Evaluating the Observer Effect for the Northeast U.S. Groundfish Fishery

Chad Demarest

updated April 18, 2019

Groundfish Plan Development Team

– DRAFT –

This information is distributed solely for the purpose of pre-dissemination review. It has not been formally disseminated by NOAA. It has no official status with the agency and does not represent final agency determination or policy.

Introduction

The commercial component of the Northeast U.S. Multispecies fishery comprises 20 individual fish stocks and 2 management units[1]. Of these, commercial fisherman are allocated quota for 15 stocks, leaving 5 for which retention is prohibited. Fishing quota is allocated to approximately 1,000 permits and actively fished by around 200 participating commercial vessels (NEFMC 2017). The majority of the commercial fishery for groundfish (~98% of landings) is managed under the sector system whereby individual vessel owners pool stock-level quota into any one of 21 sectors, each operating as a collective, pooling the quota and allocating it to individual member fisherman. Quota for allocated stocks may be traded between sectors. Trades are remunerated in three ways: single stock trades for a given amount of money (fish-for-cash), pounds of multiple stocks traded for a single value (basket trades), and pounds of quota for one stock traded for pounds of quota of another stock with no money exchanged (swaps). All regulated groundfish species have a prescribed minimum fish size and regulations prohibit retaining fish below that size, and discarding fish above it.

Observers are deployed on participating vessels to estimate discarded catch for each of the 20 fish stocks on each trip. Observer coverage levels vary but in general observers have been onboard trips accounting for between 10-35% of all trips taken in any given fishing year. Discards on observed trips are calculated by dividing the sum of observed stock-level discards on observed tows by the total amount of retained catch on these tows. For trips with no observer coverage, discards are estimated by applying the annualized observed discard rate (stock-level discards divided by the sum of kept catch), stratified by broad stock area, sector and fishing gear. Discards count against a sector's quota after adjusting for gear and stock-based discard mortality rates. Vessels are assessed estimated discards on unobserved trips based on their strata, regardless of whether or not an individual species was reported on that trip. Sectors must have adequate quota reserves for all species in a given stock area prior to any member vessels fishing in that area. Observers have also been the primary source of enforcement for mandatory retention regulations.

As observer coverage only represents a fraction of the total fishing activity in the sector component of the commercial groundfish fishery, obvious questions arise: Does data generated on observed fishing trips reflect the activities of the whole fleet? Are estimates generated from these data unbiased? Bias may be induced by either a deployment effect, where the assignment of observers to vessels is non-random, or an observer effect, where the fishing activities on observed trips vary in detectable ways from those on unobserved trips (Benoit and Allard 2009). These two effects, deployment and observer, may act separately or in combination

[1] George's Bank is divided into a "west" component for which haddock and cod stocks are assessed exclusively by NOAA fisheries, and an "east" component for which these stocks together with yellowtail flounder are jointly assessed with the Canadian Department of Fisheries and Oceans under a trans-boundary management agreement.

to render data collected by on board observers biased. This paper focuses specifically on one component of the the latter effect: do individual vessels alter their behavior in response to the presence of an observer?

Fisherman may alter their fishing behavior when carrying an observer for any one of at least five reasons: (1) people may act differently as a response to simply being watched, an established phenomena referred to as the Hawthorne Effect (McCambridge et al. 2018); (2) fisherman may not want to impart their individual discarding preferences on the other members of their sector, an effect driven primarily by within-strata fishing practice heterogeneity; (3) observers incur costs associated with slower fish processing and handling times, carrying extra food, and general inconvenience, all of which may incentivize fisherman to make shorter trips when observers are on board; (4) catch of undersized fish varies across space and fishing in areas and at times where undersized fish are relatively less abundant may minimize discard rates, though at the cost of reduced revenues; and (5) binding quota constraints impart strong economic incentives to discard legal-sized fish when an observer is not on board and to avoid these stocks in the presence of an observer, again presumably at a cost in terms of reduced trip revenues.

Methods

This paper uses an exact matching method to determine if vessel performance along several metrics vary in a detectable way when an observer is on board, and when one is not. Following a procedure laid out by Benoit and Allard (2009), same-vessel trip sequences are analyzed to test for differences among various metrics. These trip sequences take the form of either: (1) three unobserved trips in a row (UUU), or (2) one observed trip between unobserved trips (UOU). To attenuate the possibility of interpreting seasonal effects as behavioral effects, only trips occurring within 45 days of each other are included. Trips are not repeated in multiple sequences. Vessels with less than two sequences are excluded from the analysis.

Triplet sequences are winnowed to pairs by taking the difference of either the leading or lagging trip with respect to the middle trip. The variable U in equation (1) and U^1 in equation (2), below, are selected randomly as either the leading or trailing trip in the triplet sequence, while the middle trip in the sequence is always the reference trip (O or U^1, below). To mitigate against regulatory changes affecting fishing behavior within sequences while maximizing the number of OU pairs, sequences overlapping the start of a new fishing year (May 1 of each year) select only the lead or lag pair that occurs in the same FY as the reference trip.

Differences are calculated as

$$\Delta O_{yfv} = (O - U/\acute{U})_{yfv} * 100$$

(Equation 1)

$$\Delta U_{yfv} = \left(U^1 - U^2/\acute{U}\right)_{yfv} * 100$$

(Equation 2)

where y is a fishing year, f is fishing vessel and v is any one of the metrics evaluated. $\acute{U}$ is the mean unobserved value for each year, vessel and metric combination.

Metrics evaluated, v, are:

1. Trip duration
2. Kept catch
3. Total revenue
4. Kept groundfish
5. Kept non-groundfish
6. Groundfish average price
7. Opportunity cost of quota

8. Number of groundfish market categories included in kept catch

The difference between the median values for ΔU's and O's is calculated as

$$(M_{\Delta U-\Delta O})_{yfv} = median(\Delta U)_{yfv} - median(\Delta O)_{yfv}$$

.

(Equation 3)

Differences between observed and unobserved trips are tested in three ways: (1) location differences are observed in $M_{\Delta U-\Delta O}$, with 95% confidence intervals estimated using bootstrap sampling (1,000 replicates) from the U_{yfv} and O_{yfv} values, where a lack of overlap with zero implies a 95% probability that the true median values for each population are significantly different[2]; (2) the Kolmogorov-Smirnov statistic is used to test for general differences in shape of the U_{yfv} and O_{yfv} distributions; and (3) the Kuiper statistic is used to test for differences in the extremities of the distributions (Conover 1980).

Multiple hypothesis tests are performed with the Kolmogorov-Smirnov (KSA) and Kuiper (KA) statistics. For these, a p-value of 0.005 is considered to be significant. As always, statistical significance should be considered in light of the data and research question. All p-values are reported.

Data

Vessel Trip Report (VTR) and Commercial Fishery Dealer (CFDBS) data are combined to construct trip-level data using the Data Matching and Imputation System (DMIS) database [cite needed]. Trips with an Allocation Management System (AMS) declaration code of "NMS" are included in the initial dataset[3]. Only vessels fishing with trawl or gillnet gears are retained. Observer trips are matched by a step-wise algorithm, focusing on permit number, VTR serial number, days-at-sea (DAS) identification number, date and time sailed. For the sector years, both Northeast Fishery Observer Program (NEFOP) and at-sea monitoring (ASM) data are matched.

U and O values are extracted from these data, and annual fishing year (May 1 – April 30) data sets are built with same-vessel two-trip sequences.

Trips in the United States-Canada Resource Sharing Agreement Area (USCA area) are removed from the pre-sector (FY 2007-2009) dataset, as these trips were subject to observer coverage at higher rates than trips outside the area. All trips fishing with extra large mesh (ELM) and targeting non-groundfish are excluded for all years, as are all trips by vessels enrolled in the Common Pool from 2010-2017[4]. All excluded trips and their corresponding triplets are retained and, to better understand the potential drivers of observer effects, are be analyzed separately in the future.

