Choice a Sets for Spatial Discrete Choice Models in Data Rich Environments

Robert L. Hicks
rob.hicks@wm.edu
Department of Economics & School of Marine Science
The College of William and Mary

Daniel S. Holland
dan.holland@noaa.gov
National Marine Fisheries Service
Northwest Fisheries Science Center

Peter T. Kuriyama
ptrkrym@uw.edu
School of Aquatic and Fishery Sciences
University of Washington

Kurt E. Schnier
kschnier@ucmerced.edu
Department of Economics
University of California, Merced

Abstract:

Failure to properly specify an agent's choice set in discrete choice models will generate biased parameter estimates resulting in inaccurate behavioral predictions as well as biased estimates of policy relevant metrics. We propose a method of constructing choice sets by sampling from specific points in space to model agent behavior when choice alternatives are unknown to the researcher, potentially infinite, and differ according to spatial and temporal factors. Using Monte Carlo analysis we compare the performance of this point-based sampling method to the commonly used approach of spatially aggregating choice alternatives. We then apply these alternative approaches to modelling location choice in the Pacific groundfish trawl fishery which has a complex spatial choice structure. Both the Monte Carlo and application results provide considerable support for the efficacy of the point-based approaches.

JEL: C25, C53, Q22

Keywords: Discrete Choice; Consideration Set; Random Utility Model; Spatial Discrete Choice

Choice Sets for Spatial Discrete Choice Models in Data Rich Environments

Abstract

Failure to properly specify an agent's choice set in discrete choice models will generate biased parameter estimates resulting in inaccurate behavioral predictions as well as biased estimates of policy relevant metrics. We propose a method of constructing choice sets by sampling from specific points in space to model agent behavior when choice alternatives are unknown to the researcher, potentially infinite, and differ according to spatial and temporal factors. Using Monte Carlo analysis we compare the performance of this point-based sampling method to the commonly used approach of spatially aggregating choice alternatives. We then apply these alternative approaches to modelling location choice in the Pacific groundfish trawl fishery which has a complex spatial choice structure. Both the Monte Carlo and application results provide considerable support for the efficacy of the point-based approaches.

Highlights

- Grid-point model has less bias in expected revenue coefficients in most cases
- Our empirical application showed higher predictive accuracy for grid-point models
- A variety of factors can influence the absolute and relative performance of the models

JEL: C25, C53, Q22

Keywords: Discrete Choice; Consideration Set; Random Utility Model; Spatial Discrete Choice

1. Introduction

A central component of any discrete choice analysis is the selection of alternatives that determine a decision agent's consideration set. To date, research surrounding consideration sets in discrete choice frameworks assumes that the researcher can observe, or reasonably approximate, the complete set of choice alternatives. The decision-maker's problem is then to choose the optimal choice alternative in the context of costly information gathering for site attributes. The vast majority of research on choice set definition begins with the problem of paring down the complete choice set to a subset of alternatives pertinent to a given choice occasion. This paring down process is justified by assuming that individuals don't have the time or resources to learn about the full choice set. Numerous studies have found that a failure to properly specify an agent's consideration set will generate biased parameter estimates leading to inaccurate behavioral predictions and policy metrics (e.g. welfare measures or elasticities) (Feather 1994; Haab and Hicks 1997; Hicks and Strand 2000; Hicks and Schnier 2010; Parsons and Hauber 1998; Parsons et a. 2000). This has prompted researchers to delve deeper into the cognitive processes that generate one's consideration set (Roberts and Negundani 1995; Shocker et al. 1991).

In discrete choice environments where choices are made repeatedly, and where technology is available for storing time series data at precise spatial scales, it is likely that the decision-makers effectively considered all possibilities. For these types of problems, we contend the predominant issue confronting the researcher isn't paring down the full set of available choice alternatives but defining what constitutes a choice alternative and what comprises the choice set. The development of an alternative, more behaviorally realistic, consideration set formulation process is integral to our ability to predict agent behavior and inform public policy. We propose a method of constructing choice sets by sampling from a set of specific points in space that can be used to model agent behavior when potential choice alternatives are effectively infinite and must be defined over a specified spatial and temporal region.

Commercial fishing is an activity in which agents make complex spatial choices in data rich environments. Commercial fishermen make dozens if not hundreds of such repeated choices each year, have onboard geographic information systems for collecting, analyzing and storing data, and collect information such as bathymetry, past locations fished, harvests, locations of other vessels fishing, current weather and water temperature, and average weather and currents. Researchers observe the vessel choosing a point over a very large open ocean area. Most applications of spatial discrete choice models on such data have completely ignored the rich spatial data available and have rather crudely aggregated sites using convenient latitude/longitude grids or large discrete areas defined for management purposes. This is often done because it isn't at all clear what the correct spatial scale should be or fine-scale location choice data is not available. However, these areas may encompass highly heterogeneous fishing locations in terms of species composition and density or feasibility of fishing (e.g. unfishable rocky areas). While computational capabilities now allow researchers to specify larger choices sets comprised of smaller discrete areas it may still not be possible to capture the fine scale heterogeneity that determines species distributions (and co-distributions) in many fisheries when choice alternatives are arbitrarily spatially distributed.

While our application focuses on a fishery, there are other applications where choices are made at a fine scale and fine scale data on attributes are likely to be available and drive decisions. Consider for example retail store location siting or the siting of mines or wells. For these cases, decision-makers have incentives to fully understand the choice set, characterize its extent and spatial variability and make decisions using this detailed data. They are likely to collect or have access to large amounts of spatially refined data on factors that might influence the relative utility of different choices. Researchers would see only the location chosen and often characterize these choices over arbitrarily defined zones or jurisdictional boundaries, but true choices may reflect finer scale attribute data. Similarly, the choice of what house to buy is likely to be affected by the attributes of the house itself (e.g., square feet, number of bathrooms, etc.) and of neighborhood characteristics but probably also the characteristics of the houses nearby, particularly neighboring houses.

To better characterize choice alternatives and consideration sets, we consider an approach that samples points from a fine scale grid of specific locations which in totality may consist of thousands of locations. Sampling of these locations following McFadden (1978) and Parsons and Kealy (1992) can make computation tractable when the choice set gets very large without undermining the ability to identify parameters with relatively modest sample sizes (von Haefen and Domanski). The site attributes at each point are based on observations of choices within a specified time and distance of the choice alternative rather than simply observations that fall in a pre-specified discrete area which may not accurately describe conditions in that specific location. For example, if a potential choice location occurs near the corner of an arbitrarily defined discrete area, commonly used approaches will base expected utility on past realizations within that discrete area without regard to their proximity to the specific location choice[1]. In contrast our point-based approach would limit information informing the expected utility of the choice alternative to other choice realizations near that location, including information from what might have been in an adjoining arbitrarily defined area had traditional aggregate areas been used.

[1] The common approach in the commercial fisheries literature is either to be blessed with a fishery where a discrete set of points are fished or these points naturally occur in groups one might consider a site (e.g. Smith 2002) or to arbitrarily divide fishing areas into grids (e.g. Hicks and Schnier 2008, 2010; Hicks, Kirkley, and Strand 2004; and Miastean and Strand 1998). In some cases aggregation is determined by the data, .e.g. if only management areas are recorded for locations (e.g. Holland and Sutinen 2000).

The point-based method is likely to be advantageous when the resource is very patchy relative to more coarsely defined discrete areas. The advantage of our method relative to traditional approaches is likely to diminish as the scale of the traditionally defined discrete areas becomes finer or to the degree the traditional area boundaries can accurately map out the choice delineations that the agents consider (assuming the delineations are consistent across agents). As the surface of the characteristics that determine utility become smoother (e.g. density of a resource changes linearly over space rather than changing erratically within the bounds of the discrete areas) the difference between the methods should also diminish. If the analyst or the agents lack information on choice characteristics at a fine scale the point-based model is unlikely to offer a substantial advantage over a traditional discrete area method.

Using Monte Carlo (MC) analysis we compare the performance of this point-based sampling method to a commonly used method that aggregates potential choice locations into pre-defined discrete areas and thus simplifies, but potentially misspecifies, the geo-spatial complexity of an agent's choice set to facilitate estimation. Our results indicate that our proposed point-based choice set model performs well in comparison to a more traditional approach with discrete areas and has the potential to improve spatial discrete choice modeling in settings with fine-scale spatial heterogeneity. To illustrate the advantages of our proposed model we apply it to the Pacific groundfish trawl fishery, which is characterized by a complex spatial choice structure.

In the following section we contextualize our model in the existing discrete choice literature. In section three we outline the proposed model and its similarity and differences with traditional discrete choice models. Section four presents our MC analysis where we compare our proposed point-based sampling model to a traditional random utility model (RUM) that aggregates spatial choices within discrete areas. Section five presents the results from our application of the model in the context of the Pacific groundfish trawl fishery. The final section contains general conclusions.

2. Literature Review

To better understand the cognitive process that generates the choice set it is convenient to frame the activity as a two-step process (Manski 1977). In the first stage one forms their choice set, often times referred to as a consideration set, from the set of all potential alternatives available. In the second stage one looks at all the alternatives that are contained in their parsed down choice set and selects the alternative that maximizes their utility. The later step follows the random utility model (RUM) (McFadden 1978). A central issue in applied work is the assumptions made about how agents conduct the first stage of the decision process. Often times this stage is conveniently ignored, but a number of researchers have endeavored to address this issue within the literature.

The predominant methods used to reduce or restrict the choice set of agents have been site aggregation (Feather 1994; Parsons and Needelman 1992; Parsons et al. 2000), familiarity-based restrictions (Peters et al. 1995; Parsons et al. 1999; Hicks and Strand 2000), distance-based measures (Parsons and Hauber 1998; Whitehead and Haab 1999) and time-based restrictions (Banzhaf and Smith 2007), or endogenous choice set formulation (Manski 1977; Swait and Ben-Akiva 1987a, 1987b; Andrews and Srinivasan 1995; Roberts and Lattin 1991; Bronnenberg and Vanhonacker 1996; Desarbo and Jedidi 1995; Swait 2001b; Basar and Bhat 2004; Haab and Hicks 1997; and Hicks and Schnier 2010). It is important to note two things about these approaches. First, these papers argue that the issue of properly defined consideration sets is a necessary step in obtaining consistent parameter estimates; and second, the consideration set formulation process happens because information gathering for alternative specific attributes is costly for agents.

The process of site aggregation is motivated by the need to reduce the choice set to a tractable number of alternatives. However, Parsons and Needelman (1992) illustrated that site aggregation may generate biased welfare estimates with the bias increasing as one further aggregates sites. The bias arises because as one aggregates sites we do not account for the degree of heterogeneity within the site as well as the number of sites being aggregated within the empirical model (Ben-Akiva and Lerman 1985). Parsons et al. (2000) investigated the impact that different aggregation schemes have on welfare estimates in a region containing 814 unique lakes when the most highly visited five sites are removed. They considered a baseline model with the choice set consisting of nearby lakes and a random sample of lakes outside the region and compared it to alternative models that aggregated the choice set according to regions. They found that their welfare estimates varied depending on the modeling assumptions. This is similar to the findings in Feather (1994), who highlighted that the only way we know whether or not our estimates are truly biased is if we know the "true" choice set perceived by the decision agent.

