Fisheries and Oceans Canada | Pêches et Océans Canada

NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT | ***TRAC***

Comité d'évaluation des ressources transfrontalières | **Transboundary Resources Assessment Committee**

Document de référence 2010/08 | **Reference Document 2010/08**

Ne pas citer sans autorisation des auteurs | Not to be cited without permission of the authors

Determining Length-Based Calibration Factors for Cod, Haddock and Yellowtail Flounder

Elizabeth N. Brooks[1], Timothy J. Miller[1], Christopher M. Legault[1], Loretta O'Brien[1], Kirsten J. Clark[2], Stratis Gavaris[2], and Lou Van Eeckhaute[2]

[1] Northeast Fisheries Science Center
Woods Hole, MA 02536
USA

[2] Fisheries and Oceans Canada
St. Andrews, New Brunswick E5B 2L9
Canada

Ce document est disponible sur l'Internet à : | This document is available on the Internet at :
http://www.mar.dfo-mpo.gc.ca/science/TRAC/trac.html

TABLE OF CONTENTS

ABSTRACT

Data from a calibration study were analyzed for Atlantic cod (*Gadus morhua*), haddock (*Melanogrammus aeglefinus*), and yellowtail flounder (*Limanda ferruginea*) to determine appropriate factors to adjust survey data between the new *FSV Henry B. Bigelow* and the retired *FSV Albatross IV*. While some general protocols were in place, guidelines for approaching length-based calibration were lacking. A pre-TRAC working group approached the estimation of length based calibration factors for all three species together so that the criteria, and the considerations, that led to decisions on the method were consistent. After thorough evaluation of the data and comparisons of the proposed estimators, beta-binomial based estimates of length-specific calibration factors were estimated for cod, haddock, and yellowtail flounder. Data were examined for differences in seasonal (fall, spring) and site-specific calibration factors, but it was determined that all data could be pooled. Data were sparse at the smallest and the largest lengths, and calibration factors were estimated only for lengths greater than 20 cm in cod and yellowtail founder, and lengths greater than 18 cm in haddock. All lengths less than these cut-offs were assumed to have the same calibration factor. The best fit to the length data for all three species were segmented regressions where the right end point was estimated; all lengths greater than or equal to the right end point were assumed to have the same calibration factor. Numbers at length from *Bigelow* tows should be divided by the length specific calibration factors to obtain survey values on a scale that is consistent with *Albatross IV* tows.

RÉSUMÉ

On a procédé à l'analyse des données d'une étude d'étalonnage concernant la morue franche (*Gadus morhua*), l'aiglefin (*Melanogrammus aeglefinus*) et la limande à queue jaune (*Limanda ferruginea*) afin de déterminer quels facteurs de rajustement il convenait d'appliquer pour faire correspondre les données de relevé provenant du nouveau navire *Henry B. Bigelow* avec celles de l'*Albatross IV,* qui a été retiré du service. Quoique certains protocoles généraux aient été en place, il n'y avait pas de lignes directrices sur la façon de procéder à un étalonnage en fonction de la longueur. Un groupe de travail réuni préalablement à l'évaluation du CERT a procédé à l'estimation des facteurs d'étalonnage selon la longueur pour les trois espèces en même temps, si bien que les critères retenus et les éléments ayant mené aux décisions sur la méthode adoptée ont été cohérents. Après une évaluation approfondie des données et des comparaisons entre les estimateurs proposés, on a effectué des estimations bêta-binomiales des facteurs d'étalonnage selon la longueur pour la morue, l'aiglefin et la limande à queue jaune. Un examen des données en vue de déceler d'éventuelles différences de nature saisonnière (printemps, automne) ou dues au lieu a été effectué, mais il a été déterminé que toutes les données pouvaient être groupées. Les données étant peu nombreuses pour les longueurs les plus basses et les plus hautes, on a limité l'estimation de facteurs d'étalonnage aux longueurs de plus de 20 cm pour ce qui est de la morue et de la limande à queue jaune, et de plus de 18 cm pour l'aiglefin, et tenu pour acquis que le même facteur d'étalonnage s'appliquait à toutes les longueurs inférieures à ces seuils. La meilleure correspondance avec les données de longueur chez les trois espèces provenait de régressions segmentées avec estimation de l'extrémité droite; toutes les longueurs égales ou supérieures à l'extrémité droite étaient considérées comme ayant le même facteur d'étalonnage. Pour obtenir des résultats de relevé à une échelle compatible avec ceux des traits réalisés par l'*Albatross IV*, il faudrait diviser le nombre de poissons selon la longueur dans les traits du *Bigelow* par les facteurs d'étalonnage selon la longueur.

INTRODUCTION

The Northeast Fisheries Science Center (NEFSC) research bottom trawl survey has been conducted since 1963, and for most of those years the *FSV Albatross IV* was the survey vessel. In November 2008, the *Albatross IV* was decommissioned. The NEFSC research bottom trawl survey is now conducted by the *FSV Henry B Bigelow*. As vessel and gear characteristics are dramatically different between the two vessels (Table 1), a calibration study was conducted to allow estimation of a conversion factor for converting *Bigelow* observations to a scale that would be consistent with *Albatross IV* observations.

The calibration study took place during the spring and fall 2008 research bottom trawl surveys, and included additional site-specific tows in the summer of 2008. The calibration study consisted of paired tows at all of the stratified random stations in the fall and spring. The site-specific tows were added to try to obtain more tows for species that had limited positive observations (i.e. tows where the species was caught). Greater detail on the calibration study design can be found in NEFSC Vessel Calibration Working Group (2009).

After the calibration tows were completed, a comprehensive simulation study was conducted to evaluate properties of proposed estimators, and to provide estimates of preliminary calibration factors from the proposed methods (Miller et al. 2010).

A review panel was convened 11-14 August 2009 to review the analyses and make recommendations on the methodology for determining calibration coefficients (Walsh et al. 2009). Some general protocols were given by the panel, but the 2010 TRAC is the first instance within TRAC where these protocols will be applied. As such, a pre-TRAC working group met in Woods Hole, MA., during 7-8 April 2010, and carefully considered the review panel protocols with a critical examination of the available data (season-specific and site-specific tows), the available estimators (ratio and beta-binomial), and the performance of each estimator with respect to length-based conversions (separately evaluated for season-specific and site-specific trends).

The panel's general protocols were to use the beta-binomial model instead of the ratio estimator if the two methods were giving similar estimates, or if the ratio estimate was giving estimates that were greater than the beta-binomial model. For cases where the ratio estimator was giving estimates that were less than the beta-binomial model, then it was recommended to use the ratio estimates. More specifically, if the beta-binomial estimate was greater than the upper limit of the ratio estimator 95% confidence interval, then the ratio estimator was recommended. With regard to length-based conversions, the panel simply recommended that length be incorporated as a covariate where appropriate, without specifying criteria for 'appropriate.' The panel's advice for estimating seasonal calibration factors was limited to a consideration of the number of tows available: if there were less than 30 (+, +) tows (meaning at least one fish of the species was caught by both vessels at a given station) within the season, then seasonal calibration should not be attempted; if there were 30 -50 (+, +) tows within the season, then seasonal estimates could be considered if deemed necessary but should be interpreted carefully.

As the protocols were fairly general, there was ample scope for individual analysts to approach the decision about calibration differently. The pre-TRAC working group evaluated three species (cod, haddock, and yellowtail flounder) with the aim of developing a general and consistent approach that might serve to guide the estimation of calibration factors for other species.

METHODS

The number of (+, +) tows on both vessels was first examined to determine that there were sufficient observations for attempting the estimation of a calibration factor. Table 2 gives the total number of (+, +) tows, and the number of (+, +) tows for fall, spring, and site-specific stations. All three species met the criteria of having at least 30 (+, +) tows in total to attempt a constant annual calibration factor. Haddock and yellowtail had at least 30 (+, +) tows within each season and for the site-specific stations. Cod had slightly less than 30 (+, +) tows for the season-specific tows (22 in the fall, 27 in the spring), but had 45 (+, +) tows at the site-specific stations. Thus, all three species could, at a minimum, estimate a constant calibration factor.

Next, the pre-TRAC working group attempted to decide if length-based calibration factors were necessary. Lacking guidance from the review panel on how to approach this question, the group took the following steps in reaching a decision. First, the calibration factor at length point estimate and its 95% confidence interval from both the ratio of mean catches and the beta-binomial model were evaluated across all lengths for each species (Figures 1a-b, 2a-b, 3a-b). These data were examined to determine whether both estimators produced similar values. As this was generally the case, the working group followed the review panel's recommendation to use the beta-binomial estimator. The working group then tried to evaluate whether a constant or a length based calibration factor should be used. One of the differences between constant and length-specific calibration factors is that the estimated constant factor reflects where the most observations are. *If* a strong year class were present in 2008, then it is possible that the constant calibration factor might not be representative for all years. Alternatively, if the two vessels catch fish of different sizes at different rates, then the constant factor would smooth through those differences, and would again be pulled towards lengths where the most data were, regardless of year class effects. Thus, in determining if length-based calibration factors were needed, the group tried to determine if the length based point estimates reflected a consistent trend at length, rather than simply noise around a constant factor. For all three species, the length-specific calibration estimates suggested the ratio of *Bigelow* to *Albatross IV* catch was higher for small fish than large fish (Figures 1b, 2b, 3b). Consequently, the working group decided to proceed with length-based rather than constant calibration factors.

As part of the evaluation of the length-specific trends, the group noted that at the very smallest size classes, and sometimes for the very largest size classes, the data were extremely limited and the estimates were poorly determined. Preliminary analysis of different functional relationships for length-based calibration factors resulted in unreasonable estimates due to the paucity of data, especially when extrapolated beyond the range of observed lengths. Therefore, the working group decided that a minimum length class (X1) should be selected for estimating the length based calibration factors, and that all lengths below X1 would be assumed to have the same calibration factor. One consideration for determining X1 was to look at the ratio at length of the two proposed calibration estimators (the ratio estimator and the beta-binomial estimator). In general, these two estimators seemed to perform similarly when there were sufficient observations and when the observations were not dominated by large tows by one vessel compared to the other. Thus, one might expect that data are 'sufficient' where the ratio of these two estimators is near 1.0. In addition, the group also looked at the beta-binomial estimated factors at length to examine the width of the estimated 95% CI—for some of the smallest, poorly sampled lengths, the range spanned by the CI was unrealistic. After examining the ratio of length-specific estimates (ratio estimate: beta-binomial estimate, Figure 1c, 2c, 3c), and considering both the trend and uncertainty of the calibration estimates relative to length (Figure 1b, 2b, 3b), the group agreed that the minimum length class used in the estimation would be 20 cm for cod, 18 cm for haddock, and 20 cm for yellowtail haddock.