Results

Results are reported at two levels of aggregation:

- regulatory regime, as
  - pre-sector years (FY's 2007-2009),

[2] "Location" refers to the central tendency of the data, in this case the median values, and has no geographic connotation here.

[3] "NMS" is the code denoting trips made under the Northeast Multispecies Fishery Management Plan.

[4] In 2015 the New England Fishery Management Council exempt gillnet vessels fishing with mesh larger than 10 inches in certain areas near the coast from ASM coverage, as these trips had a documented history of catch very little groundfish. These trips are subject to NEFOP coverage, however.

  - initial sector years (FY's 2010-2012),
  - intermediate sector years (FY's 2013-2015),
  - contemporary sector years (FY's 2016-2018)[5]; and
- gear type, distinguishing between trawl and gillnet gears[6].

Results at the fishing year (FY) level, further disaggregated by gillnet and trawl, are estimated for context. Separate analyses have also been completed for single-day and multi-day trips, as well as a stock-level analysis of kept catch for 15 individual groundfish stocks.

Tests for differences in central tendency

Equations (1) and (2) are scaled by each vessel's mean annual values and median value differences are represented as percentages. For example, a median value of -0.04 for the kept catch variable implies that vessels catch roughly 4% less fish on an observed trip, relative to a neighboring unobserved trip by that same vessel, as measured across all vessels in the dataset. If the bootstrapped 95% confidence intervals fail to overlap with zero, the value is interpreted as significant using the confidence interval test. With eight metrics evaluated over four time stanzas, there are 32 units evaluated for observer effects. However, in the first stanza, before the sector system, there were no tradeable quota allocations.

Trawl vessels

For trawl vessels, 18 bootstrapped 95% confidence intervals failed to overlap zero. In the pre-sector years, three of seven metrics are significant under this test. In the three sector stanzas, 15 metrics are significant and nine are not.

Trawl vessels catch less fish when an observer is onboard. In the stanzas after 2009, they fish for less time and land less groundfish. Statistical significance is obtained for kept catch in all four stanzas, and for trip duration, groundfish kept catch and total revenues in the three post-2009 stanzas. Groundfish average prices are statically higher for three of the four stanzas, the exception being the period from 2010-2012. Composition of groundfish catch on observed and unobserved trips appears to be different. In the second and third time stanzas, groundfish vessels landed less high quota value stocks on observed trips, while in the final stanza the median differential is zero. Based on the reductions in catch and fishing time on observed trips after 2009, the changes in response to observer presense appear to be related to incentives embedded in catch accountability and quota constraints.

[5] FY 2018 data are complete through February 28 and inclusive of the first 10 full months of the fishing year.
[6] Trawl gears include the Vessel Trip Report (VTR) codes 'OHS','OTB','OTC','OTF','OTM','OTO','OTR','OTS', and 'OTT'. Gillnet gears include the codes 'GNR','GNS', and 'GNT'.

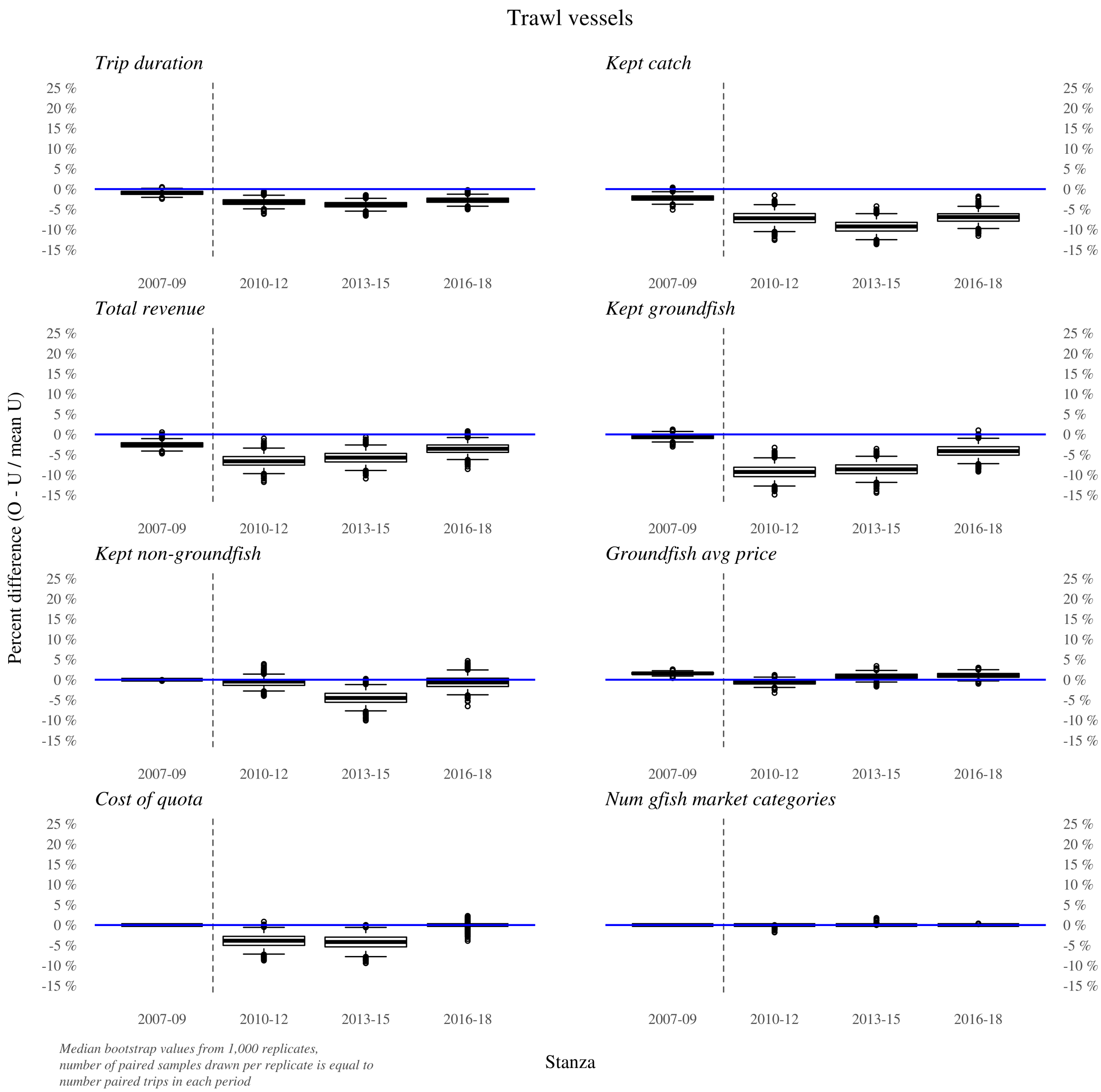


Figure 1: Results of bootstrap analysis, observed and unobserved same-vessel paired trips by stanza

Table 1: Stanza 1, 2007-2009

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Trawl	Kept groundfish		-1.9 %	-0.6 %	0.5 %	10,844	726
Trawl	Number groundfish market categories		0 %	0 %	0 %	10,844	726
Trawl	Groundfish avg price	*	0.9 %	1.6 %	2.3 %	10,845	726
Trawl	Kept catch	*	-3.7 %	-2.2 %	-0.7 %	10,845	726
Trawl	Kept non-groundfish		0 %	0 %	0 %	10,845	726
Trawl	Opportunity cost of quota		0 %	0 %	0 %	10,845	726
Trawl	Total revenue	*	-4.1 %	-2.6 %	-1.1 %	10,845	726
Trawl	Trip duration		-2 %	-0.9 %	0 %	10,845	726