The benefit of using site aggregation to reduce the choice set is that the geographic extent is preserved, however this does come at a cost of not controlling for heterogeneity within the site and the number of locations being aggregated (Ben-Akiva and Lerman 1985). An alternative tool that has been used is to reduce the geographic extent of the choice set using a travel distance metric that limits the number of sites based on what is a feasible distance to travel. Parsons and Hauber (1998) investigate alternative spatial boundaries of a consideration set defined over the travel distance from one's location. They varied the size of the choice set from sites that are 0.8 to 4 hours away from one's location and found that the models stabilized in a consistent range once the spatial choice set was defined at 1.6 hours or greater. Whitehead and Haab (1999) investigate the impact of using either a distance-based or potential catch filter to define one's choice set in a recreational fishing demand model. They found that their welfare estimates did not vary substantially as they altered their filter but using a potential catch filter altered the value of site quality preferences, mainly the value of expected catch at a site. More recent research by Scrogin et al. (2010) developed an empirically-based criterion for the distance selection metric. They construct a two-stage model that first estimates a stochastic frontier model where site quality is estimated using a distance cost frontier. Based on the efficiency rankings that result from the stochastic frontier model they select sites to be contained in the choice set using their efficiency scores. Although this model is interesting it does require that all the potential elements in the choice set are observable. This is a limitation of all of the aforementioned methods and something that we directly address in our proposed choice set model.

Researchers have also compared the different methods of restricting the choice set to investigate the impact they have on parameter and welfare estimates. Hicks and Strand (2000) investigated the sensitivity of welfare estimates to three alternative choice set specifications: full choice set, familiarity-based choice set and a distance-based choice set. They find considerable heterogeneity in results when using these different methods. However, the distance-based choice set asymptotically approaches the full choice set as the distance band increases. They concluded that if a researcher believes that an individual's choice set does not contain the full choice set they will most likely need to elicit this information from them (i.e., generate a familiarity-based choice set) or endogenously estimate it. Banzhaf and Smith (2007) conducted a rigorous meta-analysis on welfare estimates as one alters their definition of the choice set within the housing market to investigate the impact of air pollution using a hedonic property model. They define their choice sets based on three different variables: geographical scale, an agent's budget set and the length of time one searches for a home. This resulted in 128 different "choice set" definitions and for each of them they constructed a welfare estimate and then conducted a meta-analysis of their estimates. They found that the assumptions regarding the choice set can alter welfare estimates, but the estimates are less sensitive to assumptions regarding the time dimension. They hypothesize that this is true because time is not correlated with the other attributes of the model.

A limitation of all of the models discussed thus far is the assumption that the choice set is finite and tractable. In the case of spatial choice models in a number of environmental settings, including fisheries, the choice set is virtually infinite as space can be continually divided to form different "alternatives." This would render the standard estimation procedures discussed above intractable. Our point-based sampling approach attempts to circumvent this limitation by sampling from an evenly spaced fine-scale grid of points within the boundary of observed activity of the fishing fleet. Our model is similar to the time-space prism model (Yoon et al. 2012) of restricting the choice set, however we maintain that the choice alternatives form a continuous surface of spatial locations and we define space using more than one dimension in our empirical application (we incorporate depth). Each choice location draws on information from fishing events that occurred nearby in space and time and in the depth zones. Two choice locations may also utilize the same information (e.g. two points close to each other may draw on information a fishing event that occurred between them). Furthermore, Yoon et al. (2012) do not formally estimate a spatial choice model using their proposed time-space prism in order for them to investigate its impact on spatial choice modeling. Therefore, our MC analysis and application of our proposed model provides a unique contribution to the choice set (consideration set) literature.

3. Grid Point-Based Sampling Models

To motivate the development of the grid point-based model (hereafter referred to as the "Grid Model"), we will first outline the model in a generic sense as applicable to our MC analysis and later turn to additional features that are unique to our application and that further highlight the importance of its use in complex spatial choice structures. The foundation for our Grid Model is the Random Utility Model (RUM; McFadden 1978). Other than how choice sets are determined, our application of this model differs little from those used numerous times in environmental economics literature (e.g. Dupont 1993; Eales and Wilen 1986; Hicks, Kirkley and Strand 2004; Hicks and Schnier 2008, 2010; Holland and Sutinen 2000; Mistaean and Strand 1998; Smith 2002; Smith and Wilen 2003). Define $\boldsymbol{X}_{ijt}$ as vectors of individual, site, time-specific attributes related to site-specific payoffs (e.g. distance, revenues, etc.). Following the usual approaches, let ε_{ijt} be the vector of individual, site, and time-specific factors known to the decision maker i but not to the researcher. Given the preference parameters, β_x, assumed to be homogenous across decision makers, we can write individual $i's$ payoffs of choosing location j at time t as

$$(1) \qquad V_{ijt} = \boldsymbol{X}_{ijt}\beta_X + \varepsilon_{ijt}$$

The individual will choose location k having the highest payoffs,

$$(2) \qquad V_{ikt} > V_{ijt} \; \forall \; k, j \in \; S^i$$

where S^i is the choice set for individual i. Assuming that ε_{ijt} is distributed as iid generalized extreme value (GEV) I, from the researcher's perspective, the probability of observing individual i choosing location k at time t is

$$(3) \qquad Prob_{ikt} = \frac{e^{V_{ikt}}}{\sum_{j \in S^i} e^{V_{ijt}}}.$$

The choice set in the RUM outlined is defined by the elements in S^i, and the literature on restricting the choice set has predominately focused on reducing the number of alternatives in the denominator of Equation 3. The rationale for these approaches is that agents (i.e., individuals selecting a recreational site) are unlikely to be aware of every choice alternative, and therefore distance (e.g. Parsons and Hauber 1998, Hicks and Strand 2000), econometric techniques (Haab and Hicks 1997), or actual data collected on what sites were considered (Peters et al. 1995; Hicks and Strand 2000) can be used to estimate or refine the choice set before or during estimation.

The argument that decision agents are unaware of or have no incentive to gather information about potential choice locations in S^i is weak in settings where the decision agent makes repeated choices over a continuous space. Furthermore, in a number of these choice settings the decision agent often collects detailed spatial data on the observable characteristics, X_{ijt}, and are likely to have the ability to discriminate sites at a very fine, if not infinite, spatial scale. Ideally, one would define the choice set as all potential locations, both chosen and not chosen, but this would invariably generate a choice set near an infinite number of choices and render itself intractable even with modern computational capabilities. This is because as the spatial scale of what is considered a site becomes smaller, in the limit, precise geo-coordinates become sites and we encounter a curse of dimensionality in the spatial choice set.

With that in mind, we draw on McFadden's (1978) sampling approach for choice set formulation (and implemented by Parsons and Kealy (1992)) for making model estimation tractable. It is important to note that by sampling from the choice set, this approach assumes that all alternatives being sampled from are considered by the agent and the choice set is reduced purely for the sake of computational tractability. The two key assumptions are that the randomly drawn choice set, D, satisfy

(4) $Prob(D|k) = Prob(D|j)$ [The Uniform Conditioning Property]

(5) $Prob(D|j) > 0 \;\; \forall\, j \in D$ [The Positive Conditioning Property]

These properties presume that the geography and definition of all "sites" is known with certainty. Unlike the Parsons and Kealy (1992) application, where the set of choice alternatives (in their example, small lakes in Wisconsin) is known a priori, we know neither the locations considered to be an alternative, the size of these locations, nor if all decision agents agree on what constitutes a choice alternative. However, Parsons and Kealy (1992) fall back on the randomly drawn aspect of the choice set to argue that these properties are satisfied. We do the same here, except that we take an additional step and define potential locations as precise geo-coordinates. We approximate the full set of choice alternatives on a similarly fine scale by sampling from a finely spaced grid of points. Covariates for each of these grid points are defined by nearby outcomes (both in a geographic and temporal sense).

To contextualize the sampling approach we use, consider the geospatial activity of our empirical application illustrated in Figure 1 using fishing location choices (actual choices depicted as stars) made relative to a uniformly specified grid (grey and black dots). Taking a random sample of the uniform grid, we obtain a sampled choice set plus one closest to the location actually chosen (the black dots). The sampled grid points represent approximations of potential choices over a continuous set of approximately infinite choices that could be made by the decision maker, with the darker shading off the coast capturing alternative bathymetric contours which further exacerbates the infinite dimension of the potential choice set. The black dots illustrate a set of randomly sampled actual points comprising the choice set D^{it} for a randomly drawn vessel from the fleet currently in the location denoted by the black square and observed choosing for example, the grey diamond. A different choice set is randomly sampled (from the set of all observed choice locations) for every different discrete choice (i.e., vessel-specific tow in our application). In our empirical application the grid points are precise locations (e.g. points with specific latitude and longitude) but on a regular grid spaced 3.5 miles apart. This distance is shorter than the length of the vast majority of tows. The sampling is still restricted to areas open to fishing and within a feasible distance (which is delimited by observations of the geographic range of fishing by the port-based fleets modelled).

4. Monte Carlo Analysis

In our Monte Carlo (MC) analysis, we compare the grid-point sampling approach to a traditional RUM that aggregates grid points with an arbitrarily defined coarser grid overlay that divides the decision space. For the point-based sampling approach the choice set is constructed by randomly selecting sites within the decision space.

4.1 Data Generation Process

Before discussing the results of our MC analysis, we will briefly outline the data generation process utilized to test our proposed model. The data generation process utilizes a spatially patchy distribution of a revenue generating resource over a 64x64 grid (4096 individual cells). The grid of 4096 is sufficient to allow us to compare discrete area models with a range of scales. This revenue generating resource could be the spatial distribution of fish species, but in other settings might capture spatial heterogeneity of an oil reserve/precious metal, or spatial attributes of recreational sites, to cite a few potential applications. To begin the data generation process we first randomly seed a specific number of cells, defined as “clumps” (CL), and then allow the resource to diffuse from those seed cells for eight periods at a specified diffusion rate, δ. Based on the number of “clumps” selected and the rate of diffusion, this generates different spatial surfaces of the resource base that define how patchy or concentrated the resource is in our analysis. While our MC is designed with fisheries in mind, the clumps parameter determines whether there are lots of very small good locations or fewer larger ones, while the diffusion parameter determines how quickly density of the resource (or the value generated at a cite) changes as you move away from it a given location. These characteristics might also be true for other applications (e.g. mineral deposits, recreational fishing sites, recreational boater on-water site choice, back-country camping site choice, business locations, residential neighborhood quality, etc.). Each of the randomly drawn seeds in the predefined clumps was drawn from a normal distribution, $N(0, \sigma^Z)$, and used in our data construction for the variable Z^1_{ijt} described below. Figure 2 provides a graphical illustration of the resource's spatial distribution for a few of the different parameter combinations utilized in our MC analysis.

