unpublished report held in NIWA library, Wellington).~~

Table 1: Mean weight (g) and mean length of recreational fish (cm) and shellfish (mm) landed by fishstock or species from the 1999-2000 boat ramp survey used for estimating recreational harvest weights. For some fishstocks, mean weight and mean length are also provided for sub-areas.

Fish Stock/Species	Sex	n	Mean Weight (g)	*c.v.* (%)	Mean Length
BAR		269	1 971	7	70.1
BAR1		179	1 790	9	67.1
BCO1		180	605	7	33.0
BCO2		990	858	4	36.7
BCO3		6 131	733	1	35.4
BCO3 N		4 928	719	1	35.2
BCO3 S		1 203	789	3	35.9
BCO5		723	701	3	35.0
BCO7		2 426	532	2	31.9
BCO8		2 710	809	1	36.6
BUT		117	1 138	8	41.5
CRA1	F	118	646	6	69
	M	205	1 130	6	69
CRA2	F	735	584	2	67
	M	672	888	3	64
CRA3	F	24	696	12	71
	M	1	883	0	65
CRA4	F	1 072	640	2	68
	M	1 191	1 014	3	67
CRA5	F	2 974	677	1	70
	M	2 280	981	1	68
CRA7	F	33	860	12	73
	M	31	1 100	18	70
CRA8	F	2	1 552	18	93
	M	2	1 963	209	85
CRA9	F	49	712	11	71
	M	35	959	10	66
EMA		40	760	25	36.0
FLA		228	422	5	32.5
FLA1		89	384	6	31.7
FLA3		109	481	7	34.2
GMU		55	930	9	34.0
GMU1		55	930	9	34.0
GUR		6 710	481	1	34.1
GUR1		5 043	479	1	33.9
GUR2		831	607	3	37.8
GUR7		317	308	8	29.8
GUR8		515	406	5	33.1
HPB		818	8 479	7	81.4
HPB1		54	5 744	14	71.8
HPB2		341	8 220	6	81.8
HPB3		53	3 729	18	62.3
HPB7		49	7 269	9	80.3
HPB8		321	10 184	14	86.0
JDO		431	1 770	4	40.6
JDO1		415	1 767	1	40.6
JMA		189	504	15	29.8
JMA1		147	359	17	27.1

Table 1 - continued:

KAH1	5 525	1 180	1	40.9
KAH1(BPLE)	3 952	1 243	1	41.9
KAH1(ENLD)	902	1 144	3	40.3
KAH1(HAGU)	671	861	4	35.7
KAH2	481	1 624	3	47.0
KAH3	1 418	1 615	2	44.7
KAH9	2 051	1 306	2	42.4
KIN	898	7 438	6	79.1
KIN1	296	6 275	9	73.9
KIN2	94	5 443	15	70.8
KIN8	465	8 986	8	85.0
LIN2	2	2 839	200	82.5
MOK	183	1 612	9	43.0
MOK3	106	1 465	9	41.8
PAU1	66	226	3	119
PAU2	1 067	325	1	138
PAU3	95	283	2	131
PAU5	270	357	2	143
PAU7	865	317	1	137
RCO	386	388	12	32.4
RCO2	337	289	10	30.4
RCO3	28	1 349	24	50.0
SCA(QMA1)	433	110	2	108
SCA(CORO)	179	117	3	111
SCA(ENLD)	215	102	2	105
SCA(QMA7)	2 981	100	1	104
SCA(QMA9)	1 691	137	1	117
SCH	140	2 554	45	57.1
SCH1	121	2 479	53	53.3
SNA1	22 672	904	1	33.3
SNA1(BPLE)	10 135	872	1	33.1
SNA1(ENLD)	3 880	1 154	3	35.4
SNA1(HAGU)	8 657	830	1	32.7
SNA2	127	2 678	10	49.7
SNA7	288	2 148	9	44.4
SNA8	7 000	1 020	2	34.7
SNA8 N	3 180	861	3	32.9
SNA8 S	3 820	1 151	2	36.1
SPD	38	1 238	17	65.6
SPE	686	609	3	30.8
SPE2	107	490	7	28.5
SPE3	561	629	3	31.3
SPO	201	1 329	17	64.5
SPO7	74	1 017	14	62.1
TAR	7 951	558	1	30.3
TAR1	4 861	614	1	31.4
TAR2	410	616	6	31.0
TAR3				