Table 2: Stanza 2, 2010-2012

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Trawl	Kept groundfish	*	-12.6 %	-9.3 %	-5.9 %	2,787	1,413
Trawl	Number groundfish market categories		-0.4 %	0 %	0 %	2,787	1,413
Trawl	Groundfish avg price		-1.9 %	-0.6 %	0.6 %	2,787	1,413
Trawl	Kept catch	*	-10.2 %	-7.2 %	-4.1 %	2,787	1,413
Trawl	Kept non-groundfish		-3.3 %	-0.4 %	1.7 %	2,787	1,413
Trawl	Opportunity cost of quota	*	-7.3 %	-3.9 %	-0.8 %	2,787	1,411
Trawl	Total revenue	*	-9.4 %	-6.6 %	-3.4 %	2,787	1,413
Trawl	Trip duration	*	-4.9 %	-3.2 %	-1.6 %	2,787	1,413

Table 3: Stanza 3, 2013-2015

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Trawl	Kept groundfish	*	-12 %	-8.6 %	-5.4 %	2,920	954
Trawl	Number groundfish market categories		0 %	0 %	0.1 %	2,920	954
Trawl	Groundfish avg price		-0.5 %	0.8 %	2.3 %	2,920	954
Trawl	Kept catch	*	-12.3 %	-9.2 %	-6.1 %	2,920	954
Trawl	Kept non-groundfish	*	-7.9 %	-4.5 %	-1.4 %	2,920	954
Trawl	Opportunity cost of quota	*	-8 %	-4.2 %	-0.6 %	2,920	954
Trawl	Total revenue	*	-8.8 %	-5.7 %	-2.8 %	2,920	954
Trawl	Trip duration	*	-5.5 %	-3.8 %	-2.3 %	2,920	954

Table 4: Stanza 4, 2016-2018

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Trawl	Kept groundfish	*	-7 %	-4.1 %	-1.2 %	2,805	799
Trawl	Number groundfish market categories		0 %	0 %	0 %	2,805	799
Trawl	Groundfish avg price		-0.2 %	1.1 %	2.4 %	2,805	799
Trawl	Kept catch	*	-9.9 %	-6.9 %	-4.3 %	2,805	799
Trawl	Kept non-groundfish		-3.5 %	-0.7 %	2.5 %	2,805	799
Trawl	Opportunity cost of quota		-1.7 %	0 %	1 %	2,805	799
Trawl	Total revenue	*	-6.3 %	-3.5 %	-0.7 %	2,805	799
Trawl	Trip duration	*	-4.2 %	-2.7 %	-1.3 %	2,805	799

Gillnet vessels

For gillnet vessels the picture is less clear-cut. 13 units in total have 95% confidence intervals that fail to overlap with zero. Pre-sector, from 2007-2009, four metrics were significant and three were not. Under sector management, the three stanzas from 2010-2018, nine are signficant and thirteen are not. However, in the most recent stanza (FY 2016-2018), six of the eight metrics yeild significant differences in bootstrapped confidence intervals, and a seventh (number of groundfish market categories), while statistically insignificant, shows a trend toward more market categories landed on observed trips.

Gillnet vessels consistently make shorter trips, generate less revenue and appear to retain slightly less catch overall in the presence of an observer. There is a trend in later stanzas toward more groundfish and less non-groundfish on observed trips for these vessels, indicating that observers affect the mix of species landed. More groundfish market categories in the last stanza may indicate differential groundfish targeting, or perhaps high-grading of specific species. The most striking result is that, in the last stanza, with an observer on board the same gillnet vessels have a 17% higher opportunity cost of quota than when they do not. Statistically different behavior in response to an observer is nearly equally prevalent for gillnet and trawl vessels, though the nature of the response does differ between the two. This may be an artifact of smaller sample sizes (fewer number of paired trips, particularly in the later stanzas) which attenuate the model's power to discern effects. The distinction in response before and after the implementation of sectors is less clear cut for gillnetters than for trawlers, noting that gillnet vessels demonstrated a stronger behavioral response than trawlers before sectors. Finally, during the contemporary sector years (fourth stanza) a trend of less non-groundfish landed, more groundfish and, in particular, more high quota value species landed is noteworthy.

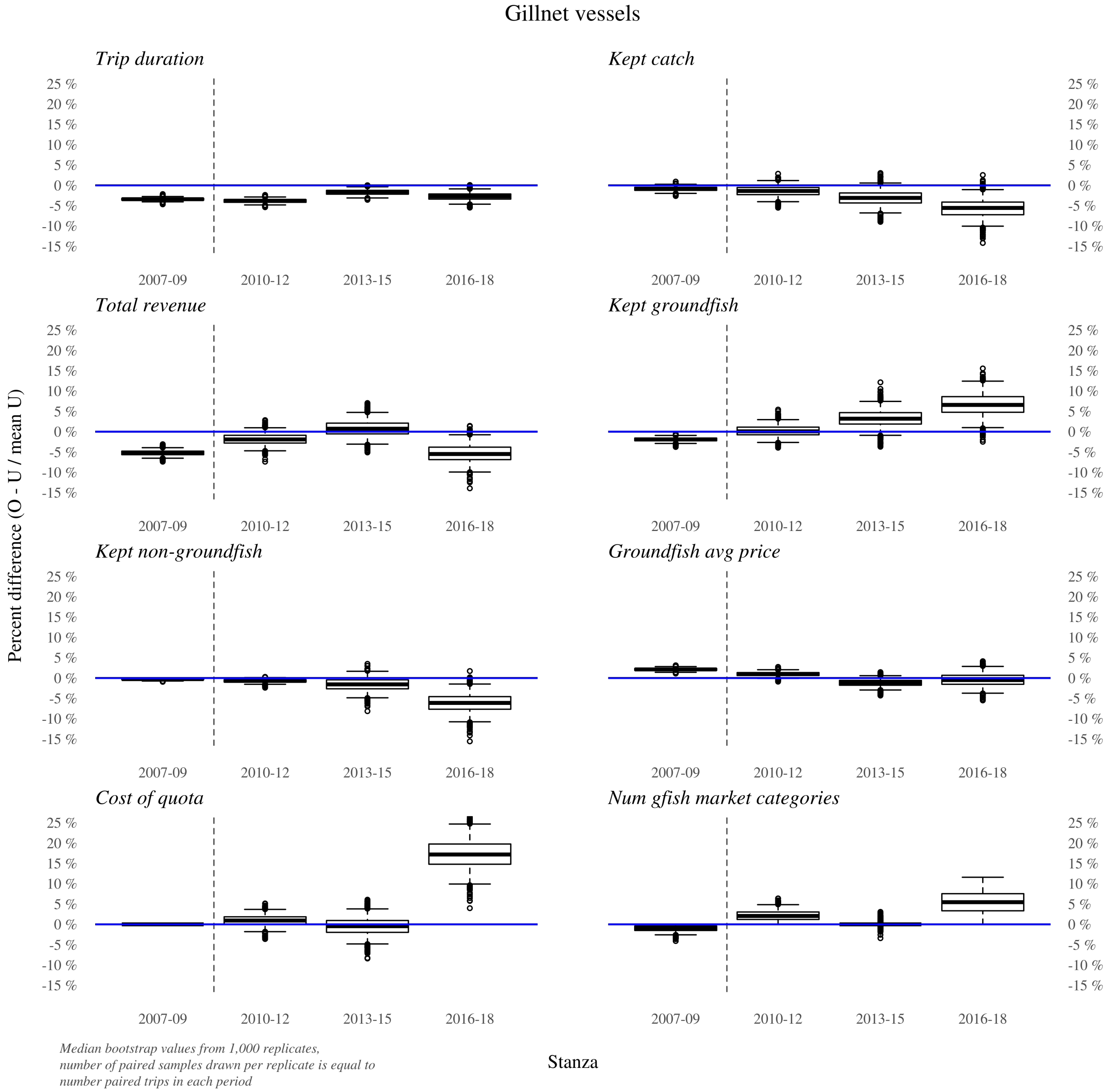


Figure 2: Results of bootstrap analysis, observed and unobserved same-vessel paired trips by stanza