Given the decision to estimate length-specific calibration factors, and a decision on the smallest length class to use in the estimation, the working group next evaluated whether season-specific length-based estimators were warranted, and further evaluated whether the site-specific stations could be pooled with the seasonal stations. To address this issue, the working group took an approach similar to that taken to establish X1 (the smallest size for which to estimate a conversion). Namely, the group examined ratios of the beta-binomial length-specific estimates for each season and for the site-specific stations, and if the ratio of those length-based estimates was close to 1.0 then it was concluded that the data could be pooled. Comparing the ratio of beta-binomial length-based estimates for sizes X1 and larger, the group decided that for all three species, all data could be pooled (i.e., all season-specific and site-specific data; Figure 1d, 2d, 3d). The noisiest of the ratios occurred for cod, which had the lowest number of season-and site-specific data (indeed, cod fell short of the review panels recommended cut-off of 30 (+, +) observations, suggesting that seasonal factors should not be attempted).

The final decision to be made with regard to the length-specific calibration factors was the functional form to fit. Two general forms were recommended for fitting to data for lengths ≥ X1: a logistic, and a "segmented-regression." For the segmented-regression, the group initially specified a length, X2>X1, where it appeared that the calibration factor was more or less constant. The segmented regression is thus a straight line with negative slope between X1 and X2, reflecting a decreasing calibration factor at length, and a constant calibration factor assumed for all lengths ≥X2. The group selected values for X2 were: 40 cm for cod, 60 cm for haddock, and 29 cm for yellowtail flounder. A second fit of the segmented regression was performed where X2 was estimated to minimize the likelihood (see Appendix). It should be noted that the length classes were pooled differently for each species to be consistent with length bins used in age-length keys. For cod, data were pooled to 3 cm length bins; for haddock, data were pooled by 2 cm length bin; yellowtail flounder data were not pooled (the observed 1 cm length classes were used). Although the decision was made to estimate length-specific calibration factors, the estimated value and fit for a constant calibration factor are provided for readers interested in comparing results of the two approaches.

RESULTS

Cod

The working group decisions were to pool all data sets (seasonal and site-specific) and to estimate two segmented-regressions and a logistic fit to data beginning at length X1=20 cm. Table 3 gives Akaike's Information Criterion (AIC) (Akaike 1973) for each of the segmented regression fits; the logistic form was inestimable. The AIC is lower for the segmented regression where X2 was estimated, which suggests that it is a better fit to the data. The model estimate for X2 was 53.377 cm. The predicted segmented regression at length is shown in Figure 1e. The higher constant calibration value for sizes < X1 is due to forcing the linear decline to the fixed X2 (40 cm); when X2 was estimated it was greater than the group's fixed guess, which allowed for a less steep trend to be fit. The length based calibration factors for cod corresponding to the segmented regression with X2 estimated are given in Table 4.

Haddock

The working group decisions were to pool all data sets (seasonal and site-specific) and to estimate two segmented-regressions and a logistic fit to data beginning at length X1=18 cm. Table 3 gives the AIC for each of the segmented regression fits; the logistic regression was inestimable. The AIC is lower for the segmented regression where X2 was estimated, which suggests that it is a better fit to the data. The model estimate for X2 was 50.052 cm. The predicted segmented regression at length is shown in Figure 2e. The two segmented regression lines are similar, but when X2 was estimated, it was less than the group's fixed guess (60 cm), so the constant calibration for lengths <X1 is slightly higher. The length based calibration factors for haddock corresponding to the segmented regression with X2 estimated are given in Table 4.

Yellowtail Flounder

The working group decisions were to pool all data sets (seasonal and site-specific) and to estimate two "segmented-regressions" and a logistic fit to data beginning at length X1=20 cm. Table 3 gives the AIC for each of the segmented regression fits and the logistic fit (the logistic regression was estimable in this case). The AIC is very similar for all 3 models. Between the two segmented regressions, the estimated X2 is very close to the groups' best guess to fix X2 (28 vs 29 cm, respectively). The logistic regression has a knife-edge transition between about 26 and 27 cm. The working group felt it was difficult to justify why the calibration factor should decrease by about a third between 1 cm lengths. Therefore, given the similarity in the AIC, the group decided to accept the segmented regression with X2 estimated (Figure 3e). The length based calibration factors for yellowtail flounder corresponding to the segmented regression with X2 estimated are given in Table 4.

DISCUSSION

Length-based calibration factors were estimated for three groundfish species: cod, haddock, and yellowtail flounder. The value estimated for the constant factor corresponded to the length classes where there were more fish sampled, whereas the length based factor accounted for the differential selection at length of the *Henry Bigelow* relative to the *Albatross IV* (in general, more smaller fish were caught). We note that in all cases, the constant calibration fit was worse (higher AIC) than the length based fits.

These calibrations were done for numbers per tow; however, no attempt was made to calibrate biomass per tow. Although Miller et al. (2010) provide a single calibration factor for biomass, it is based on the product of the constant numbers based calibration factor and the ratio of observed mean fish weight per vessel in 2008. If the length composition changes from year to year, it is possible that the mean weight would change as well. Therefore, the applicability of the single biomass conversion in subsequent years is uncertain. Several approaches for biomass calibration are presently being explored, however the analyses are still under investigation and are not presented in this document. The calibration of indices in number will suffice for the TRAC stock assessments which use indices of numbers at age per tow. However, the graphical representation of catch per tow in weight should be considered preliminary while the additional investigations are conducted.

REFERENCES

Akaike, H. 1973. Information theory as an extension of the maximum likelihood principle; pp. 267-281. In: B.N. Petrov and F. Csaki, Editors. Second International Symposium on Information Theory. Akademiai Kiado, Budapest, Hungary.

NEFSC Vessel Calibration Working Group. 2009. Design and field data collection to compare the relative catchabilities of multispecies bottom trawl surveys conducted on the NOAA

Ship *Albatross IV* and the *FSV Henry B. Bigelow*. Working Paper, Northeast Fisheries Science Center, Woods Hole, MA.

Miller, T.J., Das, C., Miller, A.S., Lucey, S.M., Legault, C.M., Brown, R.W., and Rago, P.J. 2010. Estimation of *Albatross IV* to *Henry B. Bigelow* calibration factors. U.S. Depart. of Commerce, Northeast Fisheries Science Center Ref. Doc. 10-05; 233 p.

Walsh, S., Smith, S., and Kaiser, M. 2009. Independent panel review of the NMFS Vessel Calibration Analyses for *FSV Henry B Bigelow* and *R/V Albatross IV*, August 11-14, 2009. Chair's Consensus Report. Northeast Fisheries Science Center. Woods Hole, MA.

Table 1. Vessel and gear differences between the *FSV Henry B Bigelow* and *FSV Albatross IV*. The information in this table was contributed by Russ Brown, Phil Politis, and Pete Chase of the NEFSC Ecosystems Surveys Branch.

Measure	**FSV Henry B Bigelow**	**FSV Albatross IV**
Tow speed	3.0 knots SOG	3.8 knots SOG
Tow duration	20min	30 mins
Headrope height	3.5-4m	1-2m
Ground gear (cookies, rock hoppers, etc.)	Rockhopper Sweep Total Length-25.5m Center- 8.9m length, 16" rockhoppers. Wings- 8.2m each 14" rockhoppers	Roller Sweep Total Length-24.5m Center-5m length, 16" rollers. Wings- 9.75m each, 4" cookies.
Mesh size	Poly webbing Forward Portion of trawl (jibs, upper and lower wing ends, 1st and 2nd side panels, 1st bottom belly)12cm,4mm Square aft to codend:6cm, 2.5mm Codend: 12cm, 4mm dbl. Codend Liner: 2.54cm, knotless	Nylon webbing Body of trawl= 12.7cm Codend- 11.5cm Liner (codend and aft portion of top belly)-1.27cm knotless
Net design	4 Seam, 3 Bridle	Yankee 36 (recent years)
Other comments	Wing End to Door distance= 36.5m	Wing End to Door Distance= 9m

Table 2. Number of tows where both vessels caught a given species (i.e., +, + tows).

Species	**Total (+,+)**	**Fall (+,+)**	**Spring (+,+)**	**Site-specific (+,+)**
Cod	94	22	27	45
Haddock	160	42	33	85
Yellowtail	143	38	39	66

Table 3. Results of beta-binomial length-specific calibration factors fit with segmented regression or logistic models, and the constant calibration factor for comparison.

	Segmented Regression with X1, X2 Fixed		**Segmented Regression with X1 Fixed, X2 Estimated**		**Logistic Regression**	**Constant**
Species	X1, X2 (cm)	AIC	X1, X2 (cm)	AIC	AIC	AIC
Cod	20, 40	1686.294	20, 53.377	1672.741	Not estimable	1710.848
Haddock	18, 60	7183.683	18, 50.052	7170.793	Not estimable	7246.557
Yellowtail	20, 29	4375.679	20, 28.000	4376.902	4377.079	4381.687

Table 4. Length based calibration factors from the segmented regression with X2 estimated. Numbers at length from *Henry Bigelow* tows should be divided by the calibration factor in the corresponding length bin. It is recommended to use all 6 digits to the right of the decimal point. For comparison, constant factors are: Cod (2.156782), Haddock (1.447101), Yellowtail (2.020736).