Species	Fishstock	Number of diarists harvesting the stock	Unweighted harvest reported by Diarists	cn	Estimated recreational harvest in le

Table 2 - continued:

Blue moki	MOK1[1]	33	162	81 000	37	1 612	130 607.0	381
	MOK3	24	147	36 000	32	1 465	52.5	56
	MOK5[1]	4	144	38 000	89	1 612	61 272.4	5
Red cod	RCO1[1]	25	39	21 000	36	388	8.3	3
	RCO2	39	133	39 000	25	289	11.3	130
	RCO3	87	637	207 000	25	1349	279.8	4 824
	RCO7[1]	14	49	23 000	50	388	9.0	633
School shark	SCH1	22	54	27 000	42	2 479	65.9	820

Species	Fishstock or area	Sex	Number of diarists harvesting stock	Unweighted harvest reported by diarists	Estimated harvest in numbers of shellfish	*c.v.* (%)	Mean weight (g)	Estimated harvest by sex (tonnes)	Total estimated harvest (tonnes)	1999-2000 Commercial Harvest (tonnes)
Rock lobster	CRA1	All	20	183	107 000	59			102.3	125.7
		F			39 000		646	25.3		
		M			68 000		1130	77.0		
	CRA2	All	71	730	324 000	26			235.9	235.3
		F			169 000		584	98.6		
		M			155 000		888	137.2		
	CRA3[1]	All	31	513	270 000	40			212.4	328.5
		F			141 000		696	98.3		
		M			129 000		883	114.1		
	CRA4	All	59	997	371 000	24			310.9	577.3
		F			176 000		640	112.7		
		M			196 000		1014	198.2		
	CRA5	All	43	593	151 000	34			122.3	348.7
		F			86 000		677	58.0		
		M			66 000		981	64.3		
	CRA7	All	3	9	1 000	63			1.3	56.5
		F			1 000		860	0.6		
		M			1 000		1100	0.7		
	CRA8	All	12	75	13 000	33			23.3	708.0
		F			7 000		1552	10.3		
		M			7 000		1963	13.0		
	CRA9	All	7	224	65 000	64			52.8	47.0
		F			38 000		712	26.9		
		M			27 000		959	25.9		
Paua	PAU1		24	383	262 000	32	226		59.1	<1
	PAU2		63	2121	1 277 000	46	325		415.0	123
	PAU3		23	399	127 000	29	283		36.0	90
	PAU5		40	615	149 000	32	357		53.1	327
	PAU5B		9	128	29 000	53	-		0.0	-
	PAU5D		32	487	120 000	31	-		0.0	-
	PAU6[2]		4	66	23 000	74	357		8.0	1
	PAU7		11	110	50 000	48	317		15.8	265
Scallop	SCA(QMA1)		31	1403	634 000	34	110		69.8	-
	SCA(CORO)		18	766	257 000	101	117		30.1	55
	SCA(ENLD)		9	459	322 000	32	102		32.9	

Table 4: Harvest estimates for finfish and shellfish by fishstock or species where mean weights are not available

Species	Fishstock	Number of diarists harvesting the stock	Unweighted harvest reported by Diarists	Estimated harvest in numbers of fish or shellfish	*c.v.* (%)
Albacore tuna	ALB1	6	8	6 000	47
	ALB2	6	25	9 000	53
	ALB3	1	2	1 000	102
	ALB7	1	6	1 000	101
	ALB8	3	21	7 000	81
	ALB9	5	19	6 000	51
Blue maomao	BMA1	91	524	239 000	18
	BMA2	5	27	11 000	74
	BMA8	5	21	12 000	82
	BMA9	3	5	3 000	70
Bluenose	BNS1	11	26	11 000	38
	BNS2	7	20	6 000	60
Cockle	COC1	48	6 304	2 357 000	24
	COC2	9	1 615	549 000	84
	COC3	22	3 953	1 476 000	45
	COC5	3	100	18 000	60
	COC7	16	1 361	499 000	36
	COC8	3	32	6 000	94
	COC9	8	2 640	2 197 000	95
Elephant fish	ELE2	2	4	1 000	85
	ELE3	6	10	2 000	43
	ELE7	1	1	< 500	101
Garfish	GAR1	11	349	216 000	81
	GAR2	1	14	5 000	101
	GAR3	1	4	1 000	100
	GAR7	1	12	2 000	101
Hake	HAK1	2	5	2 000	100
Hoki	HOK1	4	8	2 000	74
Kelpfish	KEL1	6	19	22 000	83
	KEL2	3	7	1 000	63
	KEL3	2	3	1 000	78
	KEL7	1	9	2 000	102
	KEL9	1	1	1 000	101
Koheru	KOH1	20	108	42 000	38
	KOH2	1	8	2 000	101
	KOH7	1	5	1 000	101
	KOH8	1	2	< 500	101
	KOH9	1	2	< 500	100
Cats Eyes	LUN1	1	40	38 000	101
	LUN2	4	195	140 000	58
	LUN8	1	110	48 000	103
Dredge Oyster	OYS5	1	200	38 000	102
	OYS7	1	3	1 000	101
Paddle crab	PAD1	5	21	14 000	75
	PAD2	5	10	4 000	56
	PAD3	1	2	1 000	103
	PAD8	4	200	44 000	72
	PAD9	5	21	9 000	57