Table 5: Stanza 1, 2007-2009

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Gillnet	Kept groundfish	*	-2.9 %	-1.9 %	-1 %	10,782	531
Gillnet	Number groundfish market categories		-2.8 %	-1 %	0 %	10,782	531
Gillnet	Groundfish avg price	*	1.5 %	2.1 %	2.8 %	10,782	531
Gillnet	Kept catch		-1.9 %	-0.8 %	0.1 %	10,782	531
Gillnet	Kept non-groundfish		-0.6 %	-0.3 %	0 %	10,782	531
Gillnet	Opportunity cost of quota		0 %	0 %	0 %	10,782	531
Gillnet	Total revenue	*	-6.5 %	-5.2 %	-4 %	10,782	531
Gillnet	Trip duration	*	-4.2 %	-3.4 %	-2.7 %	10,782	531

Table 6: Stanza 2, 2010-2012

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Gillnet	Kept groundfish		-2.4 %	0.1 %	3.2 %	2,609	1,330
Gillnet	Number groundfish market categories		0 %	2.1 %	4.9 %	2,609	1,330
Gillnet	Groundfish avg price		-0.2 %	1 %	2 %	2,609	1,330
Gillnet	Kept catch		-4.1 %	-1.4 %	1 %	2,609	1,330
Gillnet	Kept non-groundfish		-1.6 %	-0.7 %	0 %	2,609	1,330
Gillnet	Opportunity cost of quota		-1.8 %	0.9 %	3.8 %	2,609	1,330
Gillnet	Total revenue		-4.7 %	-1.9 %	1.1 %	2,609	1,330
Gillnet	Trip duration	*	-4.8 %	-3.8 %	-2.8 %	2,609	1,330

Table 7: Stanza 3, 2013-2015

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Gillnet	Kept groundfish		-0.9 %	3.2 %	7.6 %	1,622	434
Gillnet	Number groundfish market categories		-0.9 %	0 %	1.4 %	1,622	434
Gillnet	Groundfish avg price		-2.9 %	-1.2 %	0.4 %	1,622	434
Gillnet	Kept catch		-6.5 %	-3.1 %	0.4 %	1,622	434
Gillnet	Kept non-groundfish		-5.1 %	-1.6 %	1.2 %	1,622	434
Gillnet	Opportunity cost of quota		-5 %	-0.5 %	4.2 %	1,622	434
Gillnet	Total revenue		-3 %	0.7 %	4.9 %	1,622	434
Gillnet	Trip duration	*	-3 %	-1.7 %	-0.4 %	1,622	434

Table 8: Stanza 4, 2016-2018

Gear	Variable	CIs <> 0	95% CI, lower	Median	95% CI, upper	n Unobserved	n Observed
Gillnet	Kept groundfish	*	1.1 %	6.6 %	12.2 %	833	277
Gillnet	Number groundfish market categories		0 %	5.5 %	10.3 %	833	277
Gillnet	Groundfish avg price		-3.4 %	-0.5 %	2.7 %	833	277
Gillnet	Kept catch	*	-10.6 %	-5.6 %	-1 %	833	277
Gillnet	Kept non-groundfish	*	-10.8 %	-6.1 %	-1.5 %	833	277
Gillnet	Opportunity cost of quota	*	10.2 %	17.2 %	24.7 %	833	277
Gillnet	Total revenue	*	-9.6 %	-5.5 %	-1.1 %	833	277
Gillnet	Trip duration	*	-4.5 %	-2.7 %	-1 %	833	277

Tests for differences in distribution shape

The Kolmogorov-Smirnov (K-S) test, a nonparametric test evaluating the difference between cumulative distribution functions of two independent samples, U and O, is sensitive to differences in location and shape. Generally, at a 0.005 signficance level this test finds fewer significant differences in distribution shapes than the bootstrap confidence interval method for changes in location.

The Kuiper (K) test, another nonparametric test, is similar to the K-S but evaluates in an additive way both positive and negative differences in the cumulative distribution functions of the U and O values. It is more senstive, therefore, to changes in the tails of the distributions in question.

Trawl vessels

Of the 31 evaluated units, 12 are significant under the Kolmogorov-Smirnov test and 22 under the Kuiper test. In the pre-sector stanza, three of seven units have statistically significant differences in distribution shape (K-S) and, for all seven units, the tails of the U and O distributions are significantly different under the Kuiper test. In the three sector stanzas, nine units exhibit significantly different distributions under the K-S test, with 16 significanly different distributions under the Kuiper test.

The K-S test highlights similar units to the bootstrapped confidence intervals, namely kept catch, trip duration and kept groundfish. The Kuiper test, however, reveals differences in U and O distribution shapes for opportunity cost of quota (three sector stanzas) and number of groundfish market categories (all four stanzas).

Table 9: Stanza 1, 2007-2009

Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Trawl	Kept groundfish		0.179	*	0.002	10,844	726
Trawl	Number groundfish market categories	*	0.001	*	0.000	10,844	726
Trawl	Groundfish avg price	*	0.002	*	0.000	10,845	726
Trawl	Kept catch	*	0.002	*	0.000	10,845	726
Trawl	Kept non-groundfish		0.102	*	0.000	10,845	726
Trawl	Total revenue		0.169		0.031	10,845	726
Trawl	Trip duration		0.066	*	0.005	10,845	726

Table 10: Stanza 2, 2010-2012

Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Trawl	Kept groundfish	*	0.000	*	0.000	2,787	1,413
Trawl	Number groundfish market categories		0.149	*	0.000	2,787	1,413
Trawl	Groundfish avg price		0.272		0.029	2,787	1,413
Trawl	Kept catch	*	0.000	*	0.004	2,787	1,413
Trawl	Kept non-groundfish		0.625	*	0.002	2,787	1,413
Trawl	Opportunity cost of quota		0.101	*	0.000	2,787	1,411
Trawl	Total revenue	*	0.003		0.021	2,787	1,413
Trawl	Trip duration		0.007	*	0.001	2,787	1,413

Table 11: Stanza 3, 2013-2015

Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Trawl	Kept groundfish	*	0.000	*	0.002	2,920	954
Trawl	Number groundfish market categories		0.426	*	0.000	2,920	954
Trawl	Groundfish avg price		0.251		0.059	2,920	954
Trawl	Kept catch	*	0.001	*	0.004	2,920	954
Trawl	Kept non-groundfish		0.128		0.448	2,920	954
Trawl	Opportunity cost of quota		0.013	*	0.000	2,920	954
Trawl	Total revenue		0.016		0.077	2,920	954
Trawl	Trip duration	*	0.000	*	0.000	2,920	954

Table 12: Stanza 4, 2016-2018

Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Trawl	Kept groundfish	*	0.002	*	0.002	2,805	799
Trawl	Number groundfish market categories		0.127	*	0.000	2,805	799
Trawl	Groundfish avg price		0.180		0.346	2,805	799
Trawl	Kept catch	*	0.000	*	0.001	2,805	799
Trawl	Kept non-groundfish		0.649		0.443	2,805	799
Trawl	Opportunity cost of quota		0.178	*	0.000	2,805	799
Trawl	Total revenue		0.032		0.073	2,805	799
Trawl	Trip duration	*	0.000	*	0.000	2,805	799

Gillnet vessels

Only six of 31 units are significant under the Kolmogorov-Smirnov test and 9 under the Kuiper test for gillnet vessels. In the pre-sector stanza, three of seven units have statistically significant differences in distribution shape for both the K-S and Kuiper tests. In the three sector stanzas, three of 24 possible units exhibit significantly different U and O distributions under the K-S test, and 6 under the Kuiper test.