	Calibration Factor		
Length	Cod	Haddock	Yellowtail
≤18	5.723743	2.626169	3.857302
19	5.723743	2.580551	3.857302
20	5.723743	2.534933	3.857302
21	5.600243	2.489315	3.621597
22	5.476743	2.443697	3.385892
23	5.353243	2.398079	3.150187
24	5.229743	2.352462	2.914482
25	5.106243	2.306844	2.678777
26	4.982743	2.261226	2.443072
27	4.859243	2.215608	2.207367
28	4.735743	2.169990	1.971662
29	4.612243	2.124372	1.971657
30	4.488743	2.078754	1.971657
31	4.365243	2.033136	1.971657
32	4.241743	1.987518	1.971657
33	4.118243	1.941900	1.971657
34	3.994743	1.896283	1.971657
35	3.871243	1.850665	1.971657
36	3.747743	1.805047	1.971657
37	3.624243	1.759429	1.971657
38	3.500743	1.713811	1.971657
39	3.377243	1.668193	1.971657
40	3.253743	1.622575	1.971657
41	3.130243	1.576957	1.971657
42	3.006743	1.531339	1.971657
43	2.883243	1.485721	1.971657
44	2.759743	1.440104	1.971657
45	2.636243	1.394486	1.971657
46	2.512743	1.348868	1.971657
47	2.389243	1.303250	1.971657
48	2.265743	1.257632	1.971657
49	2.142243	1.212014	1.971657
50	2.018743	1.166396	1.971657
51	1.895243	1.163990	1.971657
52	1.771743	1.163990	1.971657
53	1.648243	1.163990	1.971657
54	1.601603	1.163990	1.971657
≥55	1.601603	1.163990	1.971657

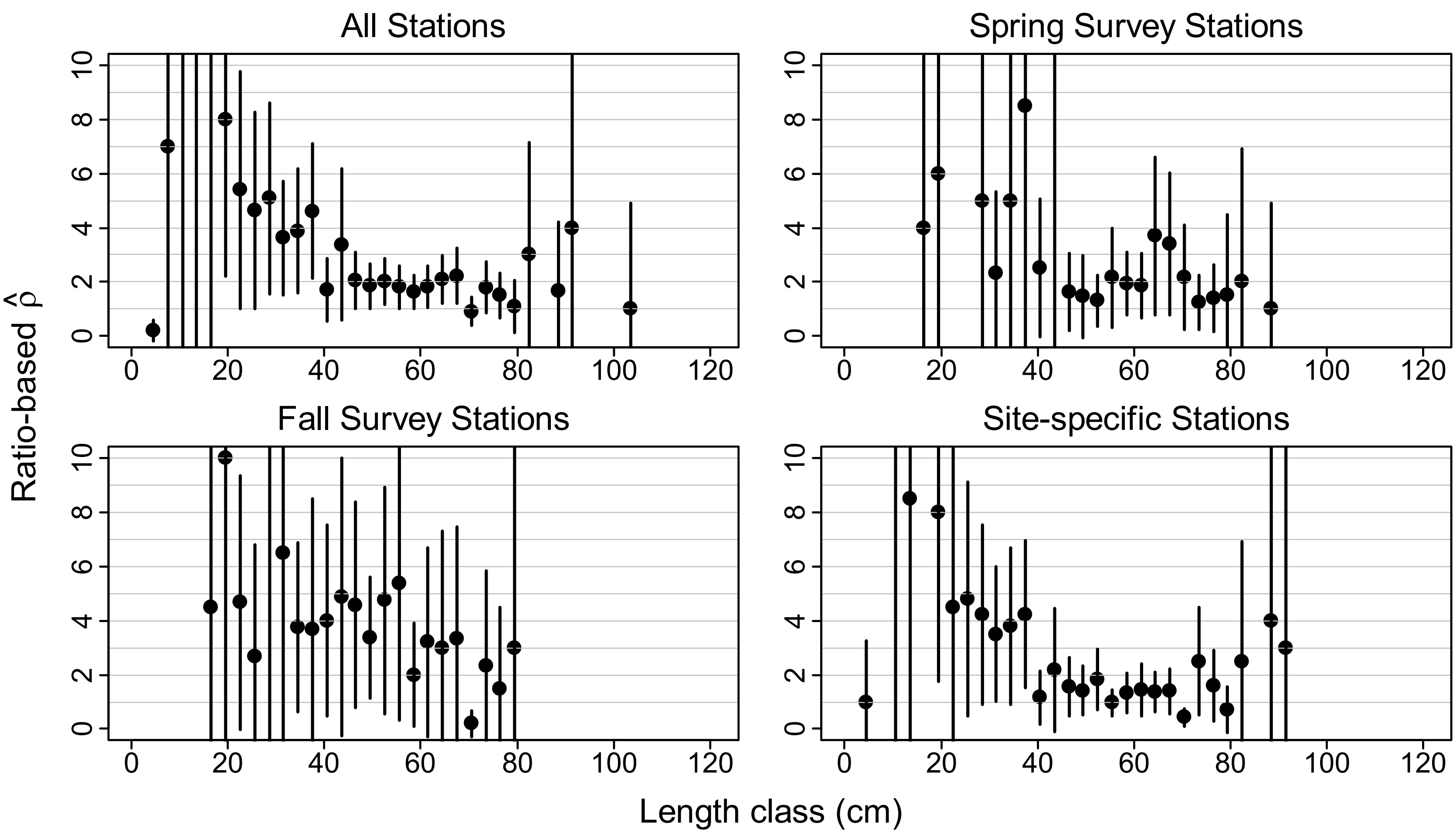

Figure 1a. Calibration factor estimates for **Atlantic cod** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of stations based on ratios of mean catches. Lengths are binned in 3 cm intervals.

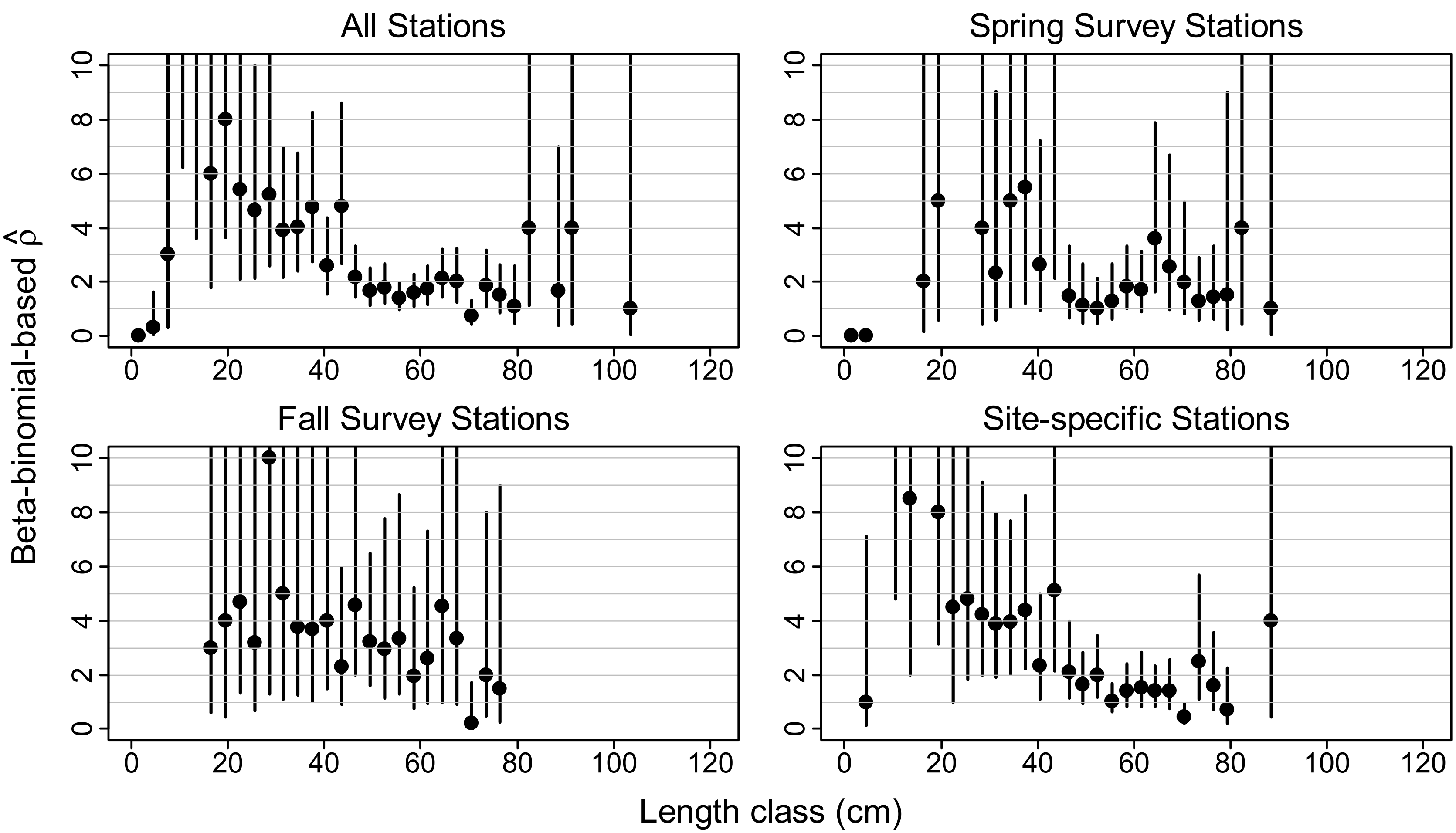

Figure 1b. Calibration factor estimates for **Atlantic cod** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of stations based on a beta-binomial model. Lengths are binned in 3 cm intervals.

