# Appendix

The Appendix is divided into six sections that provide further details on: A) the sources and processing of the data used for the applied case study; B) the alternative spatial indexes used to uncover possible correlations between each model’s results and the spatial heterogeneity of data availability; C) the tessellations with varying levels of spatial aggregation that were considered for estimating the model; D) the full methodology followed to carry out the Monte Carlo experiments; E) the complementary estimation assumptions that have been made for estimating the model in the applied case study ; and, F) complementary results for both the Monte Carlo experiments and the applied case study.

## Gulf of Mexico fishery Data

Data used to study the BLL fleet was provided by the National Marine Fisheries Service Southeast Fisheries Science Center (SEFSC). The SEFSC’s Socioeconomic Panel (SEP) combines a variety of data sources to create a rich trip-level data set. The panel includes extensive information from the Coastal Logbook on landings disaggregated by species as well as effort data with reported variables depending on the primary gear used during the trip. In the case of BLL trips, these variables include soaking time, number of hooks per line, and the number of sets during the trip[[1]](#footnote-1). Effort data such as number of days at sea, number of crew, date of landing, and dealer identifiers are also reported. The Logbook data is further supplemented with average price data from the Accumulated Landings Service to calculate trip revenues which are also disaggregated by species. Vessel technical characteristics (e.g., vessel length) from the NMFS Southeast Regional Office Permits Office are linked to the Logbook data at the trip level. VMS data were provided by the SEFSC’s Social Science Research Group.

We identified the GT-BLL fleet using the Topgear variable provided in the SEP and linked these trips to VMS using vessel identifiers along with trip start and end dates (based on reported landing date and number of days at sea). After discarding a small number of logbook entries (38 out of 4054 for the years analyzed) that had trip dates that were overlapping other entries (e.g., because of reporting mistakes or because the entries referred to a same trip), and after further subsetting GT-BLL trips to trips deriving more than 75% of the revenue from GT species, we were left with, respectively for 2008 and 2012, 816 and 420 logbook entries matched to 99,027 and 42,008 VMS observations classified as fishing, covering 362 and 350 different days of the year and representing observations from 104 and 54 vessels.

VMS pings corresponding to fishing behavior were identified using a random forest model (O’Farrell et al., 2017), trained with observer data provided by the SEFSC’s Galveston Laboratory. Specifically designed and tested for this dataset, this approach makes use of observer data to devise the best classification along factors such as vessel’s speed, heading and previous behavior (the estimated accuracy rate in predicting fishing activity on the training dataset is 92 %).

## *Spatial indexes used for the analysis*

### Index of spatial aggregation

To analyze the results of the simulations relative to the level of spatial aggregation used during the estimation of the DCM, we associate each tessellation to an index of spatial aggregation defined as the logarithm of the ratio of the area of the aggregated alternatives to the area of the “true” alternatives (i.e., those considered for the decision-making process):

means that the DCM was estimated at the same spatial scale as the one used during the decision-making process.

means that the DCM was estimated using alternatives that were times larger in terms of area, or times longer in terms of length, than the alternatives used during the decision-making process.

### Indexes of spatial distribution

In addition to analyzing the effect of aggregating spatial choices made at a more refined scale, we are also interested in analyzing the effect of the spatial distribution of the observed choices. For that reason, we have considered different sizes of fishing hotspots which induced different spatial distributions of choices, with observations being more highly concentrated with smaller hotspots.

In addition to analyzing the results of the simulations individually for each of the spatial distributions, we also analyzed the results using indexes aiming at capturing the nature of the spatial distribution of choices relatively to the spatial resolution of the estimated models. For that purpose, we computed for each pair of spatial distribution/hotspot’s size and tessellation, the associated Shannon entropy and equitability indexes, and the relative spatial resolution. The relative spatial resolution is defined as the logarithm of the ratio of the size of the hotspot (defined as the area including the 95% highest VPUE levels) with the size of an alternative:

means that a given hotspot is covered by alternatives.

The Shannon entropy index associated with a draw is:

Where, is the empirical frequency of choice for the alternative (i.e. the number of simulated fishing locations falling in alternative over the total number of simulated fishing locations); and is the whole set of alternatives considered (i.e., having ).

The more skewed the distribution of choice frequency, the smaller the corresponding entropy. If almost all of the simulated fishing locations are concentrated in only one alternative, and the other alternatives are very “rare”, the entropy approaches zero.