Once this initial resource surface is determined we generate a panel data set (termed "logbook data" below) consisting of 50 decision agents each making ten spatial discrete choice decisions. This results in a panel data set of 500 observations for each MC. In the context of a fishery this would represent 50 vessels each making ten different spatial location decisions. For each unique spatial decision, the decision agent selects the location on the grid with the highest expected utility,

(6) $V_{ikt}(\beta, Z, \epsilon) > V_{ijt}(\beta, Z, \epsilon) \ \forall k, j \in S^i$,

where $V_{ijt}(\beta, Z, \epsilon) = \beta_1 Z^1_{ijt} + \beta_2 Z^2_{ijt} + \epsilon_{ijt}$.

Within the indirect utility function, V_{ijt}, i represents the decision agent, j represents the spatial location and t represents the time period. Z^1_{ijt} is the expected revenue generated from the spatially distributed resource and Z^2_{ijt} is the distance one has to travel from their current location to obtain the resource. The error structure, ϵ_{ijt}, is a GEV error appended to the indirect utility function that is assumed to be known by the decision agent but not the researcher. On the first decision made by the decision agent we randomly assigned them a starting point within the grid, with all future distances calculated based on their current location.

In order to generate data that can be used in the MC analysis we apply a commonly used spatial aggregation heuristic (aggregating choice sites within regular grids of various sizes) with additional random noise to generate uncertainty in the data. When generating the set of initial location choices that make up the logbook data (that defines choices at precise points in the choice space) agents utilize to form expectations, a randomly distributed error term $N(0, \sigma^E)$, is added to the true density of the resource surface at each location. Locations are chosen based on this expected return observed with error and the true distance to the location according to the utility specification and multinomial probabilities in equations 2 and 3. This results in a set of location choices that includes some locations outside the areas of highest density. The actual return realized from these spatial choices (that is then recorded in the generated logbook data) is adjusted by adding another randomly distributed error term drawn from the normal distribution, $N(0, \sigma^A)$. This creates a random error between the expected return that motivated that particular location choice and the actual return that is reported in the logbook and is used in developing expectations for utility of subsequent choices. We define this actual return variable as Z_{ikt}^{1A}, where *k* is the chosen site made by decision agent *i* in time period *t*. Information on the actual return is then used in the RUM model to construct the expected returns from visiting a site by using temporally lagged observations of Z_{ikt}^{1A} within a defined spatial distance of an individual site, or coarser spatial grid overlay in the context of the traditional RUM, and then calculating the average of these observations. In the MC we used a five-period lag (i.e. observations from the prior five choice periods) for the calculation and a spatial buffer that utilizes only the adjacent cells for the grid-point model. In the traditional RUM all averages were constructed at the zone level. The travel distance variable, Z_{ijt}^{2}, was calculated using the observed sequence of choices made with the starting point for the first spatial decision being randomly selected from among the 4,096 individual cells and using unit measurement at the grid-point level. The grid of 4096 was sufficient to compare a range of scales, and this range is also reflective of the scales in the empirical application. There are other parameters that could be varied, but we focus on those we thought likely to be most important (scale, patchiness of resource, and various forms of noise). Distance is assumed to be known without error.

The preference parameters, contained in the vector β, were initialized at one for the spatially distributed resource, $\beta_1 = 1$, and negative one for the distance parameter, $\beta_2 = -1$. To generate spatial choices that were less correlated with the spatial distribution of resource, we varied the GEV scale parameter, ζ, which captures the appended GEV error structure to the indirect utility function illustrated in Equation 6. The GEV scale affects the relative weight of the model parameters versus the error term. For example, with ζ=20 the effective parameter values for revenue and distance are 0.05 and -0.05, respectively. When ζ=30 the values are 0.033 and -0.033, respectively. As the GEV scale parameter, ζ, increases, the distribution of location choices becomes less closely tied to the highest density of the resource base.

In the grid-point model, the RUM data for each observation includes distance, Z_{ijt}^{2}, and expected spatial returns, Z_{ijt}^{1}, for a choice set of 50 potential locations including 49 randomly chosen grid points and the actual location chosen. Increasing the choice set size beyond 50 has no appreciable effect on coefficient estimates and bias but slows the MC. The traditional RUM model uses the same underlying spatial choice data and process to generate the data contained in Z_{ijt}^{1} and Z_{ijt}^{2}. However, an aggregation factor is specified to create a coarser grid which is overlaid on the fine grid. This coarser grid defines a set of discrete areas. The aggregation factor is set at 16, 64, or 256 dividing the grid of 4,096 cells into a coarser grid of 256, 64, or 16 discrete areas respectively. For each observation the choice set consists of all of the discrete areas. Distances for the traditional RUM model for each potential location choice are the distances from the current location to the central grid point of each discrete area. Expected returns for each potential choice is a simple average of all observations that fall in that coarse grid cell within the prior five decision periods. This is consistent with the time window used in the grid-point models, but the definition of "location" is significantly larger.

We evaluate and compare the performance of the different choice set models for a full combinatorial experimental design of seven different variables for the MC analysis. Table 1 outlines the set of parameters used in our data generation process. This generates a total of 864 different sets of MC model variables. For each combination of MC variables we ran 200 replications using the same spatial resource surface and choice data generated to estimate each of the RUM models creating 172,800 replications in total.

4.2 Monte Carlo Performance Metrics

To compare the two models we explore both parameter bias and measures of predictive accuracy. To measure bias we calculate the average difference between the true parameter in our data generation process and that estimated in our MC analysis. To measure the predictive accuracy of our models we utilize four different model performance metrics. All metrics except the distance-based metric attempt to examine the predictive accuracy of the model in terms of how well each choice set model predicts activity into the traditional choice set. Consequently, if the grid-point model places high probabilities inside of zones defined by the traditional choice model then the predictive accuracy can be compared from the most granular to coarsest spatial resolution.

For each decision agent in our MC define the predicted site choice as,

(7) $$\hat{Y}_{it} = \underset{j \in S^i}{\operatorname{argmax}} Prob_{ijt}$$

Note that the S^i will vary over the choice set method utilized. For the traditional choice set model S^i will be defined over the coarser grids used to divide space, whereas for the grid-point models it will be individual points in the uniform grid. Note that for the grid-point model the predicted choice is based on the expected utility of <u>all</u> locations given the estimated parameters and relevant data, not just the fifty locations sampled during the estimation process. Based on this representation we estimate the following four prediction methods.

<u>Prediction Method 1:</u> The method, referred to as the Correct Prediction (CP) method, calculates the score

(8) $$CP = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{(\hat{Y}_{it} \in z_{it})}{NT}$$

where Z_{it} is the site actually chosen by individual i in time period t and $(\hat{Y}_{it} \in z_{it})$ is a logical operator that takes a value of one if the predicted site is the actual site chosen by the decision agent. The metric, CP, is simply the percent of sites that are chosen correctly by the model.

Prediction Method 2: The method, referred to as the Correct Prediction Summed (CPS) method, calculates the total probability within the chosen site, z_{it}:

(9) $$Prob_{izt} = \sum_{j \in z} Prob_{ijt}$$

This metric sums the estimated probabilities within an aggregate zone and based on that then calculates the prediction based on that summed probability,

(10) $$\hat{Y}_{it}^{A} = \underset{z \in Z}{\operatorname{argmax}} Prob_{izt}.$$

We denote this prediction with the superscript A to make explicit that we are predicting to the aggregate zone level. The previous method, CP, focuses on the point having the maximum probability (for the grid-point method). Similar to before the CPS score is calculated as follows,

(11) $$CPS = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{(\hat{Y}_{it}^{A} = z_{it})}{NT}$$

where z_{it} is the chosen aggregate zone.

Prediction Method 3: This method, referred to as the Probability Mass (PM) approach, uses the total probability mass in the chosen area defined by the traditional model, irrespective of whether it was the highest predicted probability, and reports the mean over the sample. For this and the PM metric below it is possible that more than one location fell within the correct discrete area and the probabilities for those areas would be summed. This metric is the following,

(12) $$PM = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{Prob_{izt}}{NT}$$

Prediction Method 4: The metric, referred to as the Distance (D) approach, calculates the mean distance, D, from $\hat{Y}_{it}$ to the site actually chosen, z_{it}:

(13) $$D = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{D(\hat{Y}_{it}, z_{it})}{NT}$$

4.3 Monte Carlo Results

To evaluate the performance of the models we conduct a meta-analysis of the different MC experiments using the average performance metrics over the 200 replications for each of the 864 parameter combinations. In addition to comparing the performance average across all MC parameterizations, we use a regression analysis to determine what parameters of the data generation process influence model performance.

4.3.1 Parameter Bias

The data generation process initialized the true revenue parameter, β_1, to be 0.05 with ζ=20 and 0.033 with ζ=30. The true distance parameter, β_2, is -0.05 with ζ=20 and -0.033 with ζ=30. Table 2 provides the average parameter estimates and bias of parameter estimates (averaged across all combinations of MC experiment variables) as well as the average percentage bias in parameters, broken down by the GEV scale parameter. On average both models under-estimate the revenue parameter, β_1, (all biases are negative). The average absolute bias of β_1 is smaller in magnitude for the grid-point model relative to the traditional model (i.e. the traditional model estimate is closer to zero). The standard deviation of average parameter and bias measures (shown in parentheses in Table 2) demonstrate that parameter estimates and bias vary substantially across the 864 combinations of MC parameters. For example, the traditional model gets smaller with the aggregation level (see Table A1 in Supplementary Appendix A). The bias of $\beta_1$varies for different combinations of the MC experiment variables but is nearly always negative and for most combinations of parameters the absolute bias is smaller for the grid model (Figure 3). Although absolute bias varies, the percentage bias of β_1 is similar for the different levels of the GEV scale parameter.

For the distance parameter, β_2, the average bias (across all MC parameterizations) is negative for the grid model and positive for the traditional model. However, the absolute value and percent bias of the distance parameters for all three models is much smaller in magnitude than for the revenue parameter. Also, unlike the bias in β_1, the bias in β_2 for the traditional model varies in sign for different combinations of experiment variables (Figure 3).

To determine how the different data-generation parameters impact the differences in absolute value of bias across models, we regressed the differences in the absolute value of bias between models against dummy variable indicators for the data generation parameters. We run separate regressions of the two values of the GEV scale parameter, since it changes the true model parameters and thus the absolute effect of the MC parameters. The results from these regressions are shown in Table 3. A negative regression coefficient indicates that this MC variable increases absolute bias relatively more for the traditional model than the grid model (i.e., favoring the grid model) and a positive coefficient indicates the opposite.