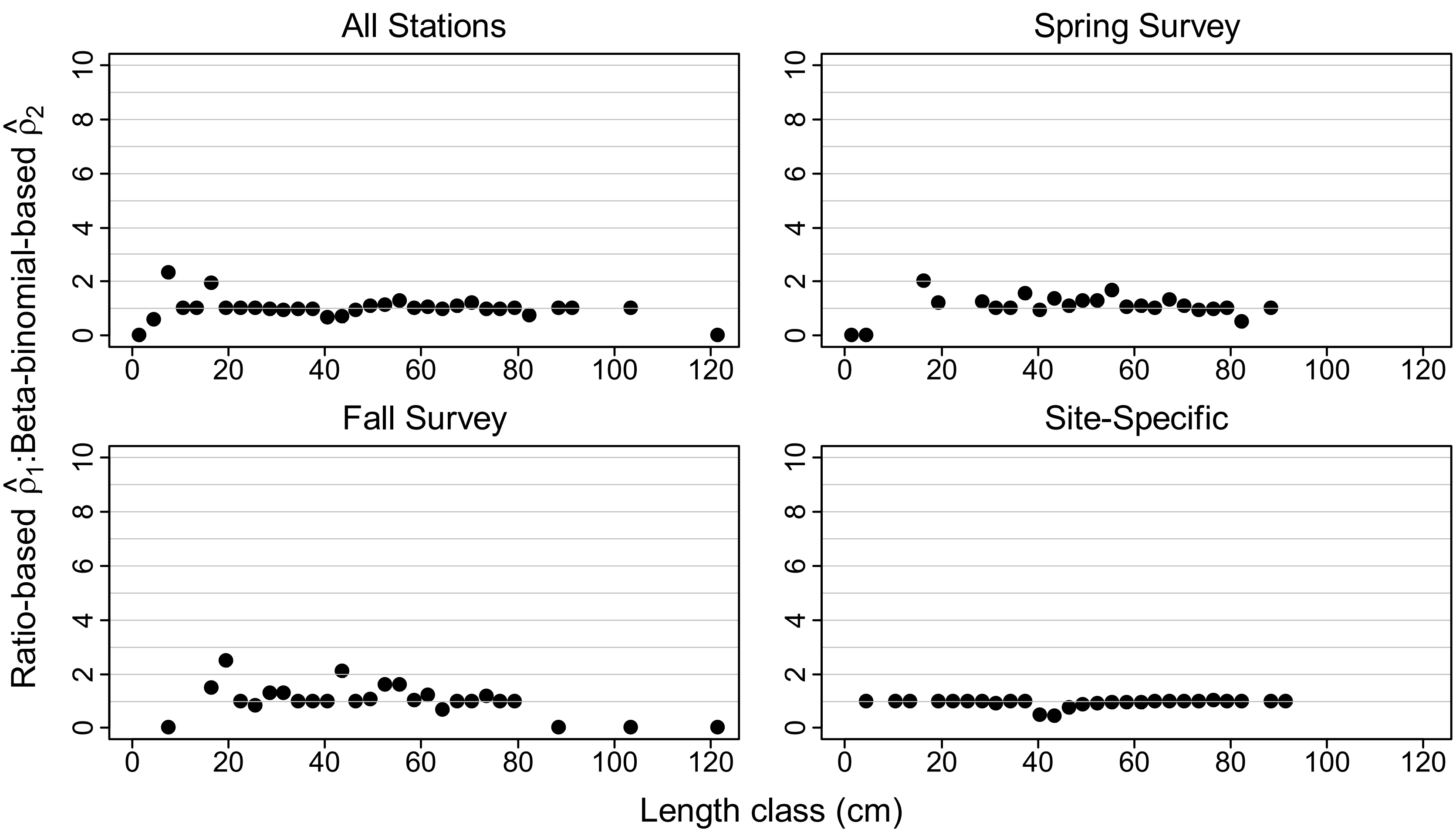

Figure 1c. Ratios between two estimators of length specific calibration factors (the estimator based on ratios of mean catches and the beta-binomial-model) for **Atlantic cod** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of data. Lengths are binned in 3 cm intervals.

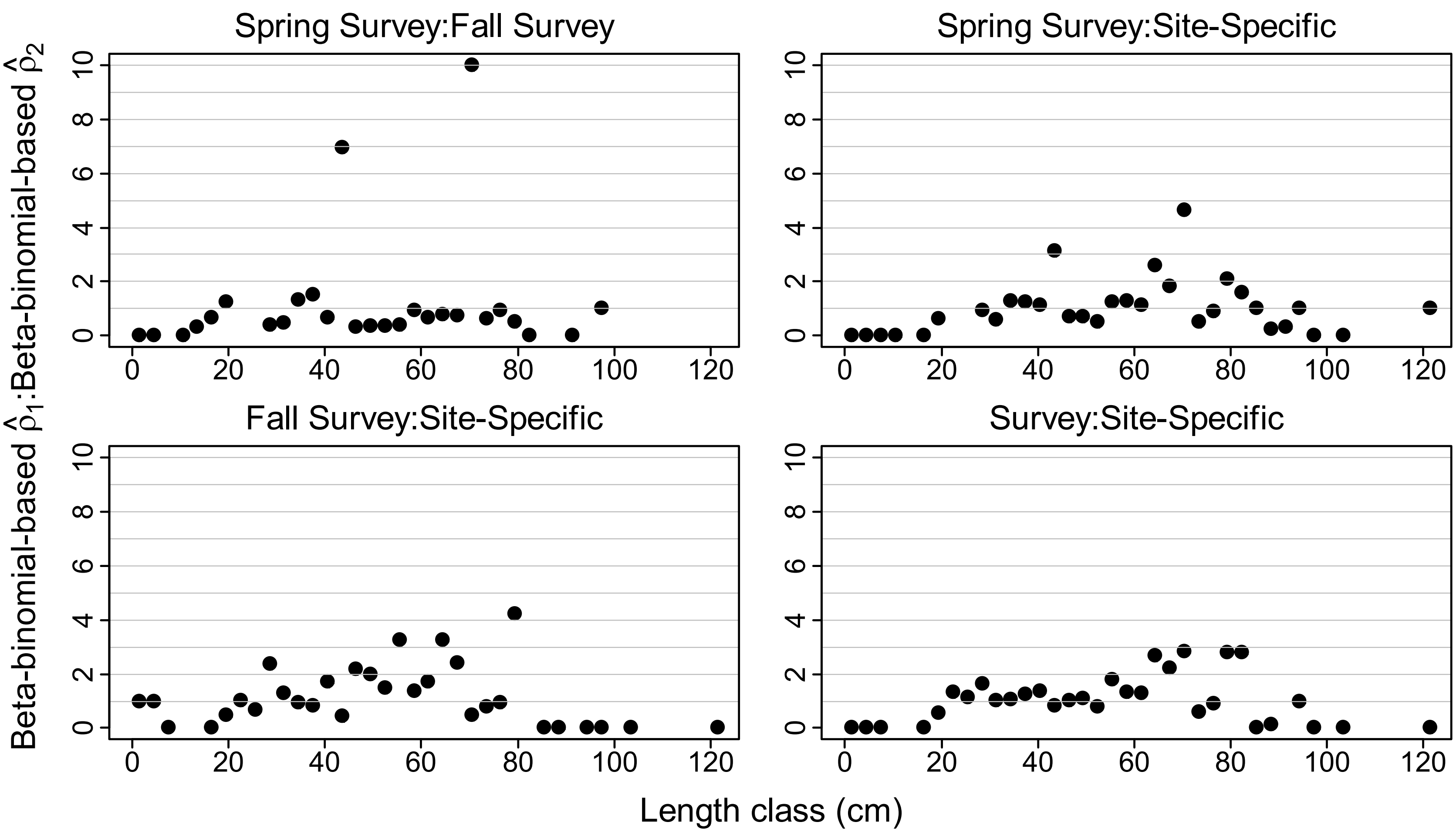

Figure 1d. Ratios of calibration factor estimates for **Atlantic cod** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of data based on a beta-binomial model. Lengths are binned in 3 cm intervals.

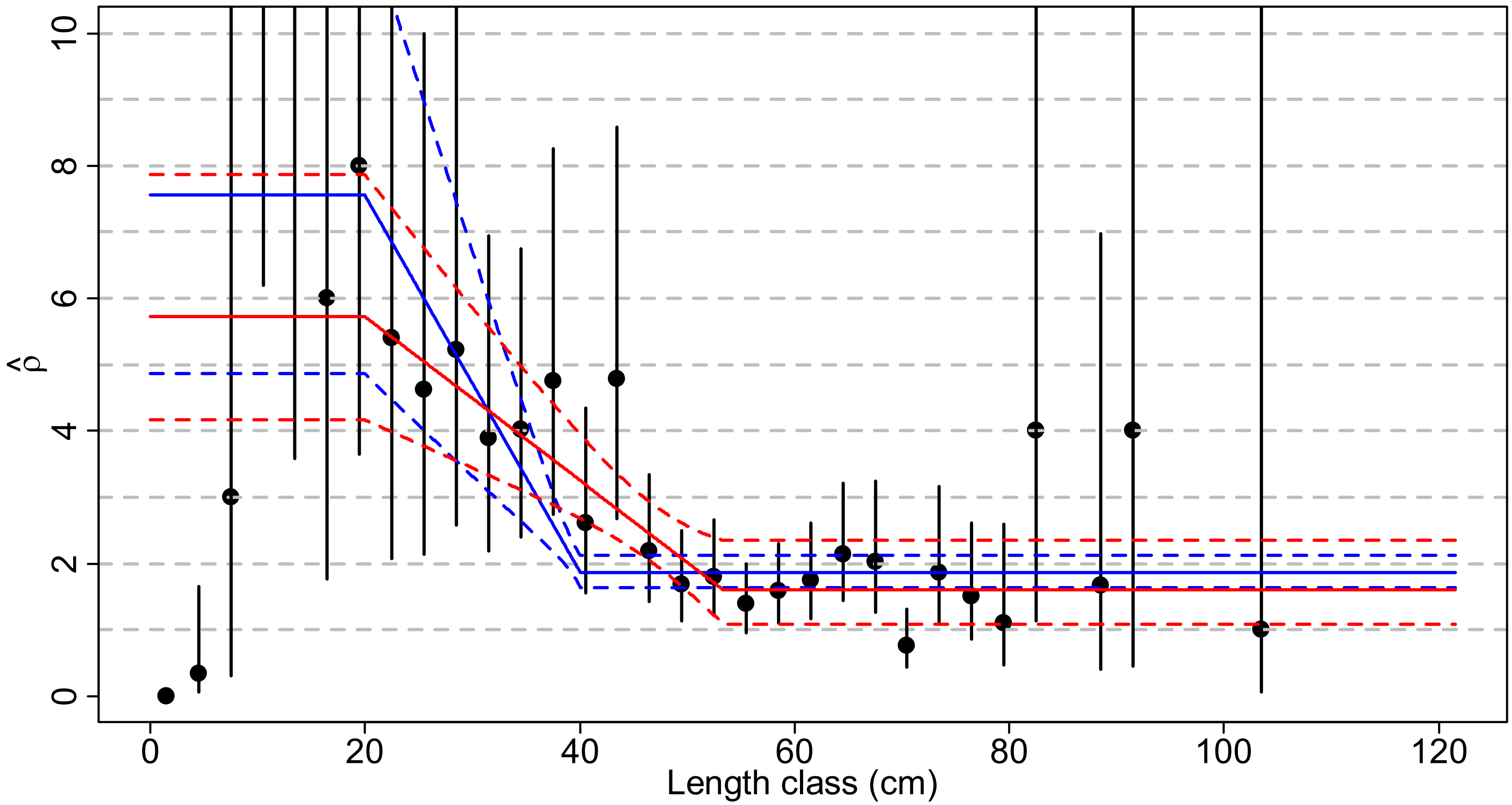

Figure 1e. Beta-binomial-based estimates of calibration factors and corresponding 95% confidence intervals by length class (3 cm bins) for **Atlantic cod**. The black points and vertical bars represent results where different calibration factors are estimated for each length class. The blue lines represent results from a segmented regression model where the two points connecting the segments are known (20 and 40 cm) and the red lines represent results from a segmented regression model where the first point (20 cm) is known but the second is estimated. Segmented regression fits are based on data from fish ≥20 cm.