Should the simulated fishing locations cover the alternatives in a perfectly balanced way, all the would be equal, and the index would take the value , where is the size of (i.e., the number of alternatives considered). Therefore, a distribution of choices having a Shannon entropy index of can be interpreted as “as diverse” as an even distribution of choices among alternatives.

The Shannon entropy index is commonly used by ecologists as a diversity index (e.g., of species, phenotypes etc.) and as a measure of the predictability of type: the higher the index, the more diversity and the less predictable the type. This interpretation of the index can be somewhat misleading in our case since it assumes that predictions are based only on empirical frequencies and that there is not underlying model.

The Shannon equitability index is simply the Shannon entropy index normalized by its maximum value:

## Maps of the partitions of space considered

|  |
| --- |
| Une image contenant dessin  Description générée automatiquementUne image contenant horloge, dessin  Description générée automatiquement |
| **Figure C.a.** Maps and indexes of spatial aggregation of the ten partitions of space considered in the Monte Carlo experiments for estimating the discrete-choice model of fishing locations, with the data-generating process set at on a 1x1 (left panels) or 10x10 (right panels) grid. The index of spatial aggregation is computed as the logarithm of the ratio of the area of the aggregated alternatives to the area of the “true” alternatives. |

|  |
| --- |
|  |
| **Figure C.b.** Maps and indexes of spatial aggregation of the nine tessellations of the Gulf of Mexico considered for estimating the discrete-choice model of fishing locations. The index of spatial aggregation is computed as the logarithm of the ratio of the area of the aggregated alternatives to the area of the alternatives considered by the NMFS, starting 2013, for the reporting of fishing locations in logbooks. |

## Monte Carlo experiments

### Utility function

We assume vessels’ decide on their fishing location according to the following utility function:

With:

* : We assume a “fixed-cost” of moving and that all vessels start their trip from the origin (0,0)
* : We assume that there is a hotspot in () with a Gaussian spatial distribution of fish abundance that depends on time.
  + We assume 3 hotspots located in the North (2.6, 2.5), Center (3.1, 0) or South (1.9, -2.4).
  + We consider 4 hotspot sizes taking
* :
  + The productivity of hotspots oscillates around with a period and reaches their maxima at .
  + We set .
* The productivity of a given point in space is the mean of the productivity of each hotspot at this given point with a stochastic error of +/- 100%:

### Expected VPUE

We assume the fishers can only form an expected VPUE at a given location based on their past observations and those of the fleet. Additional assumptions are required in this case:

* 1. Spatial extent of the expectations: we assume that fishers form their expectations on either a refined grid of 1x1 NM or 10x10 NM (1NM = , 115,200 alternatives for a x space)
  2. Temporal extent of the expectations: we assume that fishers use the VPUE records of:
     1. the past 30 days
     2. the past 30 days around the same date the year before
  3. We assume the following expected VPUE (not distinguishing individual records from fleet records):

With: ,,

, and being a dummy function valuing 1 when the argument (average of historical VPUE records) is not available.

* 1. To initiate the historical records we use the draws from one year of simulated positions under the assumption that fishers have a perfect knowledge of the VPUE maps and we discard the subsequent first year of simulated positions
  2. When a fishing site has been chosen by a fisher, its recorded position is set by randomly picking lon and lat coordinates within the site spatial extent.

|  |
| --- |
|  |
| **Figure D.2.** Maps of simulated VPUE for four different points during the year for each of the three level of spatial heterogeneity we assume. |

### Vessels and fishing days

In the Monte Carlo analyses, we consider a fleet of identical vessels (with the same preferences) fishing during 3 periods (years) . Vessels go out fishing all through the period but follow a probability density function proportional to the mean VPUE through the space (Figure D.3).

|  |
| --- |
|  |
| **Figure D.3.** Probability function for drawing the fishing days. |

Since vessels are identical, we make draws (with replacement) of vessels from a uniform distribution.

### Fishing locations

Whereas a continuous approach could have been taken to generate draws of fishing locations (e.g., approximating the spatial probability distribution function using the Metropolis-Hastings algorithm), we chose to take a discrete approach consisting in discretizing space at a refine scale and generating a field of random errors.