First, consider the results for the revenue parameter β_1. The results indicate that a larger number of clumps used to initialize the resource distribution (which leads to a more broadly distributed resource surface) tends to increase the relative advantage of the grid model in accurately identifying the revenue parameter. Higher diffusion rates tend decrease the advantage of the grid model where the difference is significant. A higher distance scale parameter has opposite effects depending on the GEV scale parameter. Higher error in expected revenue when initializing the logbook data has no significant impact on relative performance. However, error in the actual revenue recorded (which creates a gap between expectations that drove choices and the outcomes of those choices recorded in the logbooks used to create expectations of modeled choices) tends to narrow the difference in bias, i.e. decreasing the advantage of the grid-point model. While larger error increases bias for both models it affects the grid-point model more, thus reducing its relative advantage (see Appendix Table A1). Higher aggregation factors do not affect grid-point model but increases absolute bias for the traditional model thus increasing the difference in absolute bias in favor of the grid-point model. It is notable that across all the 864 different sets of MC parameter combinations absolute bias in revenue parameter is lower for the grid model in comparison to the traditional model for a given set of MC parameters.

Next we consider how MC variables effect the difference in absolute bias of the distance parameters for the two models. Recall the average absolute bias for the distance parameter is much smaller than for the revenue parameter (Table 2). Again, a negative coefficient in Table 3 indicates the advantage of the grid-point model is increased, and positive the opposite. The highest clumps value for the smaller scale parameter favors the grid model but difference is small barely significant. The higher diffusion rate favors the traditional model with the higher GEV scale parameter. The higher distance scale favors the grid model while higher error in actual revenue favors the traditional model as it did for the revenue parameter. As with the revenue parameter bias, the grid-point model is unaffected by higher aggregation factors while absolute bias increases for the traditional model thus worsening its relative performance in terms of accurately estimating the distance coefficient. While the grid model tended to have lower absolute bias in estimation of the revenue coefficient for most MC combinations, that advantage does not hold for the distance parameter where the grid model has lower absolute bias for only 46% of MC parameter combinations.

We focused above on the difference in bias for the different models. Regressions of absolute bias for each model are provided in Supplementary Appendix A[2]. These indicate that the models are mostly affected in the same way by the MC parameters but some parameters affect one model more than the other. For example, higher numbers of clumps, the highest diffusion rate, and larger error in the actual

[2] Although the ratio of revenue and distance parameters is a relevant metric in practice since it may be used in welfare analysis, we did not compare models on the basis of the average of this ratio because parameter estimates for revenue individual realizations can become very small causing the ratio to become very large. These realizations with extremely large ratios dominate the average making it unsuitable for assessing relative performance.

revenue tend to make bias in the expected revenue parameter more negative (larger in absolute value) for both models but by different amounts that also vary with the scale parameter. Another clear result is that the higher aggregation factor, which increases bias for the traditional model but does not affect the grid-point model, tends to favor the grid-point model. It is also clear that the grid-point model is more impacted by larger error in the actual revenue variable narrowing its advantage at identifying the revenue parameter correctly. General results are less clear for the bias in the distance parameter, but average bias for all the models is much smaller so perhaps less consequential.

4.3.2 Prediction Metrics

The Traditional model performs slightly better on correct prediction (CP) with 25% correct prediction compared to 24% for the grid model (Table 4). Note that for prediction metrics for the grid model, predictions are based on estimated choice probabilities for all 4096 grid cells, not just the 50 sampled choice sets. To predict the correct site the grid point with the highest predicted utility out of all 4096 sites must fall in the correct traditional grid area. In contrast the traditional model is predicting choices over a set of between 16 and 256 choices corresponding with the discrete areas.

For other predictions metrics the grid model performs better. The grid model has higher slightly higher score for correct prediction summed (CPS) at 26% relative to 25% for the grid model. Thus, while they predict the correct area less often based solely on the highest probability choice, when predictions are based on the summed probabilities of choices in each discrete area the grid-point model predicts choices more often than the traditional model. The grid-point model has a higher average score for probability mass (PM) in the correct area (summed probability associated with all choices in the correct discrete areas) at 22% compared 17% for the traditional model. The grid-point model also has lower (better) scores on the distance (D) metric which looks at the distance between the centroid of the predicted and correct areas (using the traditional model areas).

As we did for bias, we ran regressions of the difference in performance of the models as a function of the MC parameters (Table 5). Since high scores are better for CP, a positive coefficient means that variable tend to favor the grid model. While the traditional model has the highest CP score when averaged across all parameterizations, the performance of the grid-point model improves relative to the traditional model as the clumps and diffusion rates and GEV scale increase. A higher aggregation factor also improves relative performance of the grid-point model relative to the traditional model. Error in expected catch has no significant effect on relative prediction but error in actual catch favors the grid model.

The effects of different MC parameters on relative CPS scores differ from the effect on CP (Table 5). For the CPS metric, higher clump values improve the performance of the traditional model relative to the grid models as a higher GEV scale parameter. A higher diffusion rate and the higher aggregation factor favor the grid model. Error in expected or actual revenue has no significant effect on relative prediction.

Table 5 also provides regression results for the difference in the PM scores for the alternative models as a function of MC parameters. The grid-point model had the highest PM score when averaged across all MC parameters, and relative performance of grid over the traditional model was increased with higher diffusion rates, a higher distance scale, and a higher aggregation factor.

The last performance metric we regress against MC parameters is the Distance (D) score which focuses on the implied distance from the predicted site selected and the actual site selected in the data. A lower D score indicates that the model predicts a choice closer to the true choice in the data set and is thus a better score, so a negative coefficient in the regression favors the grid model. A higher clumps setting tends to improve the performance of the traditional model relative to the grid-point models as does a higher diffusion rate, higher GEV scale and larger error in actual revenues. However, a higher aggregation factor worsens performance of the traditional model relative to the grid-point model in terms of the D metric as does a higher distance scale

5. Empirical Application of Proposed Grid-Point Model

To illustrate the grid-point model in an empirical application we apply it to a fishery with a spatially heterogeneous fishing landscape, the Pacific groundfish trawl fishery. The period modeled using this data coincides with implementation of individual fishing quota (IFQ) management regime. Individuals were allocated quota shares for 28 groundfish stocks and stock complexes as well as individual bycatch quotas for Pacific halibut. Fishers must cover all catch of IFQ species with quota pounds (the annual form of quota) and 100% observer coverage ensures all catch is accounted for whether landed or discarded. Fishers may buy and sell quota, and there is an active quota market for many species. Quota prices for some rockfish species taken incidentally have quota pound prices that exceed the ex-vessel value of the fish creating incentives to avoid catching them (Holland 2016). We account for the economic tradeoffs induced by the IFQ system by netting IFQ costs from ex-vessel value for all IFQ species when estimating expected revenue from alternative location choices.[3]

[3] We originally tried estimating models with variables representing expected bycatch explicitly since anecdotal evidence suggested bycatch avoidance was a major driver of location choice. We found that signs on these variables were sometimes contrary to expectations (e.g. positivity) which we suspect is due to strong correlation between expected catch of target and bycatch species and our inability to model expected bycatch effectively at a fine enough scale since bycatch is a relatively rare event in some instances. Consequently we opted to simply include quota costs

Catch rates for both targeted species and incidentally caught rockfish species that fishers may try to avoid have been shown to be correlated with depth and latitude and can be very patchy (Holland and Jannot 2012). Because changing depths requires changing the length of tow cables, trawlers will generally avoid large changes in depth during a tow and will stay along depth contours. Depth can change dramatically over small distances in many areas when going against the contour. Fishers generally return to the same port from which they departed (along this coastline that mostly runs North-South), but may make several tows during a trip.

In this setting, the traditional discrete area choice set approach may be problematic, particularly if the definition of choice areas does not reflect the physical geography of the system and the way it affects fish distribution. To make this more concrete consider some geography and site choices made by a sample fleet in our data set (Fleet 6). In Figure 4a, we have a stretch of coastline where a quarter degree latitude grid (roughly 27 miles in north-south distance) might successfully "divide space" if combined with bathymetry data (given by the depth profiles as one goes deeper away from the coast). In this example, the small circles are actual site choices made by fishermen. For this fishery, defining discrete fishing areas with a regular grid based on latitude and longitude would make little sense unless done at a very fine scale, but a choice set structure that involves a latitudinal location and depth zone may be more appropriate. However, other stretches of coastline fished by this fleet look quite different (see Figure 4b). Notice that the depth contours (the geography that helps define choice for this fishery) do not run in any direction in a predictable way. Indeed, a vessel considering a move to the 300-fathom contour from the 500-fathom contour (like the one highlighted) can choose one of several areas within the same latitude band- each in opposite directions. As the underlying geography becomes less predictable and the importance of this geography for choice increases, the traditional site aggregation approach may have significant limitations.

The grid-point model with a fine enough grid may be better able to account for this fine scale heterogeneity of this fishing landscape. We estimate a grid-point model sampling from points spaced 3.5 miles apart. This is a shorter distance than the great majority of tows (i.e. distance from where trawl is dropped to where it is brought up). Arguably the grid-point model as applied here is a reasonable approximation of the true spatial definition of fishers' decisions since at least one grid point will typically fall within the area covered by a tow.

We model location choices for 71 trawlers participating in the groundfish limited entry ITQ trawl fishery during 2011 and 2012, the first two years of operation under ITQs. We group the vessels into 8

for both target and bycatch species in expected revenues and note that high expected bycatch of some species can drive expected revenues below zero since quota pound prices for those species can exceed the ex-vessel price.

fleets centered around ports in Washington, Oregon and Northern California (Table 6). Fishing locations as well as ancillary information such as species catch estimates, depth and gear are drawn from data collected by observers who are present on 100% of trips. Although we only model location choices in 2011 and 2012, we use 2010 logbook data to construct some of the models' explanatory variables such as expected revenue and whether the individual vessel had fished in that location the prior year. It is necessary to use logbook data for 2010 since there was only partial observer coverage in 2010 while logbook data is comprehensive.

We used data collected by the West Coast Groundfish Observer Program (WCGOP) run by the Northwest Fisheries Science Center as well as fishticket data provided by PacFIN to estimate moving average prices for species or species groups by port group. Port specific moving average prices (based on sales prior to the date of the choice being modeled) are applied to the estimated catches in the logbook data to derive estimates of expected revenue per tow. We also make use of data on quota pound transfers collected by the National Marine Fishery Service to estimate the value of quota pounds by ITQ stock. Quota pounds are the annual form of quota, thus quota pound prices are similar to lease prices. These quota pounds prices are used to estimate net revenue per pound for ITQ species. Net revenue per pound is the estimated ex-vessel value less the value of the quota pounds which might be either an opportunity cost or an out of pocket cost if the individual has to purchase the quota pounds to cover catch. As noted earlier, quota pound prices for some of the overfished rockfish species exceeded ex-vessel price during these years making net prices negative for these species.