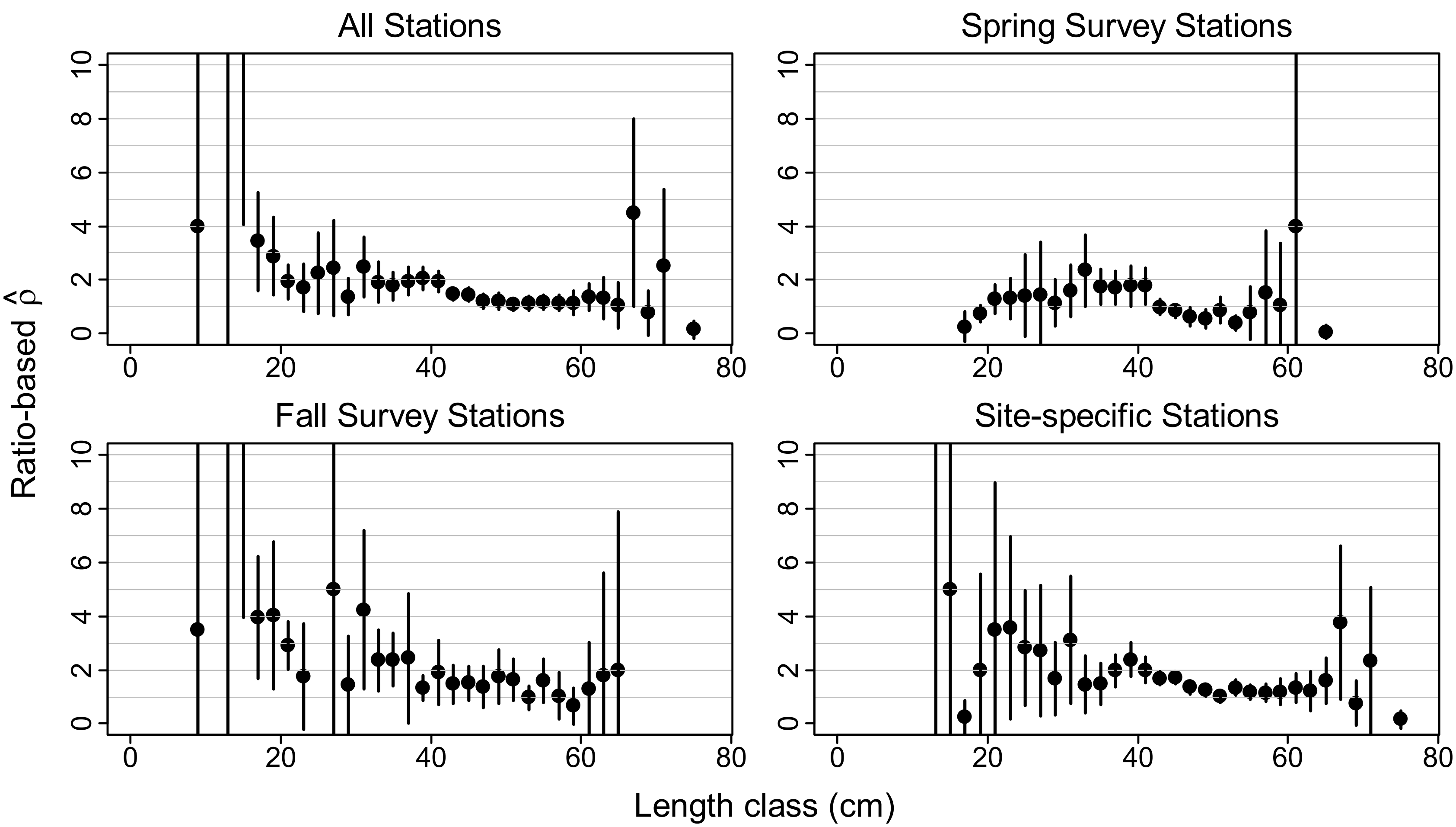

Figure 2a. Calibration factor estimates for **haddock** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of stations based on ratios of mean catches. Lengths are binned in 2 cm intervals.

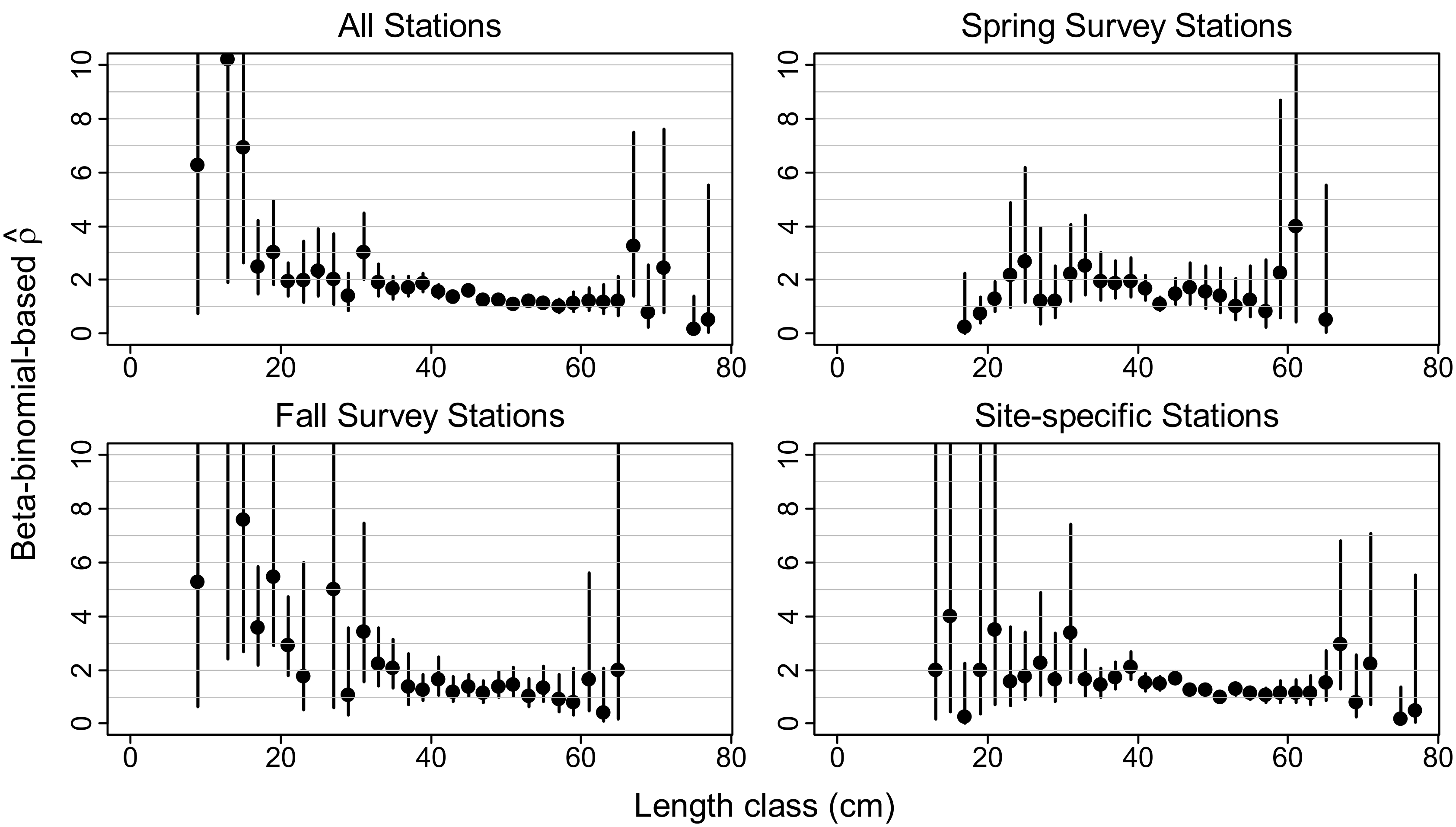

Figure 2b. Calibration factor estimates for **haddock** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of stations based on a beta-binomial model. Lengths are binned in 2 cm intervals.

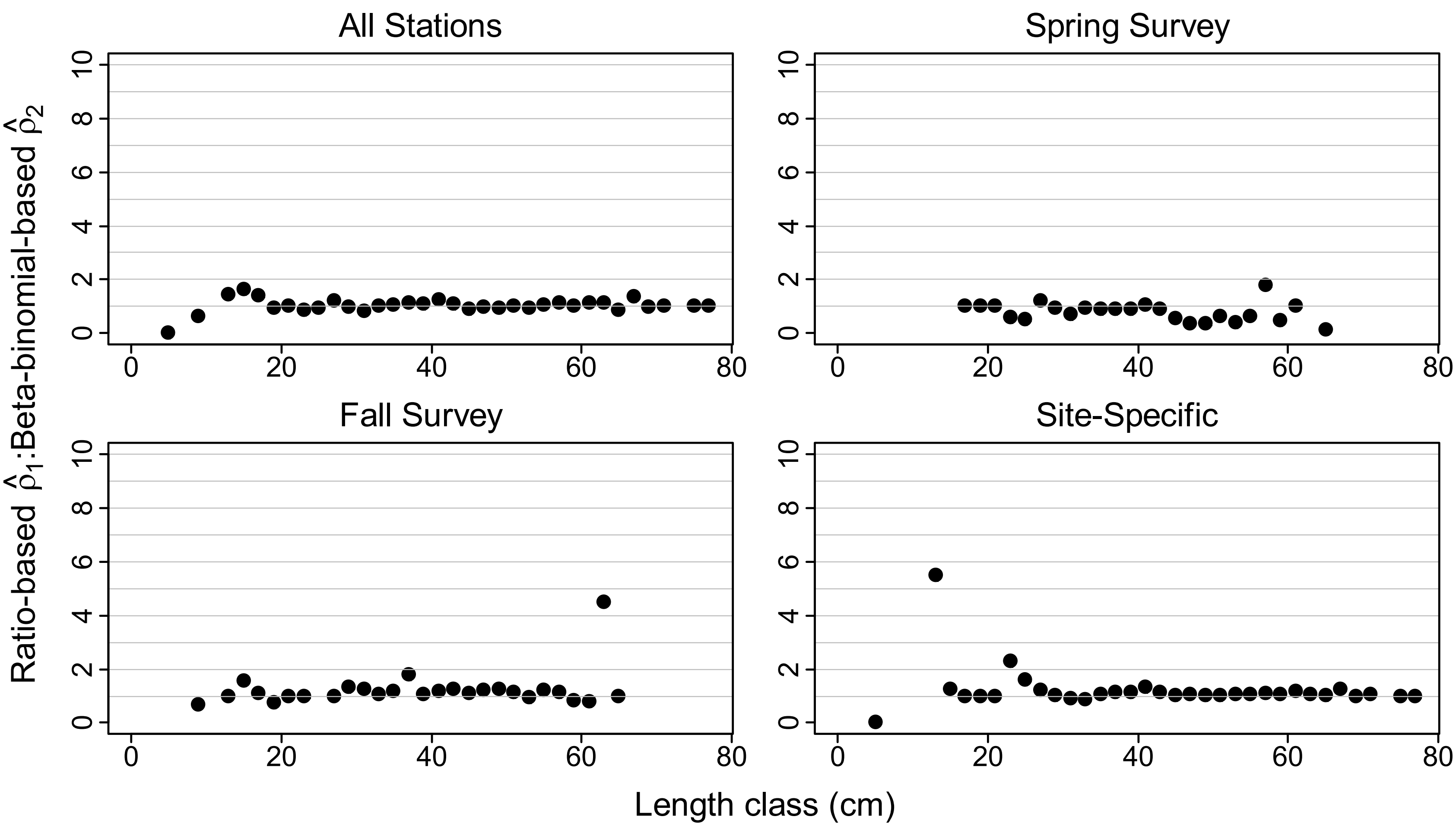

Figure 2c. Ratios between two estimators of length specific calibration factors (the estimator based on ratios of mean catches and the beta-binomial-model) for **haddock** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of data. Lengths are binned in 2 cm intervals.