Thus, for each cell of the grid, a random error is drawn from a Gumbel distribution centered in 0 and with a scale of 1. We assume that errors are independent and identically distributed which implies that the difference of two different error terms follows a logistic distribution with mean and a standard deviation . Since only differences in utilities matter, we scale the magnitude of the error terms relative to the distribution of the differences in the deterministic utilities across a whole season . Given the high skewness of the distributions of the utility differences (Figures D.4.a and D.4.b), we take the bottom 0.1% of the differences [[2]](#footnote-2) – which is 116 alternatives in average per day - as the reference point.

|  |  |
| --- | --- |
|  |  |
| **Figure D.4.a.:** PDF of the differences in the deterministic utilities, for each level of spatial heterogeneity of the VPUE and over a whole season. | **Figure D.4.b.:** PDF of the temporal distribution of the bottom 0.1% of the differences in the deterministic utilities, for each level of spatial heterogeneity of the VPUE and over a whole season. |

After having tested for differences of error terms () having a standard deviation of 10%, 15%, 20%, and 25% of the bottom 0.1% of the differences in the deterministic utilities, we chose to scale the standard deviation of as 15%, having found that starting at 15%[[3]](#footnote-3) and increasing the magnitude led to more and more random simulated fishing positions inducing significant drops in the capacity of the estimated DCM to fit the data and to recover the true parameters up to a scale factor (Figure D.4.c).

|  |
| --- |
|  |
| **A** |
|  |
| **B** |
| **Figure D.4.c.** Goodness-of-fit (Panel A) and estimated marginal utility of the short-term component of the expected VPUE (Panel B) of 15 Monte Carlo draws estimated for magnitudes of the stochastic part of the utilities varying from 10% to 25% (panels from left to right) of the bottom 0.1% of the distribution of differences in the deterministic part of the utility. Magnitudes of the error term higher than 15% lead to a significant drop in the capacity of the RUMs to fit the data, included at the proper spatial scale (level of spatial aggregation = 0). |

### Model estimation

We estimate the simple conditional logit model corresponding to the data-generating process, assuming that the weights for forming fishers’ expectations are unknown to the researcher, and distinguish configurations of information availability using dummy variables:

With:

* case 1: both short-term **and** long-term historical VPUE are available
* case 2: **only** short-term historical VPUE are available
* case 3: **only** long-term historical VPUE are available
* case 4: **neither** short-term **or** long-term historical VPUE are available

The hypothesis is that we should be able to recover the weights as well as :

## Applied case study

### Choice of the empirical specification

The selection process for the choice of the empirical specification for the applied case consisted of the following steps.

First, we undertook an extensive series of pre-analyses on the bottom longline section of the GoMRF to identify possible emerging patterns of the spatial and temporal dynamics of the fishery. Besides performing statistical descriptions of the fishery (e.g., vessel characteristics, targeted species, gear, revenue structure, homeports), we combined vessels’ trajectories with reported catches, revenues, and fishing effort per trip to produce monthly and yearly maps of catch rates, revenue rates and fishing activity. In addition to combining these maps with the history of management measures and our knowledge of the ecological dynamics of the fish resources (groupers and tilefishes), we ran correlation tests between reported statistical areas of fishing locations and vessel and trip characteristics as well as with landing prices. We confirmed the information in the literature that describes longline fishers as primarily single-gear, single-species fishers, and we did not find evidence of seasonal patterns regarding the location of fishing effort or the price for groupers and tilefish species. However, there did appear to be a break in the distribution of fishing effort after the implementation of new spatial restrictions in 2009 and 2010 and this motivated our choice of splitting the dataset into two periods. Overall, we did not find obvious environment-related patterns emerging, having the impression that habits and distance seemed to be the most important drivers of fishing locations, as has been reported in the literature.

Second, based on this pre-model analysis and an extensive literature review about discrete-choice models of fishing behavior, we selected a set of variables that we deemed the most relevant. Specifically, we retain: , the distance from one location to another ; the level of fishing effort the day before in each site for a given vessel and , for the entire fleet ; and the historical records of value per unit of effort in each site for a given vessel () and for the entire fleet (), and for a time window including either the 30 () or the 365 () days prior the fishing trip or the 30 days the year prior (), surrounding the date of the fishing trip.