Considering points as sites requires the researcher to characterize the various independent variables in Table 7 for every site that is sampled. Figure 1 demonstrates how this is done, but it requires some additional input from the researcher: the relevant spatial buffer and time dimension for supporting the calculation of these independent variables. We compared the performance of models with a variety of alternative distance-time buffers before settling on a distance-based criteria of 3 miles and a time criteria of the previous 30 days to determine which lagged observations were used to calculate expected revenue. Given the actual date of this tow, we find tows in the depth band, and within the 3 miles of the potential choice location, that occurred within the past 30-day period. We use all tows (the grey and black dots) not just the sampled tows (black dots) for making this calculation. Once we find the tows for the fleet satisfying those conditions, we average the observed net revenues using a simple mean[4]. In this example, considered to be 50-fathom alternative, we find the actual tows (stars) in the 50m fathom depth band within the radius and take a simple average of catch for our expected revenue value used in the analysis.

[4] We tried estimating models with more complex regression based methods of estimating expected revenue that included physical covariates, but found the simpler models worked better.

We do this for every sample location (black dot) for each observed tow. Defining shorter time and distance bands makes the expected revenue "surface" more spikey and adds variation. It is also important to note that the sampling method uses nearby information whether it is another latitude band or the same one but only locations in the same depth band and can therefore be thought of as approximating an expected revenue surface in continuous latitude-longitude space.

We can define the moving average revenues for any individual i, site j , and time t as

(14) $$\bar{R}_{ijt}(d^0, T^0) = \frac{(\sum_h \sum_k \sum_\tau \ (d_{jk} < d^0) * (t - T^0 < \tau < t) * R_{hk\tau})}{N(d^0, T^0)}$$

Where R_{hkt} are other tows taken by vessel h (could be equal to i), to point k, at time τ. $(d_{jk} < d^0)$ and $(t - T^0 < \tau < t)$ are logical operators equal to 1 for only the $N(d^0, T^0)$ other tows in the fleet meeting the distance and time requirements, respectively. Distance ($Distance$) from the location of the end of the last tow (or the port on the first tow) to the location of each potential choice is calculated using the standard great circle geometry formula. We calculate other site-specific information such as the habit dummy variables which take a value of one if the individual fished within 3 of miles of that location in the past 30 days ($Habit_{30}$) or within a 30-day period surrounding that date the prior year ($Habit_{Year}$).[5] The habit dummy variable are include as it has been shown that fishermen tend to fish locations that they have fished in the past (Holland and Sutinen 2000). There is also a dummy variable on missing activity, defined as $Dum_{Missing}$, when there is no information to calculate expected revenue

Our grid-point model used in the application is a modification of the grid-point model used in the MC. In the MC we divided all space up into 4,096 grids from which *any* of these points could be randomly sampled as a potential choice alternative in the grid-point model. In our application we are limited by the curse of dimensionality as the choice space is of extremely high dimension (nearly infinite) in both space and depth. To circumvent this, we generate a grid of points evenly spaced 3.5 miles apart from each other within the seven depth zones (0-50, 50-100, 100-150, 150-200, 200-250, 300-500, and 500-700 fathoms) and then randomly sampled 50 alternatives from this grid to construct the grid-point choice set.

For the traditional choice set model we divide the choice space using quarter-degree latitude bands combined with the seven depth zones (see Supplementary Appendix C for charts). Recall that the traditional choice set model is not able to fully exploit the spatial heterogeneity since vessel movements within a "zone" are treated as not moving and site-specific characteristics such as expected revenues and

[5] We use a threshold of 3 miles to limit the tows used in the calculation, but we have run models with 10 and 15-mile thresholds as well to investigate the robustness of our estimates and find our results quite robust to this specification (see Table B1 in Appendix B).

bycatch are the same within the zone. The choice set for each fleet is limited to areas that were fished by that fleet between 2010-2012. Expected revenue for the traditional model is calculated by averaging revenue from all tows that occurred within the specific zone the prior 30 days (using the reported catches and lagged moving average port prices to calculate revenue for each tow). Habit dummy variables reflect whether that zone was fished by that individual the prior 30 days or the 30 days surrounding the date the prior year. The dummy variable on missing activity takes a value of one when there were no tows in that zone the prior 30 days. Distance from the location of the end of the last tow (or the port on the first tow) to the center of each choice zone is calculated using standard great circle geometry formula.

5.1: Application Results

Like previous location choice studies in other fisheries, our models suggest that fishermen generally prefer to fish a closer site than a further one both when choosing the location for the first set of the trip and for subsequent sets within a trip. Models for all fleets, regardless of the choice set specification, have negative and significant signs on $Distance_{First}$ (distance from port to locations choices for the first tow of the trip) and on $Distance$ (distance from the endpoint of a tow to the start of the next tow) (Table 8). The magnitude of the coefficients is similar across choice set specifications. For all but one fleet, and for both model specifications, the coefficient on $Distance_{First}$ is smaller than the coefficient $Distance$ indicating that, while closer locations are preferred, vessels are less averse to going a long distance when steaming out from port and choosing a location for the initial tow. In essence, they appear to be treating the costs incurred to travel to the fishing grounds as a fixed cost (or a cost spread across the subsequent tows on the trip) and the distance traveled between tows to be the relevant marginal cost in their spatial choices. The only exception is for fleet 4 for which the coefficient $Distance_{First}$ is larger than $Distance$. Fleet 4 has substantially longer average tow lengths (measured in hours) than all other fleets and it also has the smallest coefficient on $Distance$. The longer tow length would tend to reduce the ratio of tow time to steam time, which may be why this fleet appears to be less averse to travelling further between tows.

For nearly all fleets and model specifications the coefficient on expected revenue on the first tow of the trip, $Revenue_{First}$, is positive and significant (Table 8). This suggests that vessels are attracted to areas with higher expected revenues, as would be expected. The only exceptions are for fleet 4 for which expected revenue on the first tow is not significant for the traditional model. Results are less consistent with the coefficient on expected revenue for subsequent tows, $Revenue$. For the grid-point model the $Revenue$ coefficient is significant and positive for all fleets other than fleet 6. For the traditional model $Revenue$ is significant and positive only for fleets 7 and 8 and is significant and negative for fleet 5. The lack of significant *Revenue* coefficients for the traditional model for several fleets may be due to the coarser spatial resolution in the choice set which may be aggregating a number of fishing areas of differing fishing quality. It is notable that the grid-point model works best with only a 3-mile radius to estimate expected revenue (Appendix B). This suggests fishers respond to heterogeneity at a very fine scale which is consistent with anecdotal accounts from fishers and what we know about the patchiness of the resource.

For the most part the coefficients on revenue after the first tow, $Revenue$, have a higher magnitude for the grid-point model relative to the traditional model. The coefficients for $Revenue$ in the traditional model are very small and mostly not significant. Notably, since distance coefficients are fairly similar across models, the ratio of revenue to distance coefficients tends to be smaller in magnitude for the traditional model. The ratio for coefficients on expected revenue and distance is sometimes used in the welfare analysis of management actions such as area closures, e.g., to determine the expected value of compensation needed to keep fishers' utility constant after the removal of the closed areas as a fishing choice (e.g., Curtis and Hicks 2000, Hicks et al. 2004). Negative bias in the revenue coefficient, and smaller absolute ratio of revenue and distance coefficients, could inflate welfare estimates of the cost of closures or lead to negative estimates for negative draws of the revenue coefficient if its confidence interval bounds zero. This is indeed the case in our application (see Section 5.3 below). The smaller or insignificant coefficients for expected revenue for the traditional model are consistent with the results of the MC analysis.

Like Holland and Sutinen (2000) we find that fishermen tend to visit sites that they have a prior history of visiting. The $Habit_{30}$ variables, which take a value of 1 if the vessels fished that location within the last 30 days, are positive and significant for all fleets for the grid-point model (Table 8). This variable is not significant for fleet 2 for the traditional model but is for all other fleets and models. This site fidelity also appears to be long lasting since the coefficient on fishing within 3 miles of that location in a 30-day time window around that date the prior year, $Habit_{Year}$, is also predominately positive and statistically significant.

The coefficient $Dum_{Missing}$ which indicates there has been no fishing by any vessels at that location in the prior 30 days (and thus no data to calculate the revenue variable for the site) is predominately negative and significant for the grid-point and traditional model indicating that fishermen tend to avoid these areas (Table 8). As noted above, the grid-point model and the traditional model include locations that may never have been fished or may be unfishable, and this may partly explain the negative value of $Dum_{Missing}$ for these models as it may be serving as a proxy both for areas that were not fished recently and those never fished. These areas may also have been avoided due to concerns about bycatch of rockfish.

5.2: Prediction Performance Metrics Results

We calculate the same performance metrics utilized in the MC to evaluate the predictive performance of each model in our application. For these predictions, the traditional model is predicting the choice and choice probabilities amongst all of the traditional areas. For the grid model it is predicting the choice and choice probabilities for all grid points with the range of the fleet – not just the grid points sampled in the choices. The number of grid points in the full choice set range from a low of 209 for Fleet 2 to a high of 1030 for Fleet 7. These tests show clearly that, in almost all cases, the grid-point substantially outperforms the traditional models in terms of predicting locations choices whether considering the proportion of choices correctly predicted, the correct prediction probabilities summed, or the probability mass assigned to the actual choice (Tables 9 and 10). For the first tow, the traditional model has higher prediction scores for the CP and CPS metrics for Fleet 2 and for the CP and PM metrics for Fleet 6, but grid model prediction is always superior when modeling predictions for tows other than the first tow. The grid-point model also has a smaller average distance between the predicted and actual choice in all cases except when predicting the first tow for fleets 6 and 8. This is likely due to the fact that distance plays a much strong role in subsequent tows which tends to limit the number of acceptable choices for the next tow. Both grid-point and traditional models do a poorer job predicting the location of the first tow than they do for subsequent tows. This is probably because distance does not play as strong a role in determining location choice for the first tow.

5.3: Policy Experiment and Welfare Estimates

Another policy metric we use for comparing the grid model to traditional methods is a welfare measure for area closures. Following other applications for welfare effects of area closures in commercial fisheries (e.g., Hicks and Schnier 2008; Hicks et al. 2004) we measure commercial fishers' willingness to pay (WTP) to avoid an area closure. Specifically we evaluate welfare effects of a closure of the 50-100 fathom isobaths and of the 200-250 isobath. These are hypothetical closures, but Pacific Fishery management council recently passed an essential fish habitat amendment to the groundfish fishery management play that would enable temporary closures of this type on an emergency basis in response to high levels of salmon bycatch. Denoting fisher *i*'s utility function at site *k* as[6]

$$(15)\ V(X_{ikt},\beta) = \beta_d Distance_{ikt} + \beta_r Revenue_{ikt} + \beta_{miss} D_{ikt} + \beta_h Habit_{ikt,30} + \beta_{hy} Habit_{ikt,y} + \epsilon_{ikt}$$

$$= v(X_{ikt},\beta) + \epsilon_{ikt}$$

Further, denoting K as the full set of choice alternatives and K^C as the subset of alternatives available to

[6] For the sake of brevity, we don't distinguish between first and subsequent tows here.

the fisher following an area closure policy, we can write the fisher *i*'s willingness to pay to avoid the closure for tow *t* as

$$(16) \quad WTP_{it} = \frac{\ln[\sum_{k \in K} v(X_{ikt}, \beta)] - ln[\sum_{k \in K^c} v(X_{ikt}, \beta)]}{\beta_r}$$

following the well-known compensating variation expression first derived in Hanemann (1982) assuming linear income effects. We present sample means of these measures across all individuals and tows.