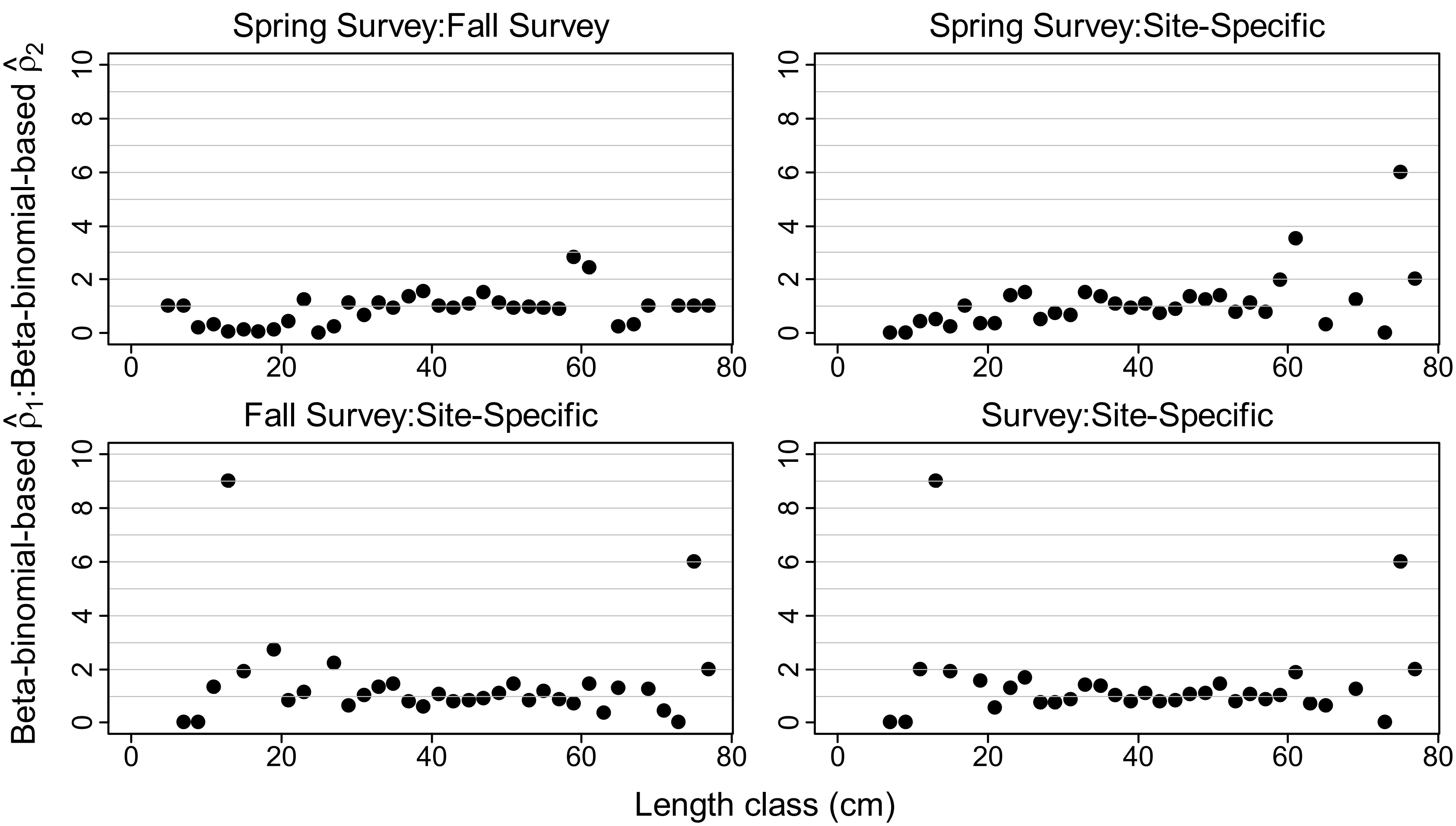

Figure 2d. Ratios of calibration factor estimates for **haddock** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of data based on a beta-binomial model. Lengths are binned in 2 cm intervals.

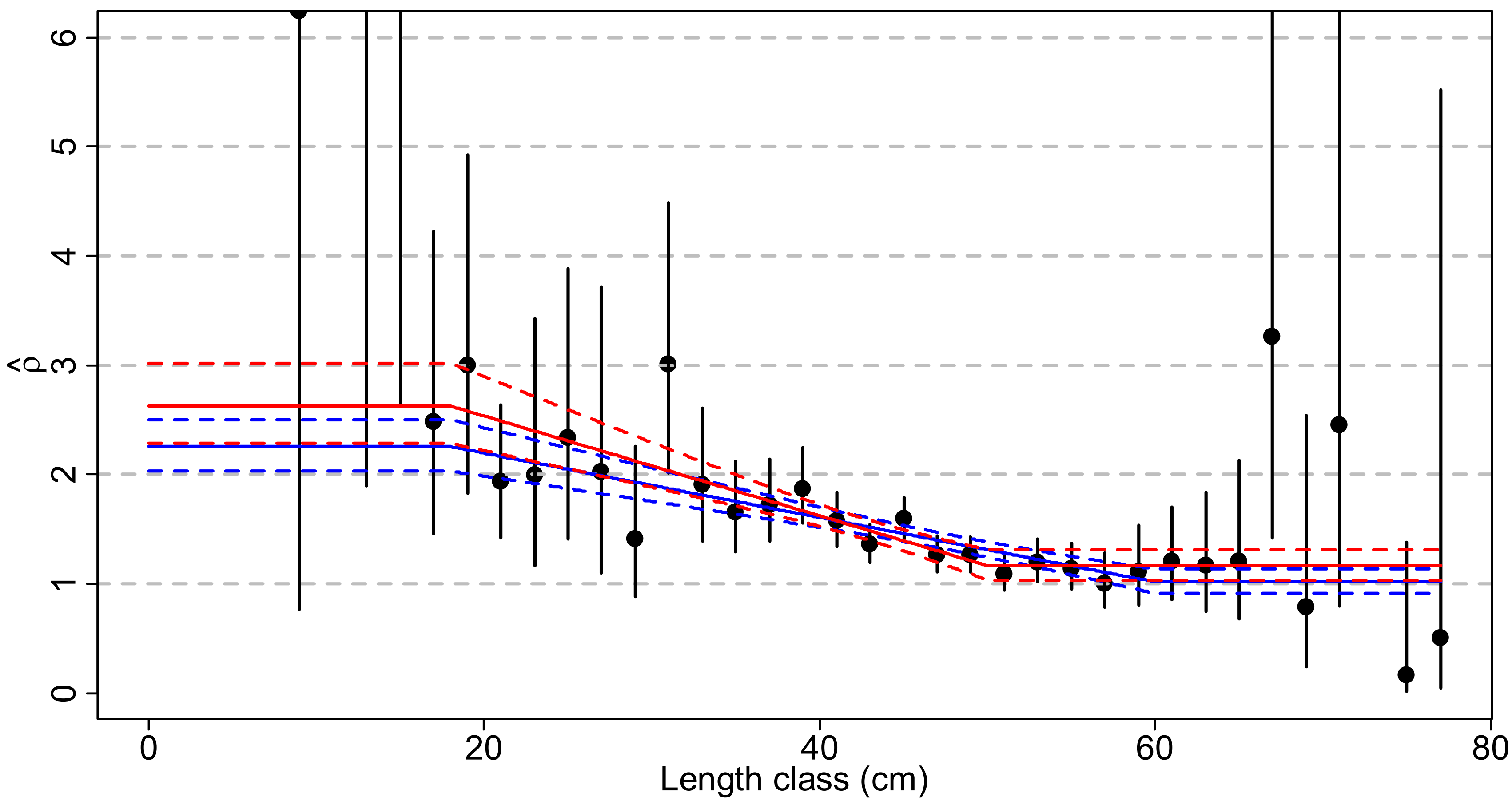

Figure 2e. Beta-binomial based estimates of calibration factors and corresponding 95% confidence intervals by length class (2 cm bins) for **haddock**. The black points and vertical bars represent results where different calibration factors are estimated for each length class. The blue lines represent results from a segmented regression model where the two points connecting the segments are known (18 and 60 cm) and the red lines represent results from a segmented regression model where the first point (18 cm) is known but the second is estimated. Segmented regression fits are based on data from fish ≥18 cm.