As with trawl vessels, the K-S test here highlights, when significant, difference similar o the bootstrapped confidence intervals. And also like with trawl vessels, the Kuiper test reveals differences in U and O distribution shapes for the number of groundfish market categories in all four stanzas.

Table 13: Stanza 1, 2007-2009

Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Gillnet	Kept groundfish		0.104		0.179	10,782	531
Gillnet	Number groundfish market categories		0.111	*	0.000	10,782	531
Gillnet	Groundfish avg price		0.012		0.027	10,782	531
Gillnet	Kept catch		0.722		0.456	10,782	531
Gillnet	Kept non-groundfish	*	0.001	*	0.000	10,782	531
Gillnet	Total revenue	*	0.002		0.007	10,782	531
Gillnet	Trip duration	*	0.002	*	0.001	10,782	531

Table 14: Stanza 2, 2010-2012

Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Gillnet	Kept groundfish		0.594		0.070	2,609	1,330
Gillnet	Number groundfish market categories	*	0.001	*	0.000	2,609	1,330
Gillnet	Groundfish avg price		0.161		0.645	2,609	1,330
Gillnet	Kept catch		0.182		0.108	2,609	1,330
Gillnet	Kept non-groundfish		0.006	*	0.000	2,609	1,330
Gillnet	Opportunity cost of quota		0.239		0.025	2,609	1,330
Gillnet	Total revenue		0.612		0.917	2,609	1,330
Gillnet	Trip duration	*	0.000	*	0.000	2,609	1,330

Table 15: Stanza 3, 2013-2015

Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Gillnet	Kept groundfish		0.137		0.018	1,622	434
Gillnet	Number groundfish market categories		0.942	*	0.000	1,622	434
Gillnet	Groundfish avg price		0.314		0.210	1,622	434
Gillnet	Kept catch		0.228		0.222	1,622	434
Gillnet	Kept non-groundfish		0.223		0.043	1,622	434
Gillnet	Opportunity cost of quota		0.167		0.028	1,622	434
Gillnet	Total revenue		0.110		0.010	1,622	434
Gillnet	Trip duration		0.034	*	0.004	1,622	434

Table 16: Stanza 4, 2016-2018

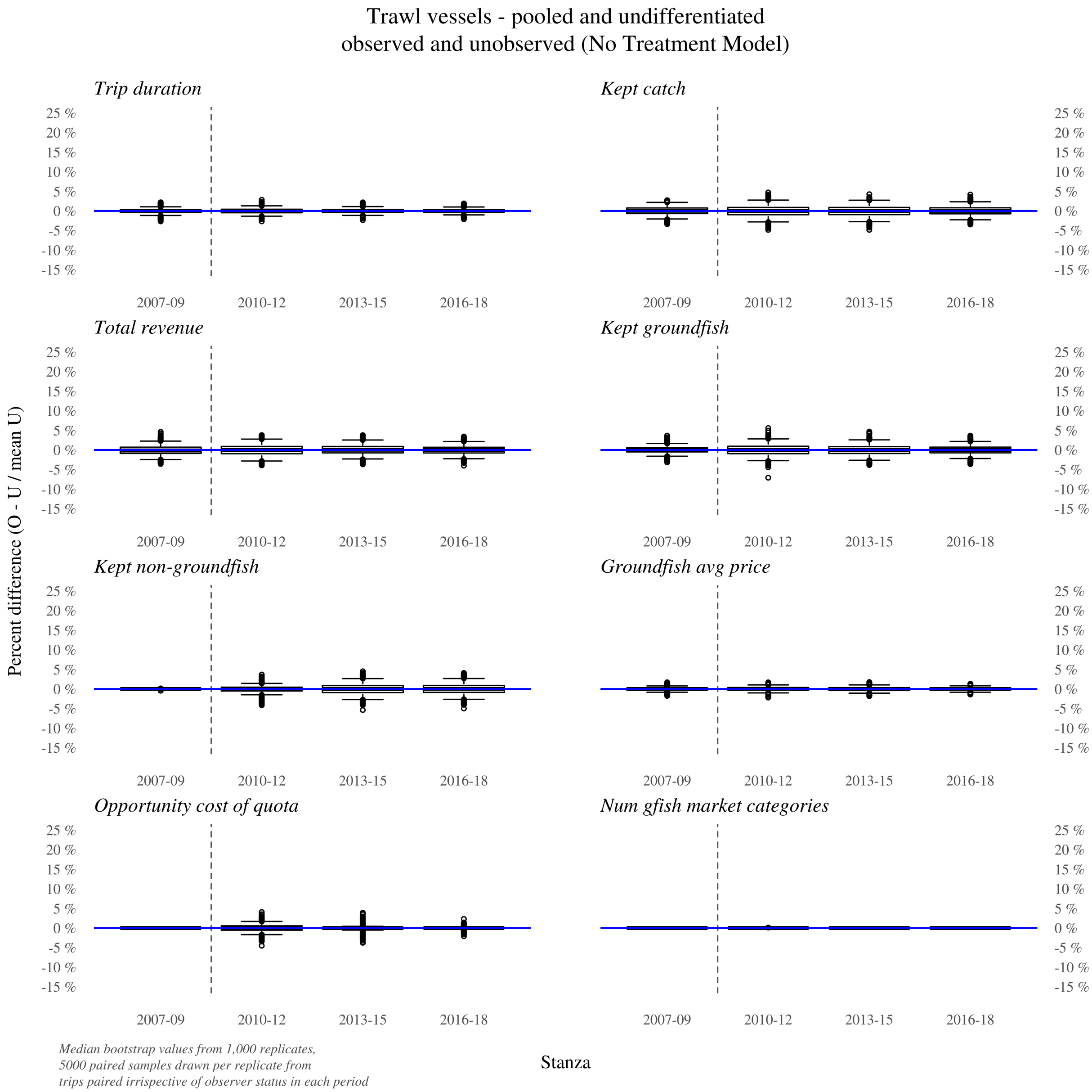
Gear	Variable	KS <= 0.005	p(KS)	K <= 0.005	p(K)	n Unobserved	n Observed
Gillnet	Kept groundfish		0.144		0.101	833	277
Gillnet	Number groundfish market categories		0.077	*	0.000	833	277
Gillnet	Groundfish avg price		0.702		0.486	833	277
Gillnet	Kept catch		0.040		0.033	833	277
Gillnet	Kept non-groundfish		0.041		0.100	833	277
Gillnet	Opportunity cost of quota	*	0.004		0.013	833	277
Gillnet	Total revenue		0.032		0.053	833	277
Gillnet	Trip duration		0.092		0.019	833	277