Table 4 - continued:

Pink maomao	PMA1	22	106	70 000	36
	PMA2	1	1	< 500	101
Parrotfish	POT1	28	60	20 000	21
	POT2	5	15	3 000	48
	POT3	22	75	28 000	43
	POT5	11	35	9 000	37
	POT7	18	68	25 000	35
	POT8	1	5	2 000	101
	POT9	1	1	< 500	101
Pacific Oyster	POY1	3	85	42 000	69
	POY7	1	12	6 000	102
Pipi	PPI1	130	15 371	6 848 000	15
	PPI2	19	1 822	861 000	36
	PPI3	17	741	316 000	37
	PPI5	4	170	88 000	78
	PPI7	9	718	176 000	63
	PPI8	17	1 372	809 000	71
	PPI9	10	1 994	1 695 000	97
Green Lipped Mussels	MSG1	37	2 527	1 308 000	30
	MSG2	10	16	8 000	42
	MSG3	9	456	402 000	74
	MSG5	3	5	1 000	62
	MSG7	4	10	3 000	53
	MSG8	14	813	242 000	44
	MSG9	3	95	25 000	87
Northern scorpionfish	RRC1	9	477	144 000	58
	RRC2	3	4	1 000	62
	RRC3	1	20	8 000	102
	RRC7	15	458	161 000	37
	RRC8	1	1	< 500	101
Blue Mussels	MSB5	6	124	66 000	55
	MSB7	11	415	188 000	44
Rock Oyster	ROY1	8	308	163 000	57
	ROY3	2	100	28 000	73
	ROY5	1	130	15 000	102
	ROY7	9	237	91 000	38
	ROY9	2	38	16 000	76
Red snapper	RSN1	24	152	79 000	32
	RSN2	1	2	1 000	102
	RSN7	1	4	1 000	101
	RSN9	2	6	1 000	77
Salmon	SAM2	1	1	< 500	101
	SAM3	23	53	22 000	37
	SAM5	4	48	9 000	62
	SAM7	1	1	1 000	102
Skipjack tuna	SKJ1	30	239	142 000	36
	SKJ2	5	29	7 000	50
	SKJ8	2	12	4 000	98
	SKJ9	5	21	6 000	65
Squid	SQU (QMA	11	18	13 000	59
Stargazer	STA2	1	5	1 000	101
	STA8	1	2	1 000	101

Table 4 - continued:

Spotty	STY1	10	68	39 000	53
	STY2	19	176	42 000	42
	STY3	12	43	10 000	34
	STY5	3	21	3 000	81
	STY7	39	251	78 000	21
	STY8	1	1	< 500	101
	STY9	1	1	< 500	102
Kina	SUR1	29	2 876	1 793 000	35
	SUR2	24	1 130	1 026 000	57
	SUR3	5	36	8 000	58
	SUR5	2	251	70 000	101
	SUR7	1	6	2 000	101
	SUR8	3	198	85 000	85
	SUR9	4	97	82 000	67
Trumpeter	TRU2	1	3	< 500	101
	TRU3	18	94	41 000	64
	TRU5	20	95	23 000	32
	TRU7	1	1	1 000	102
Tuatua	TUA1	75	8 288	3 290 000	20
	TUA2	6	266	85 000	49
	TUA3	1	200	164 000	101
	TUA5	1	150	17 000	102
	TUA7	2	69	20 000	72
	TUA8	3	354	64 000	69
	TUA9	9	1 988	1 382 000	48
Blue Warehou	WAR2	3	6	1 000	63
	WAR7	2	2	1 000	72