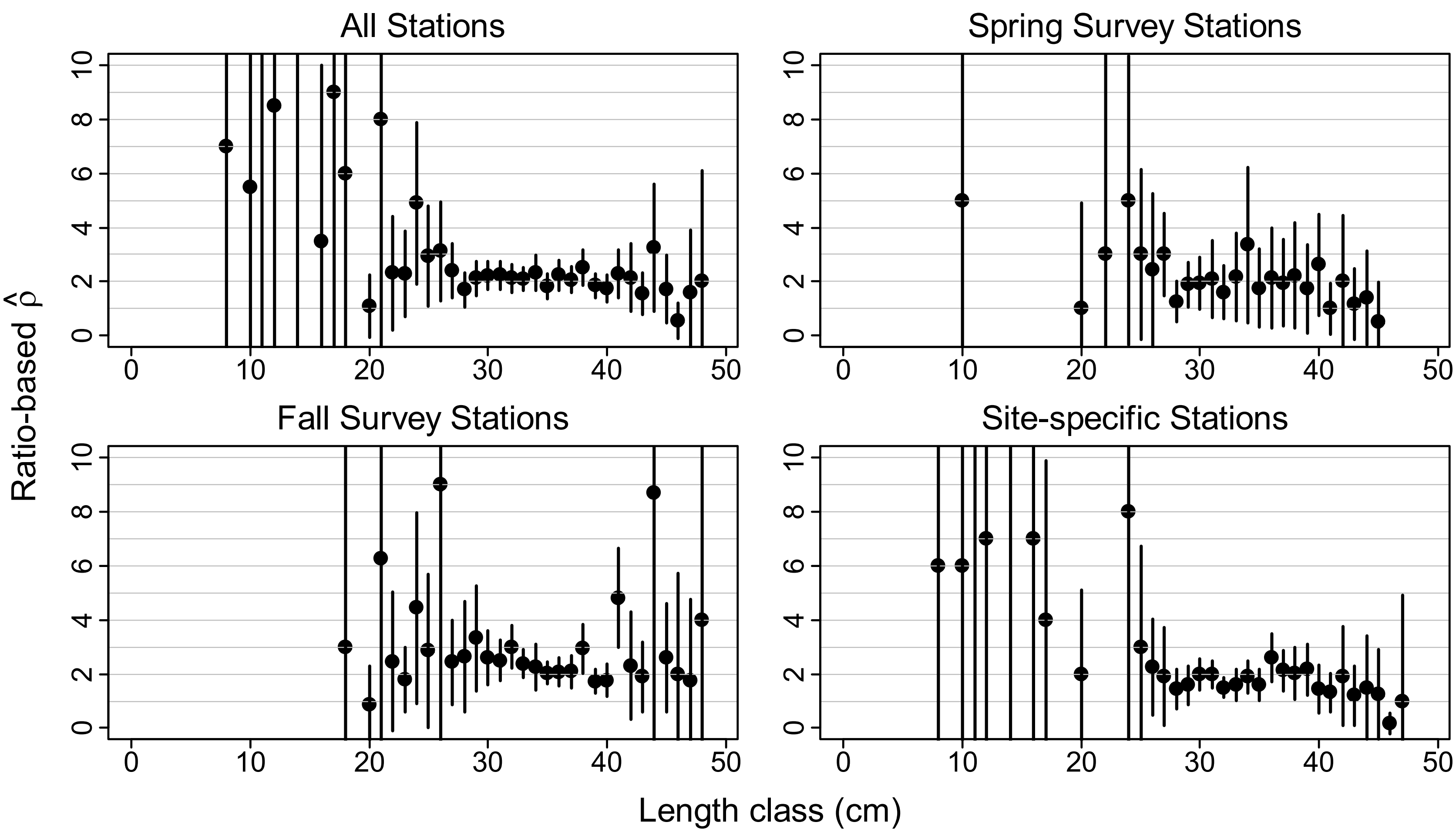

Figure 3a. Calibration factor estimates for **yellowtail flounder** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of stations based on ratios of mean catches. Lengths are binned in 1 cm intervals.

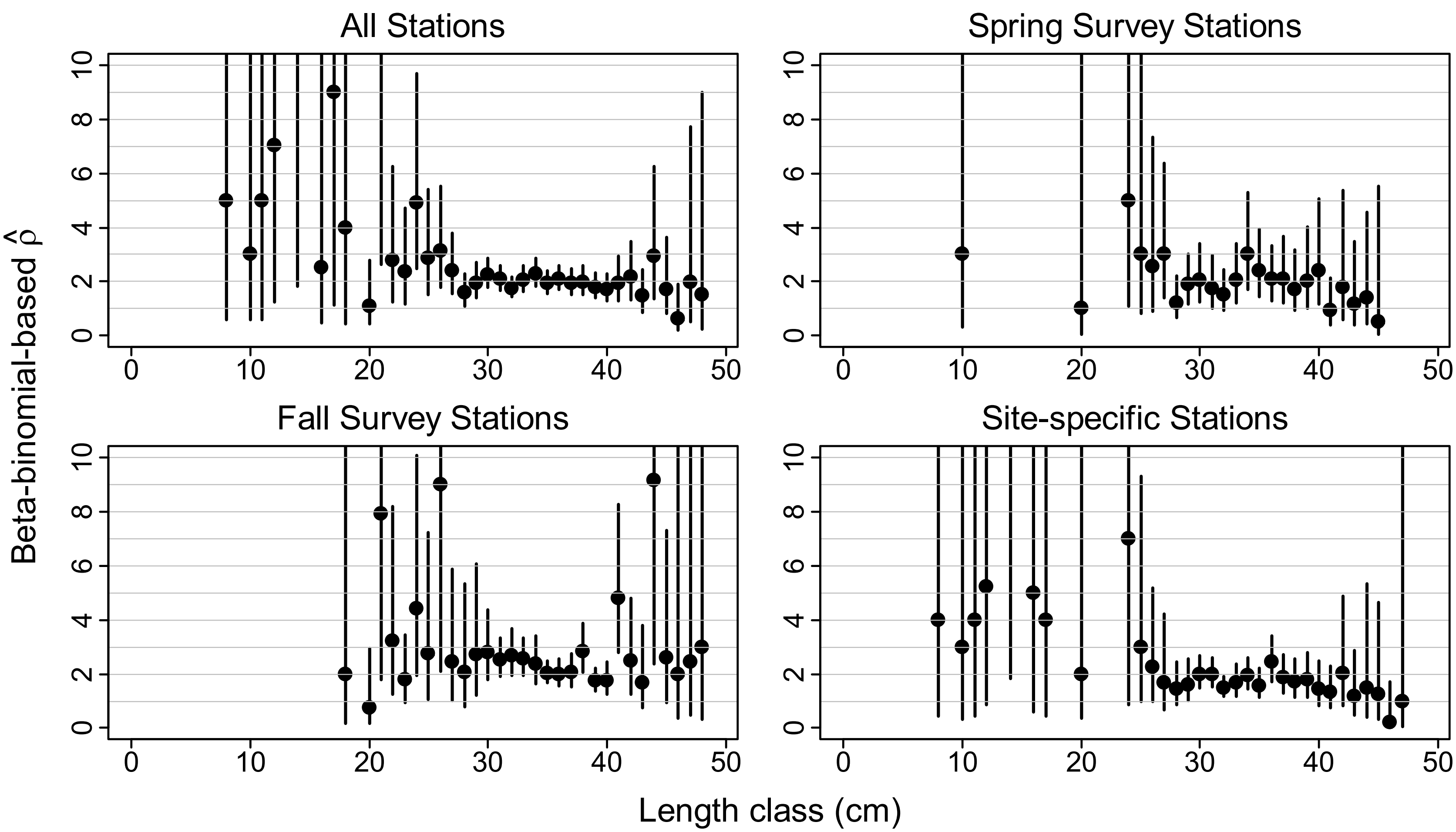

Figure 3b. Calibration factor estimates for **yellowtail flounder** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of stations based on a beta-binomial model. Lengths are binned in 1 cm intervals.

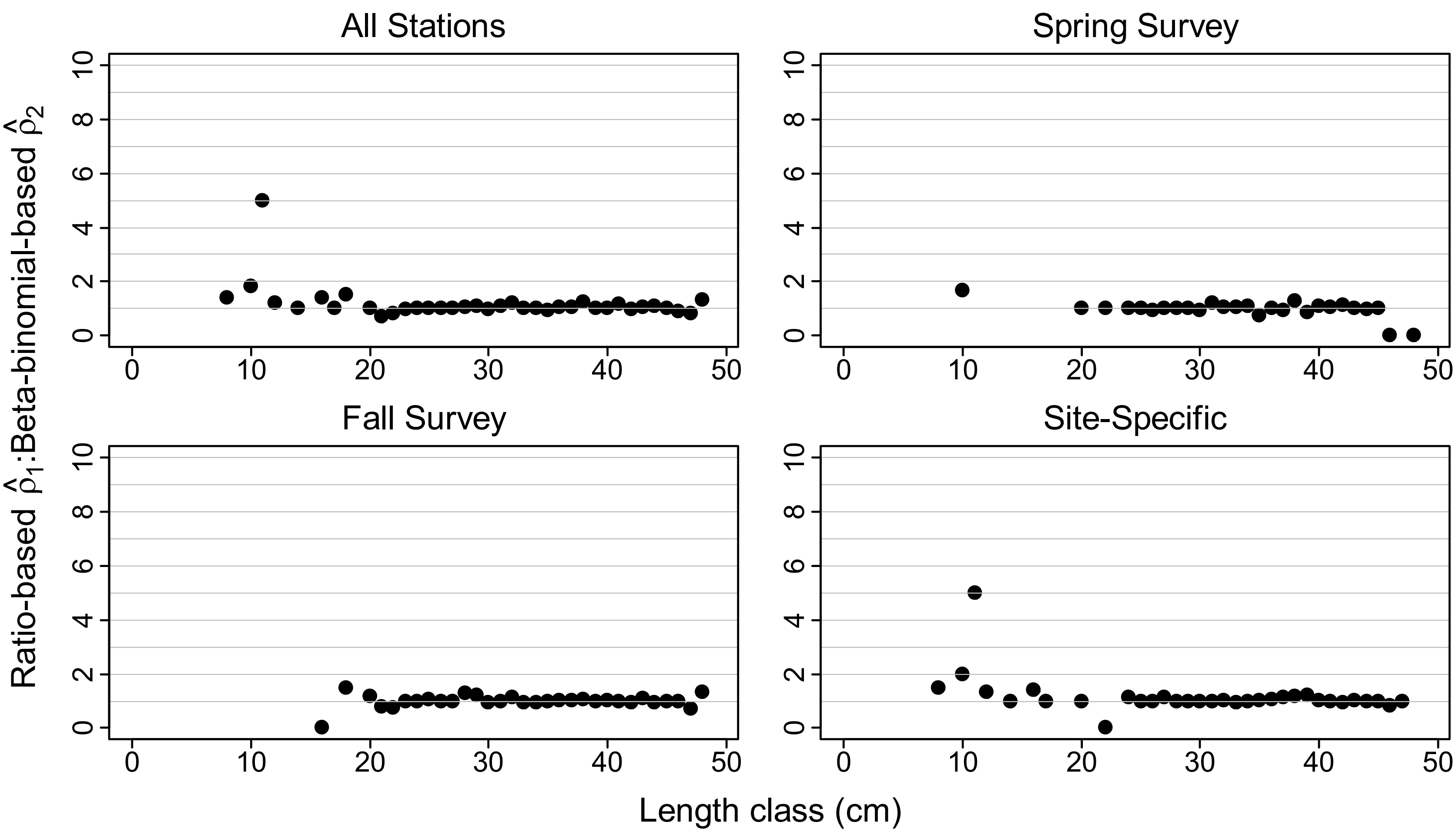

Figure 3c. Ratios between two estimators of length specific calibration factors (the estimator based on ratios of mean catches and the beta-binomial-model) for **yellowtail flounder** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of data. Lengths are binned in 1 cm intervals.