Discussion

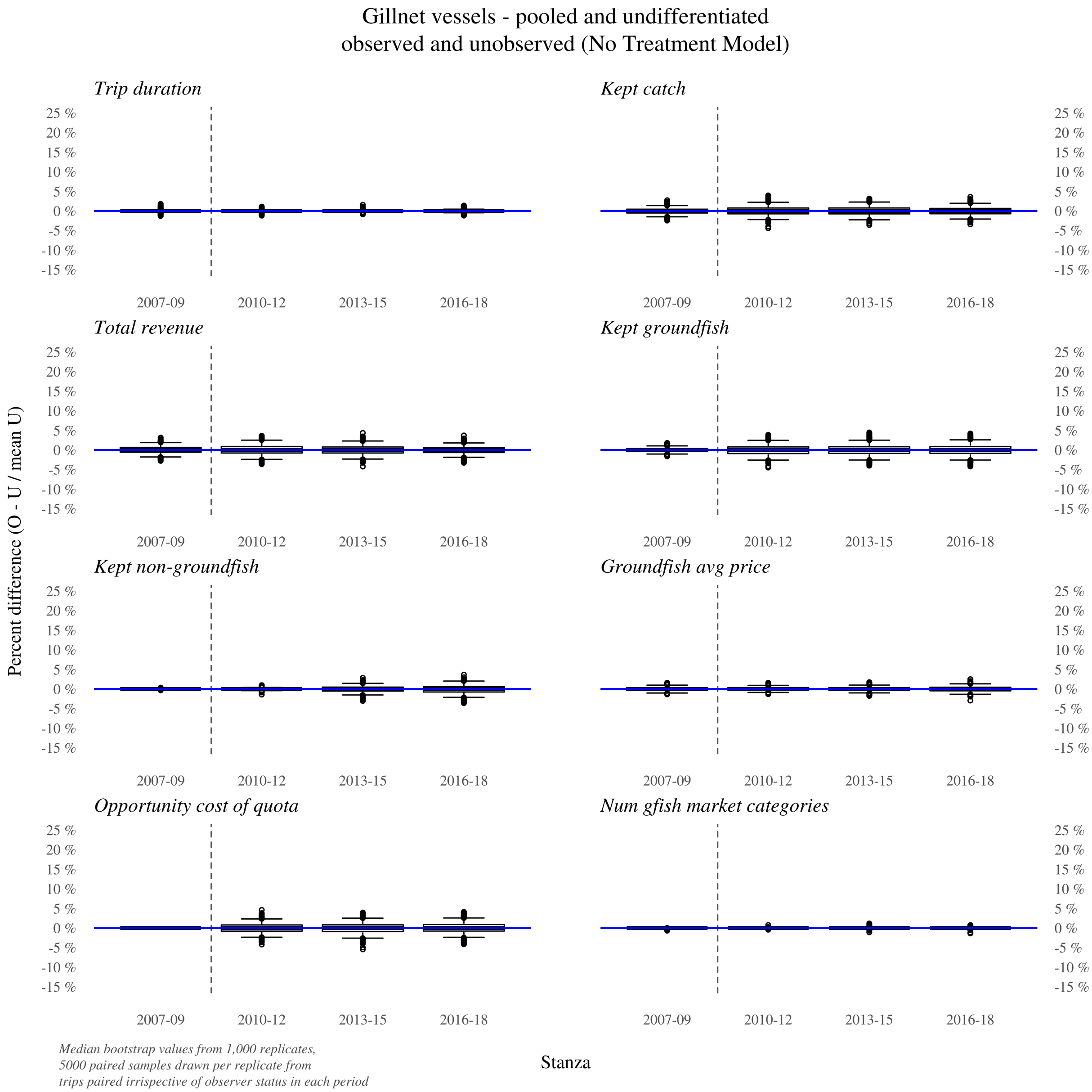
It is clear that fishing vessels engaged in the groundfish fishery alter their behavior in response to observers. Estimated confidence intervals for U and O values overlap with zero for only a handful of the metrics evaluated across stanzas or fishing years. Generally, the most pronounced effects are seen across trip duration, kept catch, kept groundfish, trip revenue and opportunity cost of quota. Observer presence has the smallest affect on the number of groundfish market categories and non-groundfish average prices, but, particulary in the former, even here we see differences in the tails of the distributions.

No treatment model

In an effort to demonstrate that the effects estimated here are, in fact, the result of observer presence and not driven by underlying variability in trip-level data driven by unobserved factors, the model was run as previously described, but with assignment to triplets (U and O) made irrespective of actual observer status. As one would expect, the No Treatment estimates across all metrics and stanzas are median-centered on zero with little variance in the two distributions. This demonstrates that the observed variation between U and O triplets in the primary (treatment) model is almost certainly a function of observer presence. See Appenix (forthcoming) for details.



Trawl vessels - pooled and undifferentiated
observed and unobserved (No Treatment Model)
Trip duration
Kept catch
Total revenue
Kept groundfish
Kept non-groundfish
Groundfish avg price
Opportunity cost of quota
Num gfish market categories
25 %
20 %
15 %
10 %
5 %
0 %
-5 %
-10 %
-15 %
2007-09
2010-12
2013-15
2016-18
Percent difference (O - U / mean U)
Stanza
Median bootstrap values from 1,000 replicates,
5000 paired samples drawn per replicate from
trips paired irrispective of observer status in each period



Differences across time

Incentives to alter fishing behavior have varied across time. Prior to sector implementation discards had no direct cost to fisherman and trip limits required discarding certain species. These factors may have reduced the incentive to alter fishing practices in response to an observer, noting that gillnet vessels did demonstrate a significant behavioral response prior to sectors. Gillnet vessels, however, are also more likely to have encounters with marine mammals and have other gear-specific requirements (i.e. pingers) that may further affect responses to observers independent of quota-based management and associated regulatiosn.

After full sector implementation, the accountability of discards and the application of sector/gear specific discard rates to unobserved trips, together with the potential catch of constraining stocks and the high opportunity cost of quota associated with landing such stocks, increased the incentive to change behavior. We see this most dramatically in the contemporary sector stanza for gillnet vessels, but the trend from lower quota costs on observed trip toward zero difference on trawl vessels may reflect a similar response.

The two-sided problem

Incentives to alter behavior in response to an observer may induce less effort, catch, etc…or more, as some vessels fish longer (or shorter) trips or otherwise alter their fishing practices due to quota allocations, fishing preferences, or other factors. One vessel may attempt to minimize observed discarding of flatfish at the expense of cod, while another vessel may take the exact opposite approach. Such offsetting behavior could change the central tendency of the $M_{\Delta U - \Delta O}$ distribution very little, but affect it's shape, particularly at the tails. Number of market categories for groundfish and opportunity cost of quota differ at the tails for both gillnet and trawl vessels. These distribution differences may point toward highgrading and/or circumventing mandatory fish retention regulations.

More broadly, the two-sided nature of the problem is important to understand because directionally opposite responses to observer presence attenuates the central tendency test and some may view location differences on the order of 5-10% as trivial when, taken in context, they represent large and statistically significant differences between observed and unobserved populations.

To better understand the influence of positive and negative observer responses, we estimated median annual (FY) values across each of the eight metrics for all vessels represented in the matched pair data, subtracting each vessel's annual median U value from it's median O to get a median difference in observed behavior. An example of the distribution of vessel-level observer effects by FY, in this case for opportunity cost of quota, can be seen below.

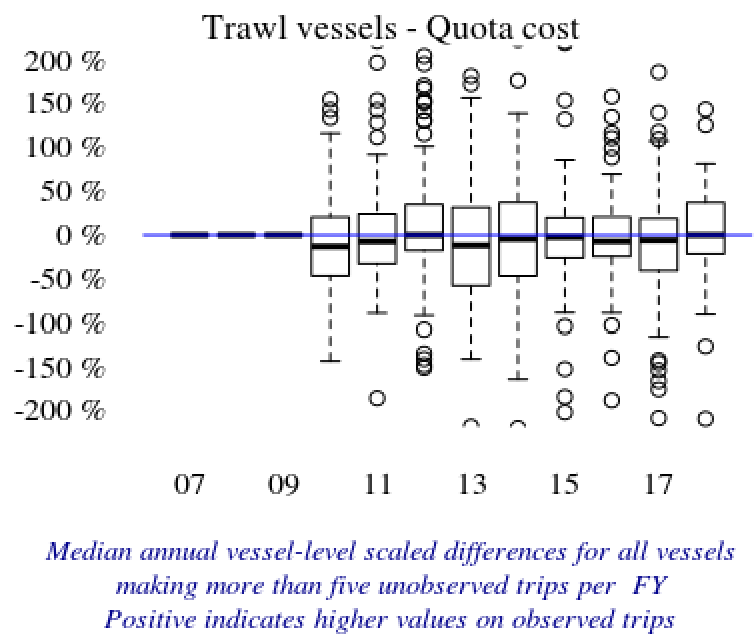


Figure 3: Distribution of vessel-level median annual observer effects, trawl)

These plots make clear the point that over the course of a year, some vessels persistently shift their behavoir in response to observer in a positive direction, others the opposite.