Working with this set of variables, we explore 12 possible specifications of the GoM model (cf. Table E.1) that we estimated with both the 2007-2008 and 2011-2012 datasets for all the 9 levels of spatial aggregations described in the paper. We also assumed either constant or random coefficients (i.e., either considering a conditional logit model or a mixed logit model) in order to account for vessels’ heterogeneity.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Table E.1** Combinations of information signals considered for the specification of the expected VPUE | | | | | | | |
| Model # | Info. Source | Individual level | | | Fleet level | | |
| Time span | [t;t-30] (m-1) | [t;t-365] (y-1) | [t-350;t-370] (ym-1) | [t;t-30] (m-1) | [t;t-365] (y-1) | [t-350;t-370] (ym-1) |
| 1 |  | N | N | N | Y | N | N |
| 2 |  | N | N | N | N | Y | N |
| 3 |  | N | N | N | N | N | Y |
| 4 |  | Y | N | N | Y | N | N |
| 5 |  | N | Y | N | N | Y | N |
| 6 |  | N | N | Y | N | N | Y |
| 7 |  | N | N | N | Y | Y | N |
| 8 |  | N | N | N | Y | N | Y |
| 9 |  | N | N | N | N | Y | Y |
| 10 |  | N | N | N | Y | Y | Y |
| 11 |  | Y | N | Y | Y | N | Y |
| 12 |  | N | N | N | N | N | N |
| *Shaded line: Specification presented in the main body of the paper. t-30 for example means we used the information from the month prior.* | | | | | | | |

Figure E.1.a shows the goodness of fit for the 12 model’s specifications estimated with 2007-2008 data, by level of spatial aggregation and for either fixed (left panel) or random parameters (right panel). Results are similar using 2011-2012 data and comparing models’ prediction errors instead of models’ goodness-of-fit. Figure E.1.b. shows the AIC differences of the 12 models relatively to the lowest AIC each level of spatial aggregation. Model #11 has (almost) systematically the lowest AIC, despite having the largest number of variables (50).

Overall, we found very little difference between the conditional and mixed logit models, and between models distinguishing individual and fleet-level information. In order not to lose the reader with unnecessary complexity and to focus the paper on the issue of spatial aggregation, we chose to present only the parsimonious model (model #8) that we describe in the paper.

|  |
| --- |
|  |
| **Figure E.1.a.** Goodness-of-fit for the 12 model’s specifications estimated with 2007-2008 data, by level of spatial aggregation and for either fixed (left panel) or random parameters (right panel). Model #8 with fixed parameters is the model presented in the main body of the paper and is highlighted with the black line.  Computational limitations did not allow to present the estimates for the most refined level of spatial aggregation for all model specifications. Convergence issues in the estimation algorithm for the most complex specification (model 11) suggest that estimates for this specification may not correspond to a global optimum. |

|  |
| --- |
| Une image contenant carte, texte, capture d’écran  Description générée automatiquement |
| **Figure E.1.b.** AIC differences for the 12 model’s specifications estimated with 2007-2008 data, by level of spatial aggregation and for either fixed (left panel) or random parameters (right panel). Differences are computed relatively to the lowest AIC obtained for each level of spatial aggregation and each model type. Model #8 with fixed parameters is the model presented in the main body of the paper.  Computational limitations did not allow to present the estimates for the most refined level of spatial aggregation for all model specifications. Convergence issues in the estimation algorithm for the most complex specification (model 11) suggest that estimates for this specification may not correspond to a global optimum. |

### Complementary estimation assumptions

A couple of assumptions were necessary to estimate the RUM. To begin with, we chose a daily time scale for choice occasions. This time scale is a compromise between the most refined time scale that would be based on fishing sets (but that would be much more data intensive and require a more refined analysis of the VMS pings) - and the coarser time scale of trips. Although this latter time scale may be very well suited for single-day trip fisheries (such as the urchin fishery studied extensively by Smith, e.g., Smith 2002, Smith, 2005), it is not appropriate here given the average duration of a fishing trip is approximately of one week. However, as the resolution of models becomes more spatially refined, the assumption of the uniqueness of choice becomes sometime violated (thereby emphasizing once more the trade-off between the spatial resolution of models and estimation issues). We followed a standard assumption in the literature (Girardin et al., 2015) of designating the “chosen” alternative as the one where most of the fishing effort was allocated[[4]](#footnote-4). Effort, catches and revenues were re-assigned accordingly. Depending on the tessellations, between 4% and 7% of effort was re-assigned according to that process. In total, we obtained 6,406 and 2,944 unique choice occasions (for 2008 and 2012 respectively), defined as a combination of a logbook trip with a day of the year.

We assume that the fishing effort remains constant over each fishing trip. Whereas the number of hooks per line is clearly fixed for a trip, the number of fishing sets[[5]](#footnote-5) per day may vary from. However, we assumed it did not affect the decision of where to fish and, when allocating effort on a daily basis, that the total number of fishing sets was homogenously distributed across the different days of the same trip.