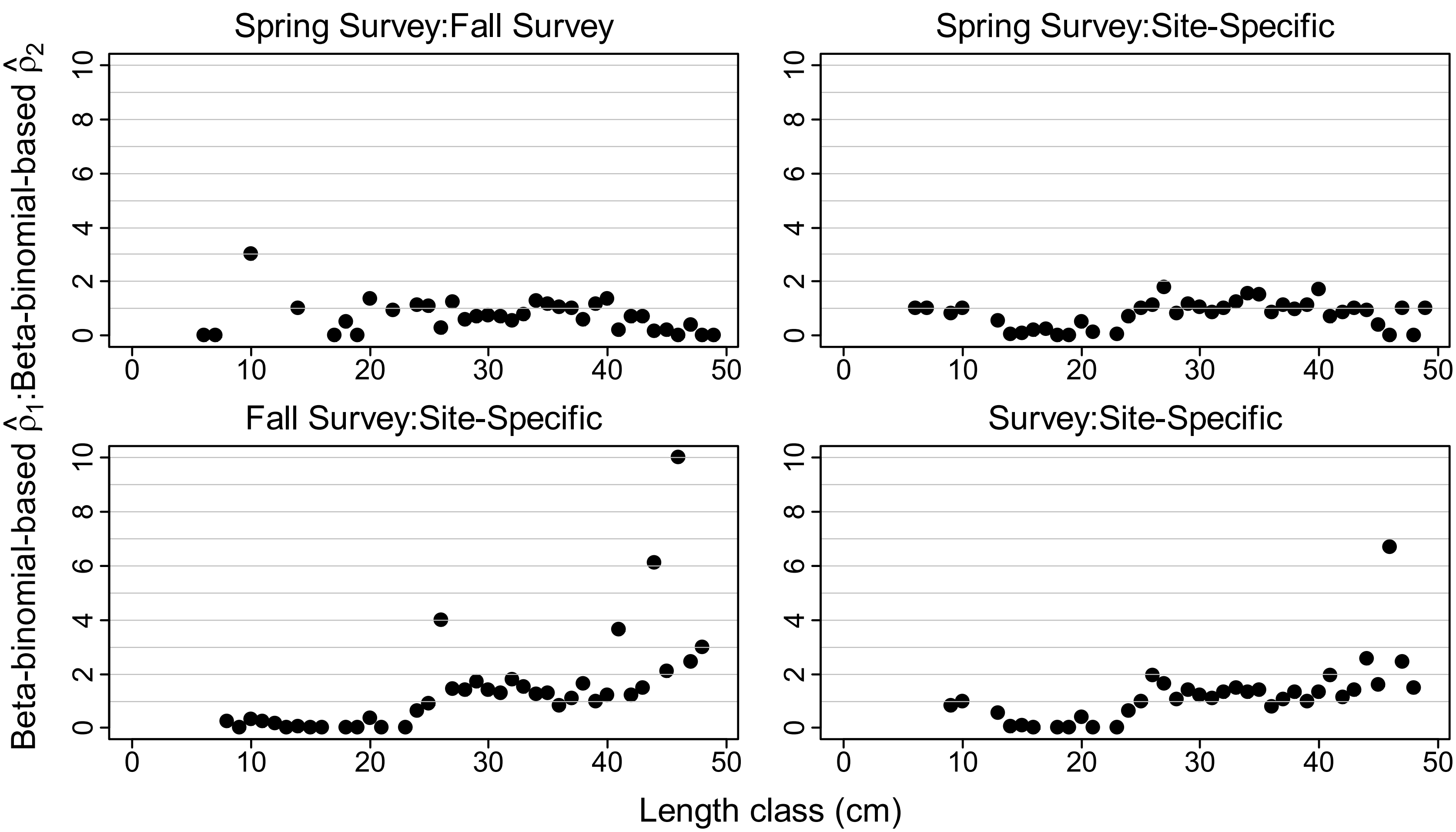

Figure 3d. Ratios of calibration factor estimates for **yellowtail flounder** catches from the *Bigelow* and *Albatross IV* by length bin in different sets of data based on a beta-binomial model. Lengths are binned in 1 cm intervals.

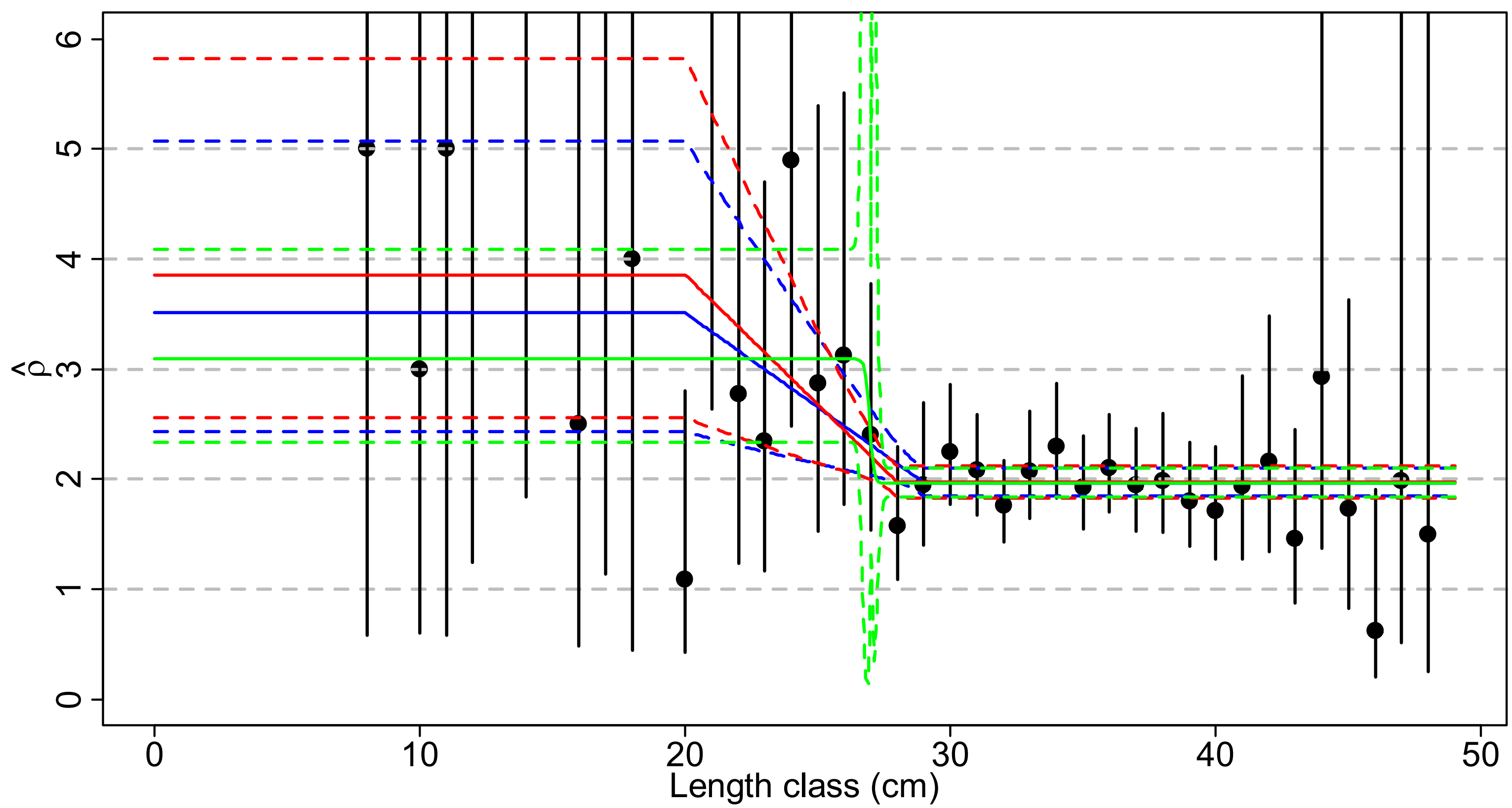

Figure 3e. Beta-binomial based estimates of calibration factors and corresponding 95% confidence intervals by length class (1 cm bins) for **yellowtail flounder**. The black points and vertical bars represent results where different calibration factors are estimated for each length class. The blue lines represent results from a segmented regression model where the two points connecting the segments are known (20 and 29 cm), the red lines represent results from a segmented regression model where the first point (20 cm) is known but the second is estimated, and the green lines represent results from the logistic model. Segmented-regression and logistic model fits are based on data from fish ≥20 cm.

APPENDIX

In the "segmented-regression" model, the model for the calibration factor ρ as a function of length is:

$$\rho(l) = \begin{cases} e^{\gamma} + e^{\delta} & \text{if } l \le X_1 \\ e^{\gamma} + e^{\delta} \dfrac{X_1 + e^{\chi} - l}{e^{\chi}} & \text{if } X_1 < l \le X_1 + e^{\chi} \\ e^{\gamma} & \text{if } l > X_1 + e^{\chi} \end{cases}.$$

Between the two points X_1 and $X_2 \equiv X_1 + e^{\chi}$ the calibration factor is decreasing with length and it is constant when length is less than X_1 or greater than X_2. When the second point X_2 is not assumed known, the parameters χ , γ, and δ are estimated whereas only γ and δ are estimated when X_2 is fixed.

In the logistic model, the calibration factor ρ as a function of length is:

$$\rho(l) = e^{\gamma} + \frac{e^{\alpha}}{1 + e^{-(\beta_0 + \beta_1 l)}}.$$

This form allows the calibration factors at the smallest sizes to asymptote at a value greater than zero and the difference between the minimum and maximum ρ to be different than 1.

Letting the full set of calibration factor parameters be θ (which depends on which of the above models is used), the beta-binomial likelihood we maximized is:

$$L(\theta,\phi) = \prod_{i=1}^{S} \prod_{j=1}^{M} \frac{\mathrm{Beta}\left(a_j + N_{Bij}, b_j + N_{Aij}\right)}{\mathrm{Beta}\left(a_j, b_j\right)} \binom{N_{Aij} + N_{Bij}}{N_{Bij}}.$$

In the likelihood d above, $\mathrm{Beta}()$ is the beta function, N_{Aij} and N_{Bij} are the numbers caught at station i in length class j by the *Albatross IV* and *Bigelow*, respectively. The calibration factor is $a_j = \rho(l_j \mid \theta)\phi$, $b_j = \phi / (1 + \rho(l_j \mid \theta))$ and ϕ is an estimated dispersion parameter.