The effect of these off-setting behaviors may be that a large amount of catch can be taken by vessels that persistently alter behavior in one direction or the other. To test this, and to better understand how much fishing activity may be affected, we take two sub-sets of vessels–those that exhibit a +/- 15% median annual

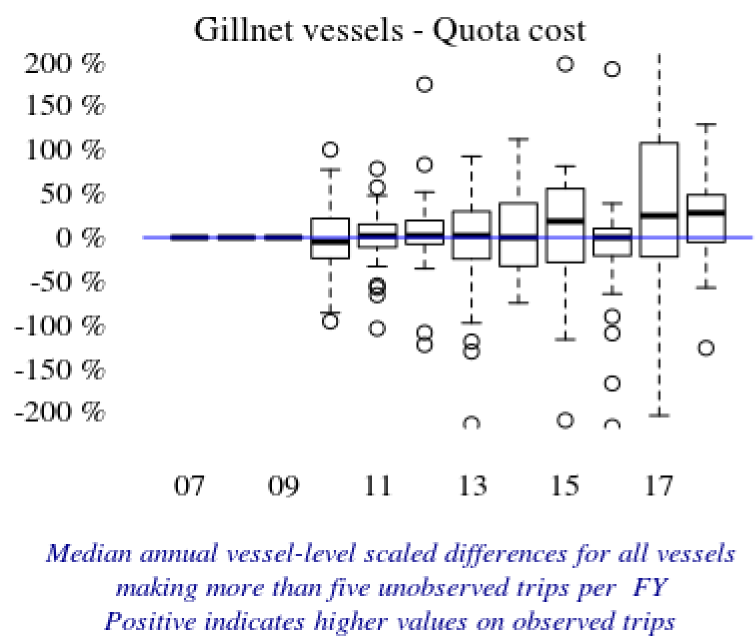


Figure 4: Distribution of vessel-level median annual observer effects, gillnet)

difference in behavior (oserver effect) for each metric, and those with a +/- 30% difference–and estimate the proportion of vessels and groundfish catch accounted for annually by these sets. We find that across a range of metrics, vessels with an annual observer effect response of +/- 15% or more account for roughly 20-30% of the groundfish vessels, and roughly 50-60% of the groundfish catch. Vessels with a +/- 30% response account for 10-20% of the vessels and 30-40% of the catch. Vessels exhibiting these levels of observer effect for the opportunity cost of quota metric, in particular, represent the largest share of groundfish catch, from 40-80% depending on threshold and year. It is important to note that, even in the case of no observer effect, the nature of fishing and it's underlying variability would likely result in some vessels fitting into one or both of these threshold categories. Further analysis of, for example, the extra-large mesh fishery, which has no quota-based incentives that may benefit from observer effects, may shed more light on the question of underlying variability versus strategic behavioral responses.

Last word

These analyses point toward a consistent pattern of different fishing behaviors when an observer is on board. The Benoit and Allard method isolates vessel effects by focusing on the differences in behavior in response to an observer *for the same vessel*. The data show a clear trend for three key metrics–in almost all circumstances vessels appear to retain less fish, fish for less time and obtain lower revenues when an observer is on board. Gillnet vessels retain substantially more groundfish, at a higher opportuntiy cost of quota, in the most recent time stanza. The distributions of U and O pairs is substantially different at the tails for the number of groundfish market categories landed, pointing toward highgrading by a subset of the fleet. Persistent differences such as higher average groundfish prices with an observer on board (trawl vessels) and emerging differences like a greater number of market categories retained with an observer (gillnet vessels) indicate that the composition of catch on observed trips is different. This suggests that data collected by observers are not merely a compressed representation of unobserved fishing practices but, rather, they are non-representative along critical dimensions such as proportions and quantities of discarded fish, legally and perhaps illegally, and fish retained.

\begin{table}[t]

\caption{Vessel median observer effects > +/- 15% and 30%, proportion of total and proportion of groundfish landed}

FY	Variable	N vsls	Vsls, > +/-15%	% gfish caught +/-15	Vsls, > +/-30%	% gfish caught +/-30
2007	gfish_lbs	564	125	0.35	90	0.27
2007	gfish_mcat	564	91	0.22	53	0.11
2007	gfish_price	564	77	0.29	32	0.13
2007	k_all	564	114	0.38	86	0.28
2007	non_gfish_lbs	564	92	0.26	75	0.23
2007	total_value	564	124	0.39	91	0.28
2007	trip_dur	564	89	0.30	57	0.17
2008	gfish_lbs	527	129	0.31	91	0.23
2008	gfish_mcat	527	117	0.27	61	0.12
2008	gfish_price	527	81	0.25	54	0.17
2008	k_all	527	137	0.35	95	0.26
2008	non_gfish_lbs	527	113	0.38	80	0.28
2008	total_value	527	134	0.38	90	0.25
2008	trip_dur	527	101	0.30	59	0.15
2009	gfish_lbs	476	114	0.51	79	0.35
2009	gfish_mcat	476	107	0.33	60	0.18
2009	gfish_price	476	88	0.36	48	0.24
2009	k_all	476	120	0.51	86	0.33
2009	non_gfish_lbs	476	118	0.48	93	0.33
2009	total_value	476	124	0.46	86	0.30
2009	trip_dur	476	102	0.40	63	0.25
2010	gfish_lbs	377	96	0.55	56	0.26
2010	gfish_mcat	377	72	0.27	33	0.14
2010	gfish_price	377	56	0.36	22	0.18
2010	k_all	377	95	0.48	66	0.33
2010	non_gfish_lbs	377	82	0.49	64	0.37
2010	quota_cost	377	103	0.53	76	0.43
2010	total_value	377	99	0.49	63	0.32
2010	trip_dur	377	64	0.43	31	0.22
2011	gfish_lbs	362	113	0.54	80	0.43
2011	gfish_mcat	362	61	0.23	22	0.09
2011	gfish_price	362	49	0.29	18	0.08
2011	k_all	362	98	0.41	58	0.30
2011	non_gfish_lbs	362	79	0.41	55	0.29
2011	quota_cost	362	99	0.45	61	0.30
2011	total_value	362	108	0.48	68	0.28
2011	trip_dur	362	64	0.35	32	0.22

\end{table}

\begin{table}[t]

\caption{Vessel median observer effects > +/- 15% and 30%, proportion of total and proportion of groundfish landed}