Finally, we assume that the decision to go fishing and the decision on effort level fishing were independent from the decision of the fishing location. We tested those assumptions using 2008 logbooks and analyzed the correlation between fishing sites (reported for a given trip as the statistical area that yielded the highest revenue) and landing prices (a major driver for the decision to start a fishing trip or not) as well as the correlation between fishing sites and effort levels. In both cases, we found only a weak correlation.

## Complementary results

### Monte Carlo experiments

#### Absence of correlations between spatial indexes and model’s performance

Figure F.1.a, F.1.b and F.1.c show the results of the correlation analyses that we carried out between the spatial indexes of data heterogeneity that we considered – the Shannon entropy index and the Shannon equity index (see Section B of the Appendix) – and each model’s performance in terms of goodness-of-fit (Figure F.1.a), prediction capability (Figure F.1.b) and capacity to recover the model’s true parameters (Figure F.1.c).

|  |
| --- |
|  |
|  |
| **Figure F.1.a.** Goodness-of-fit of 600 Monte Carlo experiments in function of Shannon entropy (upper panels) and equity (lower panels) index. Panels reflect for the ten levels of spatial aggregation tested. The solid black lines show the results of a simple linear regression. |

|  |
| --- |
|  |
|  |
| **Figure F.1.b.** Prediction errors of 600 Monte Carlo experiments in function of Shannon entropy (upper panels) and equity (lower panels) index. Panels reflect the ten levels of spatial aggregation tested. The solid black lines show the results of a simple linear regression. |

|  |
| --- |
|  |
|  |
| **Figure F.1.c.** % error in the estimation of the short-term component of the expected VPUE for each of 600 Monte Carlo experiments with respect to Shannon entropy (upper panels) and equity (lower panels) index. Panels reflect the ten levels of spatial aggregation tested. The solid black lines show the results of a simple linear regression. |

#### Average Marginal Effects of explanatory variables

Here we show the average marginal effects of explanatory variables in estimating Eq (1).

|  |
| --- |
|  |
| **Figure F.1.d.** Average marginal effects (AME) on choice probabilities of the distance variable estimated in the Monte Carlo experiments. Effects are computed for an increase of one standard deviation. The box edges are the 25th and 75th percentiles of the distribution of the AME and whiskers are located at +/- 1.5 times the interquartile ranges. |
|  |
| **Figure F.1.e.** Average marginal effects (AME) on choice probabilities of the four components of the expected VPUE estimated in the Monte Carlo experiments. The top and bottom left panels show, respectively, the short-term and long-term components of the expected VPUE when both information types are available, the top right panel shows the case when only short-term information is available, and the bottom right panel shows the case when only long-term information is available. Effects are computed for an increase of one standard deviation of the corresponding variable. The box edges are the 25th and 75th percentiles of the distribution of the AME and whiskers are located at +/- 1.5 times the interquartile ranges. |