Some important caveats should be mentioned with respect to this measure. Utility units are converted to willingness to pay by way of our model's coefficient on expected revenue. An ideal modeling approach would be to include site-specific profits at each alternative that would net costs such as travel costs and operating expenses from revenues. Due to data limitations we are unable convert travel distance to costs in defensible way since costs per distance unit travel isn't known. Therefore we enter distance in the utility function directly as a separate argument. Other costs that vary by site are not included in our specification but we believe that travel distance to site is the primary site-specific variable cost of concern. Additionally, we assume that site covariates won't change as a result of the closures. Despite these limitations, we believe that the welfare measure provides a useful additional performance metric for comparing to traditional methods.

The welfare estimates for closures (WTP to avoid the closure) when applied to tows other than the first tow (Tables 11) vary substantially across fleets but even more so across choice set specifications (grid vs. traditional). The estimates from the grid model tend to have narrower bounds and be positive - as they should be theoretically. Estimates for the traditional model often have very wide bounds with negative mean results in some cases and bounds that include negative values in others. The reason for this is that coefficients for expected revenue from the traditional model are often much closer to zero and are often not significant having confidence intervals that span zero. Draws from the covariance matrix can result in the denominator in equation 16 being very small (leading to very large magnitudes) or negative (leading to negative estimates of welfare effects).

The welfare estimates for first tows (Table 12) are much more similar across models, and estimates are positive and have smaller bounds for both models. This is because revenue coefficients for the first tow are generally significant and positive for both models, though revenue coefficients are smaller for the traditional model than the grid model. Since the revenue coefficient is in the denominator of equation 16, we might expect mean welfare estimates for the traditional model to be inflated relative to the grid model. However this is not necessarily the case. While the smaller revenue coefficient does tend to increase the welfare estimate all else equal, the numerator in equation 16 is also impacted by distance and habit variables which can make the relative effect of the deep or shallow closures vary for different fleets and models. This accounts for differences in relative welfare impacts of shallow and deep closures for the two models for some fleets.

5.4 Robustness checks for Effects of Heterogeneity

The RUM models presented in the prior sections included habit variables in the utility function which, since they are positive and significant, show that, all else equal, individuals tend to prefer fishing locations they have fished in the past. While these variables are intended to capture state dependence, they can confound state dependence and heterogeneity (Heckman 1981, Smith 2005). To the extent that the habit variables capture heterogeneity this would suggest that, for two areas with the same expected revenue based as estimated in our model from fleet activity, an individual's expectation of profit will be higher for an area they have fished. This might reflect personal knowledge of how to fish that area. It is possible that this could cause bias in other model parameters, particularly the expected revenue parameter, and this could in turn bias welfare estimates (e.g. estimating the cost of a closure) which are strongly influenced by the expected revenue coefficients. To explore the robustness of our base model we run models without habit variables and also models with a mixed logit formulation allowing coefficients to vary randomly with a normal distribution. The results from these models are presented Supplementary Appendix D. We summarize the findings below.

For models run without habit variables, coefficients on distance and expected revenue generally exhibit only small changes in magnitude and levels of statistical significance are unaffected in all but a few cases. Likelihood ratio tests reject eliminating the habit variables, though, as we note below, they do support allowing for heterogeneity in the utility function for some parameters. Predictive accuracy generally declines with elimination of habit variables as expected but the degree of decline varies by fleet. Decline in prediction power without habit variables is much greater when predicting location of the first tow than for subsequent tows suggesting that considering past choices is less important for predicting behavior within the trip.

Welfare measures also change with elimination of habit variables. Changes are relatively small for grid model welfare estimates other than for models where expected revenue parameters are not close to zero and not significant. Welfare estimates for traditional models exhibit much larger changes which is consistent with the much higher volatility and wider bounds of those estimates due to the revenue coefficients being close to zero and often insignificant. For both grid and traditional welfare estimates changes are much smaller for first tows. This is due to the fact that expected revenue parameters tend to be larger and more significant and more robust to model specification.

For the mixed logit models we again find that mean coefficients for distance and expected revenue generally exhibit only small changes in magnitude and levels of statistical significance are mostly unaffected. In many, but not all, cases the expected revenue coefficients from the mixed logit model are smaller relative to the base model but this effect is not consistent across fleets. For most fleets the standard deviation of the coefficients on expected revenue and distance for the first tows is not statistically significant suggesting there in not significant heterogeneity in these coefficients. However, the coefficients for tows other than the first tow often do exhibit statistically significant variation suggesting heterogeneity across vessels. Likelihood ratio tests support the mixed logit model over the base model in terms of overall fit. However, prediction scores for the mixed logit models are slightly lower than for the base model in almost all cases. It is not clear why this would be the case but the result is consistent. Welfare estimates for our policy experiments modeling the cost of closures tend to be more erratic with wider bounds as compared to the base model. Welfare estimates for tows other than the first tow, which were almost always positive for the base conditional logit grid model, are sometimes negative for the mixed logit model and have wider bounds. This is consistent with the smaller coefficient (closer to zero) which can cause welfare estimates to be negative when draws of the revenue coefficient are negative in some policy simulations. Given that our main focus is on comparing the grid and traditional choice formulation opt to use our simpler conditional logit model with habit variables in the main paper and present the mixed logit results in an appendix. In particular, with the prediction scores and policy experiments and welfare estimates this base model provides a more clear illustration of how negative bias in the revenue coefficients can impact ability to predict choice and welfare impacts. We note that Smith (2005) cautions more against modeling preference heterogeneity in isolation (e.g. with random parameters) than against modeling state dependence in isolation.

6. Conclusions

Spatial discrete choice modeling is challenging because it requires the proper specification of the subjects' consideration set to obtain unbiased parameter estimates. Given the breadth of its use within the environmental economics literature, developing estimation methods that can more accurately capture the underlying choice structure will advance the application of these methods. Furthermore, with fine scale geo-spatial data becoming more readily available it is imperative that we develop estimation tools that are capable of meeting this growth in data availability when individuals are likely to be aware of alternatives at very fine spatial resolutions. This research expands the discrete choice literature by developing a grid-point choice structure that is not fettered by the traditional and arbitrary aggregation of space into large areas used extensively in the literature to facilitate estimation.

To investigate the properties of our grid-point choice model we compare it to the traditional choice model that arbitrarily divides space into pre-specified grids using MC analysis. The results from the MC suggests that the grid-point model may perform the best if the objective is accurate estimates of coefficients. This might be the case when they are to be used to estimates the costs of a closure for example. The MC showed, however, that a variety of factors can influence the absolute and relative performance of these models. Some of these may be within the control of the analyst, e.g. the size of discrete areas for a traditional model. The study suggests the advantage of the grid-point model will diminish as the size of discrete areas for the traditional model is reduced (i.e. defined on a finer scale). For other factors that affect the relative performance of the models such as the sizes and number desirable choices and how quickly utility changes over space (determined in our MC by the clumps and diffusion rate parameters), the analyst may have some information about them and these might help inform model choice. There may be circumstances where the traditional choice set model outperforms the grid-point model. Those conditions are what we have defined as relatively spatially homogeneous choice structures, in which the arbitrary gridding of space does not impact the analysis.

The MC analysis provided less clear results with regards to predictive accuracy which in some cases is the primary focus of an application (e.g. if it is to be used to predict how changes in regulations affect effort distribution). However, unlike coefficient bias, where true values are unknown, we are able to evaluate predictive accuracy empirically over actual data. Our empirical application suggested a clear advantage for the grid-point models over the traditional model in terms of predictive accuracy.

Our application focused on a fishery where both the physical geography and the resource is known at a very fine scale. The econometric method effectively builds a localized surface around sampled points for characterizing fishing quality. In constructing this surface, we find that a radius of 3 miles proved to be the best alternative for including past tows for calculating expected revenue. A short radius around sample points such as this greatly reduces the number of observations that can be used to calculate expected revenue for a given location choice, but nevertheless appears to provide a more accurate signal when the information is available which is consistent with the patchy and uneven spatial distribution of areas driving catch in the fishery. Even though the traditional model implements areas shaped according to bathymetry which is known to drive resource distributions, it seems clear that the artificial aggregation of space based on geographic boundaries (in our case latitude cutoffs and depth zones) masked underlying heterogeneity that drives fishers' decisions. Many fisheries and other choice settings likely also exhibit fine scale heterogeneity and our point-based considerations sets may provide more effective strategies for modeling choice in these situations.

Acknowledgements

Funding for this research was provided by NOAA, National Marine Fisheries Services. Funding sources were not involved in the study design, analysis, or writing of this paper or the decision to submit the article for publication.

References

Andrews, Rick L. and T.C. Srinivasan. 1995. Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data. *Journal of Marketing Research* 32(1): 30-41.

Banzhaf, H.Spencer and V. Kerry Smith. 2007. Meta-Analysis in Model Implementation: Choice Sets and the Valuation of Air Quality Improvements. *Journal of Applied Econometrics* 22(6): 1013-1031.

Ben-Akiva, Moshe and Steven Lerman. 1985. *Discrete Choice Analysis*. Cambridge, Mass. MIT Press.

Bronnenberg, Bart J. and Wilfried .R. Vanhonacker. Limited Choice Sets, Local Price Response and Implied Measures of Price Competition. *Journal of Marketing Research* 33(2): 163-173.

Curtis, Rita and Robert L. Hicks. 2000. "The Cost of Sea Turtle Preservation: The Case of Hawaii's Pelagic Longliners," *American Journal of Agricultural Economics* 82(5): 1191-97.

Desarbo, Wayne S. and Kamel Jedidi. 1995. The Spatial Representation of Heterogeneous Consideration Sets. *Marketing Science* 14(3): 326-342.

Feather, Peter M. 1994. Sampling and Aggregation Issues in Random Utility Model Estimation. *American Journal of Agricultural Economics* 76(4): 772-780.

Dupont, Diane P. 1993 "Price Uncertainty, Expectations Formation, and Fishers' Location Choices," *Marine Resource Economics* 8(3): 219-247.

Eales, James S. & Wilen, James E., 1986. An Examination of Fishing Location Choice in the Pink Shrimp Fishery, *Marine Resource Economics* 2(4) 331-351.

Feather, Peter M. Sampling and aggregation issues in random utility model estimation. *American Journal of Agricultural Economics* 76.4 (1994): 772-780.