Table 5: Harvest estimates for finfish and shellfish by generic species and QMA

Species	Fishstock	Number of diarists harvesting the stock	Unweighted harvest reported by Diarists	Estimated harvest in numbers of fish or shellfish	*c.v.* (%)
Cod	QMA1	2	4	1 000	77
	QMA2	6	22	4 000	46
	QMA3	4	22	4 000	68
	QMA5	1	8	3 000	101
	QMA7	8	91	31 000	63
	QMA8	4	103	17 000	79
Dogfish	QMA1	1	1	< 500	101
	QMA2	4	12	4 000	72
	QMA3	18	166	53 000	44
	QMA5	1	1	< 500	101
	QMA7	4	8	2 000	54
	QMA8	1	1	< 500	101
	QMA9	3	17	5 000	87
Mackerel	QMA1	17	68	26 000	31
	QMA2	3	10	5 000	77
	QMA3	4	18	5 000	86
	QMA7	4	13	7 000	85
	QMA8	4	28	10 000	56
Mussels	QMA1	38	2 026	1 008 000	26
	QMA2	1	50	137 000	100
	QMA3	18	1 070	420 000	44
	QMA5	9	605	175 000	51
	QMA7	15	637	344 000	43
	QMA9	9	690	536 000	70
Mullet	QMA1	10	78	44 000	40
	QMA2	4	9	4 000	57
	QMA3	4	60	13 000	66
	QMA5	1	25	3 000	102
	QMA7	2	36	11 000	90
	QMA9	5	146	61 000	55
Oyster	QMA1	16	770	462 000	40
	QMA3	2	105	53 000	101
	QMA5	9	881	174 000	44
	QMA7	8	159	105 000	55
	QMA9	4	405	184 000	73
Perch	QMA2	1	5	1 000	102
	QMA3	8	91	20 000	58
	QMA5	1	4	2 000	101
	QMA7	5	15	4 000	55
Shark	QMA1	9	18	9 000	51
	QMA2	8	27	10 000	49
	QMA3	5	13	3 000	55
	QMA5	1	1	< 500	101
	QMA7	5	8	4 000	54
	QMA8	3	5	2 000	62
	QMA9	6	6	2 000	45
Skate	QMA1	1	1	1 000	102
	QMA2	5	6	1 000	46
	QMA3	3	3	1 000	60
	QMA5	1	1	< 500	102
	QMA7	1	1	< 500	101

~~Tierney, L.D., Kilner, A.R., Millar, R.B., Bradford, E. and Bell, J.D. 1997. Estimation of Recreational Harvests from 1991-92 to 1993-94. New Zealand Fisheries Assessment Research Document 97/15.~~

Appendix 1: Estimation of c.v.s for harvest estimates

Coefficients of variation (c.v.s) for recreational harvest estimates have been calculated using fairly simple formulae in previous national and regional recreational fishing surveys (see Bradford (1998a) and Tierney *et al.* (1997) for details). These formulae required specific assumptions about the distribution of catch among diarists, and in particular they assumed that diarists' harvests followed a Poisson distribution. That assumption was computationally convenient, since a Poisson random variable with mean *A* has a c.v. of $1/\sqrt{A}$, but it fails to hold for three reasons. First, for most fishstocks there are a substantial number of diarists who target the fishstock but do not catch any fish, generally many more than a Poisson distribution would predict. Secondly, there are often more small catches than expected. Thirdly, catch distributions are usually highly skewed with a small proportion of diarists reporting large harvests over the year, which would not occur in a Poisson distribution.

Because the catch distribution is usually much more dispersed than a Poisson distribution, calculating c.v.s based on a Poisson approximation is likely to result in c.v.s that are too small. Figure A1 below plots the distribution of annual diarist harvests from the 1999-2000 diary survey for the BCO3 fishstock, against a Poisson distribution with the same mean (with both variables on a log scale).