#### Root-mean squared-errors of parameter estimates

|  |
| --- |
| **Table F.1** Root-mean squared errors (RMSE) of parameter estimates for each level of spatial heterogeneity and level of spatial aggregation. RMSE are computed taking the mean over all the 600 simulations of the squared difference between a parameter estimate and its true value. |
| |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Het.** | **Parameters** | **Spatial aggregation** | | | | | | | | | | |  |  | **0** | **1.4** | **2.2** | **2.8** | **3.2** | **4.6** | **5.4** | **6** | **6.4** | **6.8** | | High |  | 1821 | 957 | 12 | 3 | 1 | 165 | 10610 | 2676 | 17509 | 112284 | |  | 110 | 323 | 9 | 9 | 9 | 26 | 1743 | 150 | 844 | 18267 | |  | 570 | 413 | 24 | 23 | 21 | 965 | 2904306 | 8561 | 81660 | 150951 | |  | 3803 | 2068 | 28 | 7 | 3 | 810 | 35144 | 6175 | 25791 | 432064 | |  | 6199 | 4316 | 33 | 630 | 5 | 27 | 1137 | 1005 | 88075 | 142373 | | Med. |  | 0 | 0 | 1 | 1 | 2 | 23 | 307 | 171 | 14399 | 20259 | |  | 9 | 9 | 9 | 9 | 9 | 11 | 46 | 28 | 2555 | 2475 | |  | 11 | 12 | 13 | 15 | 20 | 100 | 1443 | 773 | 66996 | 83385 | |  | 1 | 1 | 1 | 2 | 5 | 62 | 806 | 366 | 45013 | 23347 | |  | 8 | 7 | 6 | 7 | 10 | 109 | 4492 | 12636 | 155390 | 88426 | | Low |  | 0 | 0 | 1 | 2 | 3 | 23 | 586 | 23325 | 1360 | 2956 | |  | 9 | 9 | 9 | 9 | 9 | 11 | 52 | 4574 | 127 | 401 | |  | 11 | 11 | 14 | 16 | 22 | 106 | 1758 | 192534 | 5271 | 13058 | |  | 1 | 1 | 2 | 3 | 4 | 58 | 937 | 141904 | 2697 | 9316 | |  | 6 | 6 | 5 | 5 | 8 | 612 | 1468 | 60123 | 9892285 | 174352 | |
| |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Het.** | **Parameters** | **Spatial aggregation** | | | | | | | | | | |  |  | **-4.6** | **-3.2** | **-2.4** | **-1.8** | **-1.4** | **0** | **0.8** | **1.4** | **1.8** | **2.2** | | High |  | 81400 | 18 | 1970 | 2672 | 7350 | 1 | 1063 | 140 | 286 | 3279 | |  | 1756 | 28 | 1017 | 2580 | 3044 | 9 | 317 | 23 | 67 | 419 | |  | 3 | 13 | 2695 | 4650 | 14628 | 12 | 607 | 321 | 1121 | 7040 | |  | 23 | 15 | 8689 | 2820 | 5045 | 2 | 913 | 523 | 403 | 5553 | |  | 14252 | 115 | 64243 | 18654 | 62282 | 6 | 845 | 27 | 384 | 2709 | | Med. |  | 1114 | 31 | 5612 | 226 | 328 | 0 | 1275 | 20 | 1063 | 2916 | |  | 633 | 43 | 9706 | 110 | 63 | 9 | 141 | 10 | 218 | 451 | |  | 1 | 0 | 11798 | 98 | 273 | 11 | 1785 | 85 | 1130 | 13661 | |  | 15 | 13 | 8776 | 231 | 439 | 2 | 3537 | 56 | 1603 | 89888 | |  | 3418 | 1151 | 35286 | 2932 | 3861 | 4 | 10904 | 51 | 17988 | 2710 | | Low |  | 2086 | 37 | 54 | 1948 | 3054 | 1 | 22876 | 1526 | 3098 | 371 | |  | 265 | 26 | 77 | 186 | 347 | 9 | 1911 | 298 | 416 | 61 | |  | 1 | 7 | 217 | 969 | 2719 | 12 | 15931 | 4032 | 8780 | 1840 | |  | 15 | 31 | 28 | 666 | 4428 | 2 | 26396 | 2975 | 4335 | 1426 | |  | 6 | 6 | 5 | 5 | 8 | 612 | 1468 | 60123 | 9892285 | 174352 | |

### Applied case study

Here we show the average marginal effects of explanatory variables in estimating Eq (2).

|  |
| --- |
|  |
| **Figure F.2.a.** Average marginal effects on choice probabilities of the distance to fishing grounds, of the fishing effort of other vessels, and of vessel’s own fishing effort the day before estimated with GoMRF data. Segments indicate the 95% confidence interval of the corresponding point estimate. Effects are computed for an increase of one standard deviation of the corresponding variable. |

|  |
| --- |
|  |
| **Figure F.2.b.** Average marginal effects on choice probabilities of the four components of the expected VPUE estimated with GoMRF data. The top and bottom left panels show, respectively, the short-term and long-term components of the expected VPUE when both information are available, the top right panel shows the case when only short-term information is available, and the bottom right panel shows the case when only long-term information is available. Segments indicate the 95% confidence interval of the corresponding point estimate. Effects are computed for an increase of one standard deviation of the corresponding variable. |

1. The information about soaking time was highly unreliable (fishers in some cases reported either the soaking time for the entire trip or the mean soaking time per set) and thus were discarded. Complementary analyses exploiting the observer data showed that soaking times were fairly similar across fishers - about an hour – and did not affect the level of catches of trips. Therefore, we defined and computed fishing effort for BLL as the total number of hooks having been soaked during the trip (i.e., number of hooks per line multiplied by the number of sets). [↑](#footnote-ref-1)
2. i.e., the threshold for the 0.1% alternatives that are the closest to the alternative with the highest utility [↑](#footnote-ref-2)
3. In practice, we multiply the errors terms by . [↑](#footnote-ref-3)
4. When different sites had the same levels of effort, we randomly selected one of them. [↑](#footnote-ref-4)
5. i.e., the number of times the longline is soaked and hauled back. [↑](#footnote-ref-5)