Haab, Timothy C., and Robert L. Hicks. Accounting for choice set endogeneity in random utility models of recreation demand. *Journal of Environmental Economics and Management* 34.2 (1997):127-147.

Hanemann, Michael W. 1982. Applied welfare analysis with qualitative response models. Working Paper No. 241, *CUDARE Working Papers*, University of California, Berkeley, 1982.

Heckman, James J. 1981. Statistical models for discrete panel data in *Structural analysis of discrete data with econometric applications*, ed. C. F. Manski and D. McFadden. Cambridge: pp. 114-178. *MIT Press*

Hicks, Robert L. and Ivar E. Strand. 2000. The Extent of Information: Its Relevance for Random Utility Models. *Land Economics* 76(3): 374-385.

Hicks, Robert L., James Kirkely, and Ivar Strand. 2004. Short-Run Welfare Losses from Essential Fish Habitat Designations for the Surfclam and Ocean Quahog Fisheries, *Marine Resource Economics* 19(1): 113-129.

Hicks, Robert L. and Kurt E. Schnier. 2008. "Eco-labeling and dolphin avoidance: A dynamic model of tuna fishing in the Eastern Tropical Pacific," *Journal of Environmental Economics and Management* 56(2):103-115.

Hicks, Robert L. and Kurt E. Schnier. 2010. Spatial Regulations and Endogenous Consideration Sets in Fisheries, *Resource and Energy Economics*, vol. 32(2): 117-34

Holland, Daniel S. and Jason E. Jannot. 2012. Bycatch risk pools for the US West Coast Groundfish Fishery. *Ecological Economics*, 78: 132-147.

Holland, Daniel S. and Jon G. Sutinen. 2000. Location Choice in New England Trawl Fisheries: Old Habits Die Hard,*Land Economics* Vol. 76(1):133-149 1.

Holland, Daniel S. 2016. Development of the Pacific Groundfish Trawl IFQ Market. *Marine Resource Economics.* 31(4):453-464.

Manski, Charles 1977. The Structure of Random Utility Models. *Theory and Decision* 8: 229-254.

McFadden, Daniel. 1978. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, eds. A. Karlqvist et al., 75-96. Amsterdam: North-Holland Publishing.

Parsons, George R. and Mary Jo Kealy. 1992. Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation, *Land Economics* 68.1 (1992): 93-106.

Parsons, George R., and Michael S. Needelman. 1992. Site Aggregation in a Random Utility Model of Recreation. *Land Economics* 68(4): 418-433.

Parsons, George R., and A. Brett Hauber. 1998. Spatial Boundaries and Choice Set Definition in a Random Utility Model if Recreation Demand. *Land Economics* 74(1): 32-48.

Parsons, George R., D. Matthew Massey, and Ted Tomasi. 1999. Familiar and Favorite Sites in a Random Utility Model of Beach Recreation. *Marine Resource Economics* 14: 299-315.

Parsons, George R., Andrew J. Plantinga, and Kevin J. Boyle. 2000. Narrow Choice Sets in a Random Utility Model of Recreation Demand. *Land Economics* 76(1): 86-99.

Peters, Thomas, Wictor L. Adamowicz, and Peter C. Boxall. 1995. The Influence of Choice Set Consideration in Modeling the Benefits of Improved Water Quality. *Water Resources Research* 613: 1781-1787.

Roberts, John H., and James M. Lattin. 1991. Development and Testing of a Model of Consideration Set Composition. *Journal of Marketing Research* 28(4): 429-440.

Roberts, John, and Prakash Nedungadi. 1995. Studying Consideration in the Consumer Decision Process: Progress and Challenges. *International Journal of Research in Marketing* 12: 3-7.

Scrogin, David, Richard Hofler, Kevin Boyle, and J. Walter Milon. 2010. An Efficiency Approach to Choice Set Formation: Theory and Application to Recreational Destination Choice. *Applied Economics* 42: 333-350.

Shocker, Allan D., Moshe Ben-Akiva, Bruno Boccara, and Prakash Nedungadi.1991. Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions. *Marketing Letters* 2(3): 181-197.

Smith, Martin D., 2002. Two Econometric Approaches for Predicting the Spatial Behavior of Renewable Resource Harvesters, *Land Economics* 78(4), 522-538.

Smith, Martin D. 2005. State dependence and heterogeneity in fishing location choice. *Journal of Environmental Economics and Management* 50 (2): 319-40.

Smith, Martin D. and James E. Wilen. 2003. Economic Impacts of Marine Reserves: The Importance of Spatial Behavior, *Journal of Environmental Economics and Management* 46(2): 183-206.

Swait, Joffre, and Moshe Ben-Akiva.1987a. Incorporating Random Constraints in Discrete Choice Models of Choice Set Generation. *Transportation Research Part B* 21: 91-102.

Swait, Joffre, and Moshe Ben-Akiva. 1987b. Empirical Test of a Constrained Discrete Choice Model: Model Choice in Sao-Paulo. *Transportation Research Part B* 21: 103-115.

Swait, Joffre 2001a. A Non-Compensatory Choice Model Incorporating Attribute Cutoffs. *Transportation Research Part B* 35: 903-928.

Swait, Joffre 2001b. Choice Set Generation Within the Generalized Extreme Value Family of Discrete Choice Models. *Transportation Research Part B* 35: 643-666.

von Haefen, Roger H. and Adam Domanski, A., 2018. Estimation and welfare analysis from mixed logit models with large choice sets. *Journal of Environmental Economics and Management*. 90:101-118.

Whitehead, John C., and Timothy C. Haab.1999. Southeast Marine Recreational Fishery Statistical Survey: Distance and Catch Based Choice Sets. *Marine Resource Economics* 14: 283-298.

Yoon, Seo Youn, Kathleen Deutsch, Yali Chen, and Konstadinos G. Goulias. 2012. Feasibility of Using Time-Space Prism to Represent Available Opportunities and Choice Sets for Destination Choice Models in the Context of Dynamic Urban Environments. *Transportation* 39: 807-823.

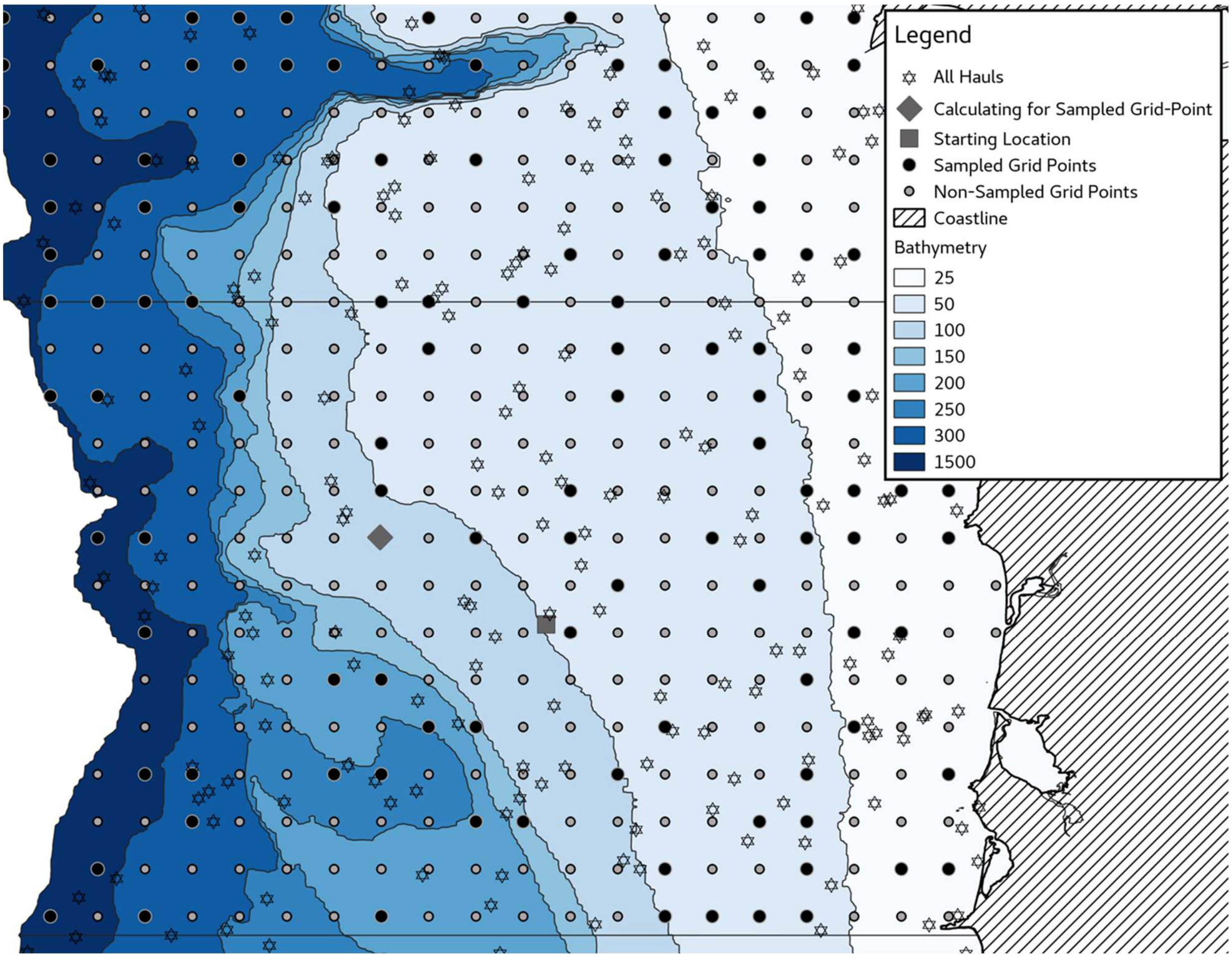

Figure 1: Illustration of grid point-based choice set formation.