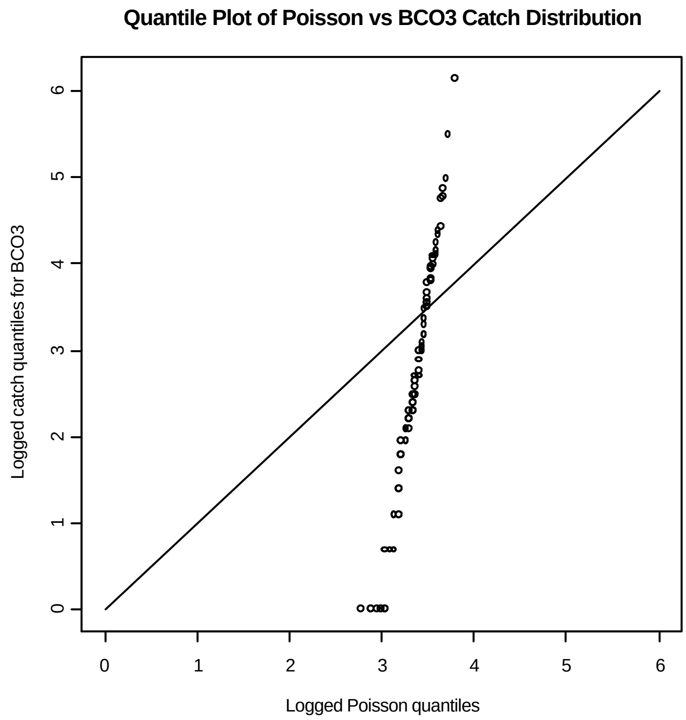

Figure A1: Quantile plot of the BCO3 harvest distribution versus a Poisson distribution with the same mean. (A log transformation of the quantiles has been used, and zero catches were not included in the plot or when calculating the mean.)

If the Poisson distribution was a good approximation to the diarist harvest distribution, the points would fall close to the continuous diagonal line shown in Figure A1. Instead the harvests are much more widely spread, differing from the Poisson values by more than a factor of ten at the extremes of the distribution.

Coefficients of variation for the 1999-2000 harvest estimates have therefore been calculated using the extended delete-a-group jack-knife (EDAGJK) method[1]. This involves repeatedly calculating harvest estimates on sub-samples carefully selected from the original data set to reflect the original sample design. These results are then combined to estimate the sampling errors (expressed in this report as c.v.s).

The EDAGJK method was applied to data from the National Readership Survey (NRS), the telephone recruitment survey and the diary survey[2]. Data from these surveys was combined to calculate the harvest estimates, so we need to incorporate the contribution of each survey to the c.v.s. The full data sets from the NRS and the telephone recruitment survey were initially divided into 120 variance groups, as described further below. Since the diarists are a subset of the recruitment survey respondents, this effectively selected 120 groups of diarists as well. Then 120 jack-knife replicates were obtained for each survey by removing each group from the full sample. Harvest estimates were then recalculated for each jack-knife replicate, after re-weighting the data as described by Reilly (2002). Finally the results for all the replicates were combined as suggested by Kott (1999), to estimate the c.v.s.

Because the NRS uses stratified systematic sampling of Nielsen Area Units as the first stage of its sample design, the 120 groups were formed for the NRS by systematically selecting every 120th Nielsen Area Unit from the NRS data set within each stratum[3]. (Nielsen Area Units are the primary sampling units, or PSUs, for the NRS.) The telephone recruitment survey was assumed to be a stratified random sample of households, with equal probabilities of selection within strata[4], so the 120 groups for this survey were created by randomly selecting 1/120 of the diarist recruitment survey respondents within each of the strata.

Replicate weights were created by first adjusting the inverse probability weights for the NRS respondents and recruitment survey respondents to create the jack-knife replicate weighting factors as described by Kott (1999). Briefly, if respondent k in PSU j within stratum h has initial weight w_{hjk}, the weight for this respondent in replicate r remained unchanged if no PSUs were in stratum h for variance group r. However if some PSUs did fall within stratum h for variance group r, the r-replicate weights for respondents in these PSUs were given by $w_{hjk}\left(1-(n_h-1)\sqrt{Z_h}\right)$, while the r-replicate weights for respondents in the other PSUs were $w_{hjk}\left(1+\sqrt{Z_h}\right)$. Here n_h is the number of PSUs selected in stratum h, and $Z_h = R/\left((R-1)n_h(n_h-1)\right)$ where R is the number of jack-knife replicates. The EDAGJK weighting factors are unusual for a jack-knife method in that they are non-zero even for observations outside the current jack-knife replicate.