FY	Variable	N vsls	Vsls, > +/-15%	% gfish caught +/-15	Vsls, > +/-30%	% gfish caught +/-30
2012	gfish_lbs	352	131	0.67	87	0.44
2012	gfish_mcat	352	75	0.27	29	0.09
2012	gfish_price	352	77	0.44	41	0.20
2012	k_all	352	122	0.62	75	0.45
2012	non_gfish_lbs	352	115	0.59	91	0.48
2012	quota_cost	352	113	0.61	79	0.43
2012	total_value	352	125	0.65	72	0.37
2012	trip_dur	352	90	0.53	52	0.34
2013	gfish_lbs	305	102	0.62	67	0.43
2013	gfish_mcat	305	62	0.26	31	0.10
2013	gfish_price	305	65	0.49	27	0.25
2013	k_all	305	100	0.63	72	0.49
2013	non_gfish_lbs	305	95	0.66	62	0.36
2013	quota_cost	305	105	0.73	84	0.60
2013	total_value	305	92	0.61	52	0.35
2013	trip_dur	305	64	0.55	36	0.31
2014	gfish_lbs	280	85	0.70	60	0.45
2014	gfish_mcat	280	52	0.32	26	0.14
2014	gfish_price	280	57	0.51	32	0.24
2014	k_all	280	80	0.64	48	0.39
2014	non_gfish_lbs	280	71	0.53	55	0.41
2014	quota_cost	280	95	0.71	72	0.49
2014	total_value	280	90	0.67	56	0.39
2014	trip_dur	280	66	0.54	31	0.21
2015	gfish_lbs	250	75	0.55	56	0.37
2015	gfish_mcat	250	50	0.18	27	0.11
2015	gfish_price	250	46	0.42	24	0.19
2015	k_all	250	76	0.52	63	0.41
2015	non_gfish_lbs	250	82	0.63	63	0.45
2015	quota_cost	250	80	0.46	59	0.36
2015	total_value	250	76	0.47	51	0.28
2015	trip_dur	250	63	0.52	41	0.35
2016	gfish_lbs	230	67	0.56	46	0.29
2016	gfish_mcat	230	39	0.14	19	0.05
2016	gfish_price	230	46	0.42	20	0.16
2016	k_all	230	82	0.70	51	0.40
2016	non_gfish_lbs	230	69	0.56	53	0.32
2016	quota_cost	230	78	0.74	44	0.41
2016	total_value	230	73	0.54	41	0.35
2016	trip_dur	230	50	0.66	20	0.12

\end{table}

\begin{table}[t]

\caption{Vessel median observer effects > +/- 15% and 30%, proportion of total and proportion of groundfish landed}

FY	Variable	N vsls	Vsls, > +/-15%	% gfish caught +/-15	Vsls, > +/-30%	% gfish caught +/-30
2017	gfish__lbs	213	73	0.63	50	0.35
2017	gfish__mcat	213	42	0.17	14	0.06
2017	gfish__price	213	48	0.43	24	0.12
2017	k__all	213	67	0.59	43	0.28
2017	non__gfish__lbs	213	73	0.63	48	0.44
2017	quota__cost	213	76	0.60	54	0.43
2017	total__value	213	72	0.61	49	0.44
2017	trip__dur	213	52	0.66	25	0.46
2018	gfish__lbs	198	50	0.31	39	0.25
2018	gfish__mcat	198	45	0.20	13	0.05
2018	gfish__price	198	37	0.25	15	0.09
2018	k__all	198	58	0.51	28	0.34
2018	non__gfish__lbs	198	51	0.64	27	0.39
2018	quota__cost	198	58	0.69	39	0.44
2018	total__value	198	51	0.46	33	0.20
2018	trip__dur	198	36	0.42	18	0.22

\end{table}

References

Benoît H. P., Allard J. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? Canadian Journal of Aquatic Sciences 2009; 66:2025-2039.

Conover, W.J. 1980. Practical non-parametric statistics. 2nd ed. John Wiley and Sons, Inc, New York.

Faunce C. H., Barbeaux S. J. The frequency and quantity of Alaskan groundfish catcher-vessel landings made with and without an observer. ICES Journal of Marine Science 2011; 68 (8): 1757-1763.

McCambridge, Jim, John Witton, and Diana R. Elbourne. "Systematic Review of the Hawthorne Effect: New Concepts Are Needed to Study Research Participation Effects." Journal of Clinical Epidemiology 67.3 (2014): 267–277. PMC. Web. 30 Apr. 2018.

New England Fishery Management Council (NEFMC). Framework Adjustment 46 to the Northeast Multispecies Fishery Management Plan. New England Fishery Management Council, Newburyport, MA. 2012. 299 pp.

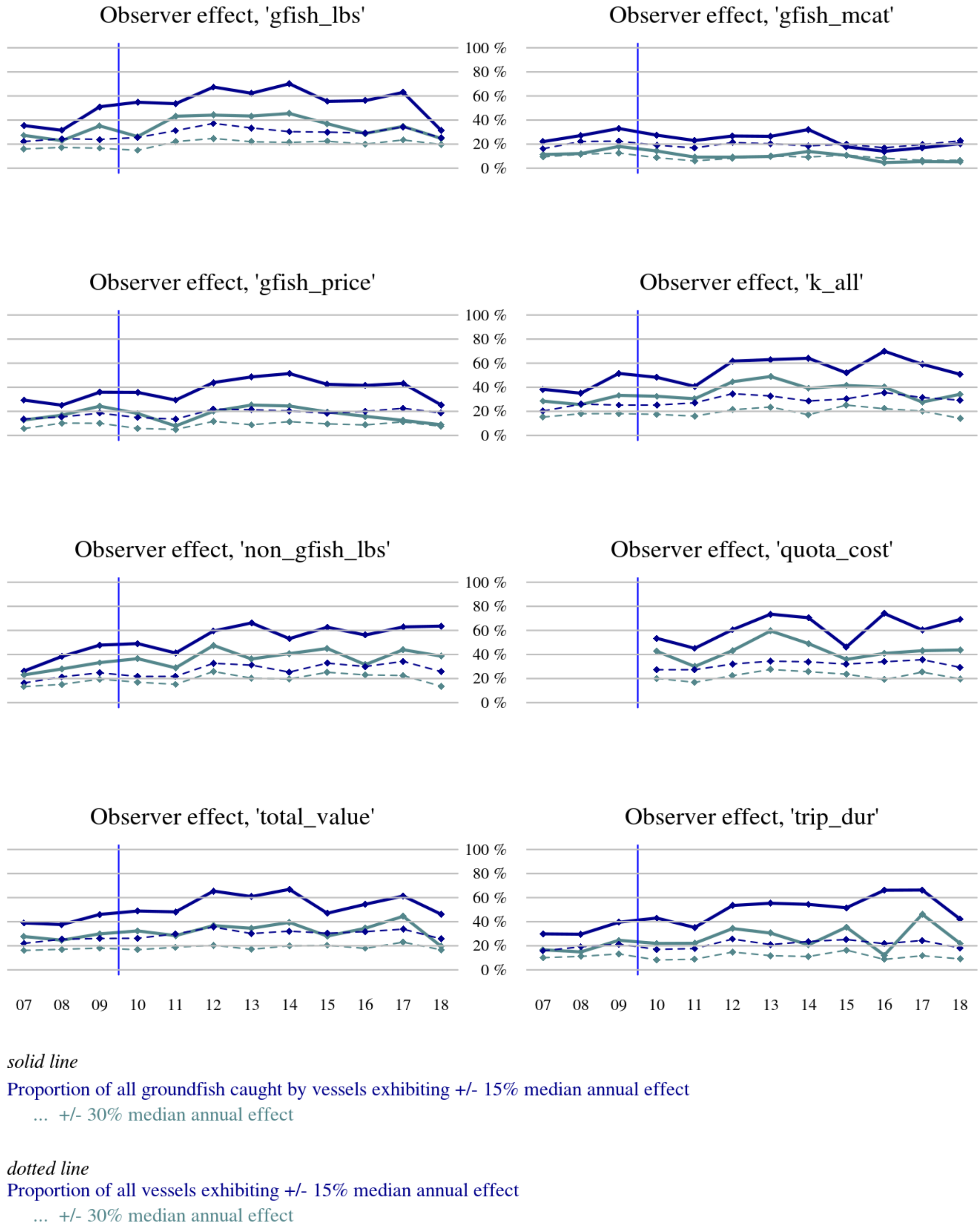


Figure 5: Proportion of vessels and catch accounted for by vessels with median annual observer effect greater than +/- 15 and 30%