All the remaining weighting steps described by Reilly (2002) were then repeated for each replicate, producing final replicate weights based on the initial replicate weights described above. This weighting process comprises several steps, including non-

[1] See Kott (1999) and Kott (2001) for details of this method.
[2] See Reilly (2002) and Ransom (2000) for more details of these surveys.
[3] The NRS sample is divided into 94 strata, which split the country into detailed regions (more detailed than regional council) and by the level of urbanisation within these regions.
[4] The "stratification" for this survey was implemented through centrally managed quotas for the number of interviews conducted in each of 40 area codes. These 40 areas divided the country by regional council, and also separated main urban areas from the rest of each region. Auckland, Wellington and Christchurch were each split into 4, 4, and 3 areas respectively.

response adjustment, post-stratification of data from each survey against reliable population totals, and an adjustment for fishers entering the fishery. Each set of final replicate weights is used to calculate harvest estimates for that replicate. The variation amongst these replicate harvest estimates enables the calculation of sampling variances for harvests, as described in Kott (1999). The c.v.s for the harvest estimates are then simply calculated from the harvest estimate and its sampling variance.

Using the EDAGJK approach for calculating the harvest c.v.s has several advantages, including that:

- there is no need for distributional assumptions, and in particular, the results will reflect the impact of skewed diarist harvest distributions;
- the EDAGJK requires many fewer replicates than a bootstrap approach; and
- EDAGJK can also handle designs with varying numbers of units selected from each stratum in the first stage of sampling, which is difficult to handle using BRR.

BRR and jack-knife techniques have been used to handle complex survey data for decades, with early literature dating back to the 1950's and 1960's. Both have been widely used in survey research since the 1970's, and have been adapted to handle a wide variety of sample designs. The bootstrap has only been used in survey research in relatively recent times, with articles first appearing in the late 1980's and early 1990's. The available literature on the bootstrap does not appear to cover as wide a range of sample designs, and in particular there does not seem to be any description available of how it could be used for systematic samples such as the NRS. This is another factor in favour of using a jack-knife method to calculate harvest c.v.s, instead of a bootstrap approach.

However there are some disadvantages to the EDAGJK approach, particularly regarding the NRS component of the data. The EDAGJK tends to overestimate variances if stratum sample sizes are less than five (which applies to 5% of areas selected for the NRS). It also assumes that the sample formed by leaving out a systematic sub-sample consisting of every 120th area (as was done to form the jack-knife replicates) will have similar statistical properties to the full NRS sample selected by systematic sampling. Some assumptions are always needed to calculate sampling variation based on a single systematic sample. The assumption made here seems reasonable given the slowly varying geographical trends observed in recreational marine fishing prevalence (highest in the upper North Island, dipping in the lower North Island, rising again in the upper South Island, and lowest in the far South). Finally, the c.v.s do not include any adjustment for sampling from a finite population, although 28% of Nielsen Area Units were used in the NRS measure of fishing prevalence. As a result, the c.v.s for fishing prevalence are probably over-estimated and may be approximately 15% higher than necessary[5].

However these issues affect only the prevalence component of the c.v.s. On its own, individual fishing prevalence has a jack-knife c.v. of only 3.3%, and this is relatively small compared to the much larger variability in harvest estimates arising from the skewed distribution of diarists' harvests. These issues are therefore believed to have a negligible impact on the overall harvest c.v.s.

[5] This figure of 15% is based on a simple formula for the effect of finite population sampling, which states that the sampling variance is reduced by the sampling fraction f (28% in this situation). C.v.'s are therefore multiplied by the factor $\sqrt{1-f}$. This formula is not accurate in general, but should give an indication of how large the effect might be in this situation.

It appears that the calculated c.v.s may underestimate the true c.v. if the calculation is based on a very low number of diarists who reported harvesting that fishstock. This is particularly evident where there is only one diarist who harvested that fishstock, but some caution is probably advisable when interpreting the reported c.v.s for fishstocks with fewer than five successful diarists reporting they harvested that stock.

There are conflicting suggestions in the literature regarding the number of groups that should be used for this type of jack-knife method. Although Kott suggests that 15 groups is sufficient for the delete-a-group jack-knife, Smith (2001) found that 40 groups were advisable for the New Zealand Household Labour Force Survey (and that using 120 groups gave even better results). The harvest c.v.s reported here are based on 120 groups. Calculations were also originally done using 30 groups, which gave similar results.