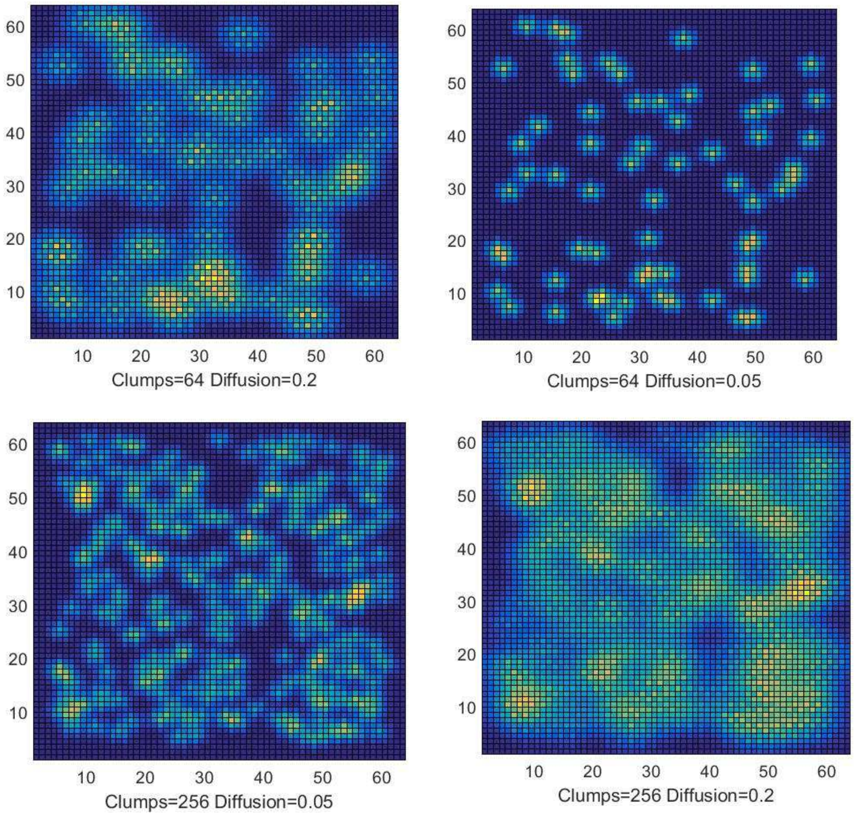

Figure 2: Graphical illustration of the resource spatial distribution, Z^1_{ijt}, for alternative clumps and spatial diffusion rates, δ, of the resource.

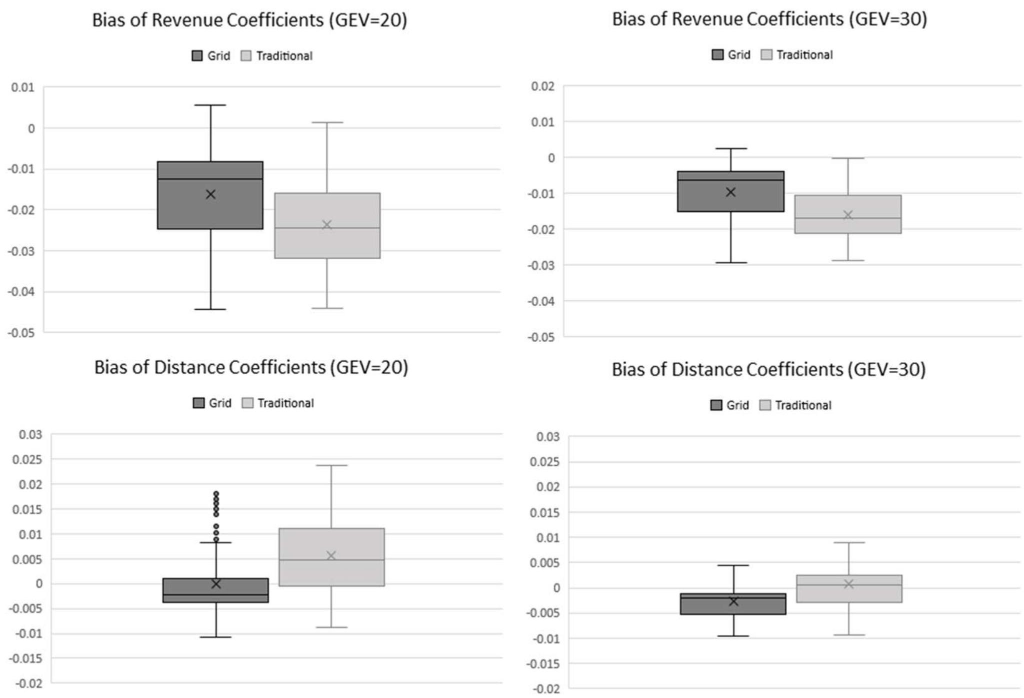

Figure 3: Box plots of bias of in revenue and distance coefficients from Monte Carlo analysis.

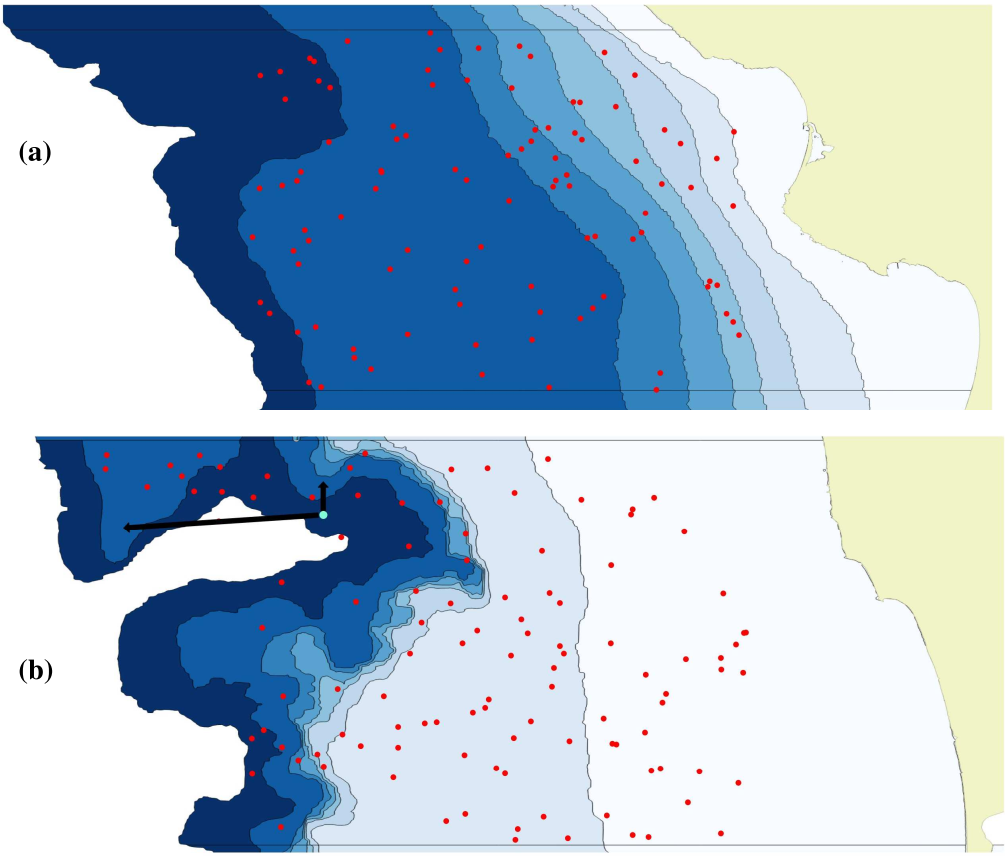

Figure 4: (a) Sample of a simple geography; (b) Sample of complicated geography.

Table 1: Parameters utilized in our data generation process for the Monte Carlo. A full combinatorial of data parameters was constructed resulting on 864 unique data set parameters.

Model Variable	Settings
Number of Clumps (seeded locations on fish surface)	64, 128, 256
Diffusion Rate, δ	0.05, 0.075, 0.10, 0.20
Standard Deviation of Error Added to Expected revenue, σ^E	10, 20
Standard Deviation of Error Added to Actual Catch, σ^A	10, 20, 40
Distance Scale (scales relative distance between grid points)	1, 3
GEV Scale Factor, ζ	20, 30
Aggregation Factor (number of cells in discrete areas for traditional model	16, 64, 256

Table 2: Average parameter estimates and mean bias across the Monte Carlo parameterizations for the revenue coefficient (left column), β_1, and the distance coefficient (right column), β_2. The first row of each section contains the results for GEV scale parameter of 20 ($\zeta = 20$) and the second row contains the results for the GEV scale parameter of 30 ($\zeta = 30$). Bias is defined as $(\hat{\beta} - \beta)$ where $\hat{\beta}$ is the estimated coefficient and β is the true parameter estimate. The true parameter estimate, β, is the coefficient within the data generation process divided by the GEV scale parameter, ζ. Standard deviations of parameters averages and bias are in parentheses. Note that standard deviations reflect variation if scores across MC parameterizations.

	Grid		Traditional		Grid		Traditional	
GEV Scale	Average Coefficient Estimate for β_1				Average Coefficient Estimate for β_2			
$\zeta = 20$	0.034	(0.013)	0.026	(0.011)	-0.050	(0.006)	-0.044	(0.007)
$\zeta = 30$	0.023	(0.008)	0.017	(0.007)	-0.036	(0.003)	-0.032	(0.004)
GEV Scale	Average Bias B_1				Average Bias B_2			
$\zeta = 20$	-0.016	(0.013)	-0.024	(0.011)	-0.000	(0.005)	0.006	(0.007)
$\zeta = 30$	-0.011	(0.008)	-0.016	(0.007)	-0.003	(0.003)	0.001	(0.004)
	Average Percent Bias for β_1				Average Percent Bias for β_2			
$\zeta = 20$	-33%		-47%		0%		11%	
$\zeta = 30$	-29%		-48%		-8%		2%	

Table 3: Differences in the absolute value of bias between the Grid and Traditional models for the revenue coefficient β_1 (left two columns) and for the distance coefficient β_2 (right two columns). The dependent variable is mean difference in the absolute value of bias between the two models (Grid bias – Traditional bias). For both coefficients results are shown for the GEV scale parameter of 20 ($\zeta = 20$) and the GEV scale parameter of 30 ($\zeta = 30$). A negative regression coefficient indicates that this MC variable increases absolute bias relatively more for the traditional model than the grid model (i.e., favoring the grid model) and a positive coefficient indicates the opposite.

Monte Carlo Variable	Grid-Traditional$^{\beta_1}$				Grid-Traditional$^{\beta_2}$			
	$\zeta = 20$		$\zeta = 30$		$\zeta = 20$		$\zeta = 30$	
Constant	-0.007	***	-0.005	***	0.003	***	0.002	***
Clump 128	-0.006	***	-0.005	***	0.000		0.000	
Clump 256	-0.002	***	-0.001	**	-0.001	*	0.000	
Diffusion 0.075	0.000		0.000		0.001	*	0.001	*
Diffusion 0.10	0.002	***	0.000		0.000		0.001	***
Diffusion 0.20	0.007	***	0.004	***	0.000		0.002	***
Distance Scale 3	-0.001		0.000		-0.004	***	-0.003	***
Std. Dev. Exp. Rev. 20	0.000		0.000		0.000		0.000	
Std. Dev. Act. Rev. 20	0.003	***	0.002	***	0.000		0.000	
Std. Dev. Act. Rev. 40	0.009	***	0.006	***	0.001	**	0.000	
Agg. Factor 64	-0.003	***	-0.004	***	-0.002	***	0.000	
Agg. Factor 256	-0.005	***	-0.005	***	-0.008	***	-0.002	***
R-square	0.65		0.48		0.58		0.55	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Table 4: Average prediction and distance performance scores for different models average across all Monte Carlo parameterizations. Standard deviations of scores are in parentheses. Note that standard deviations reflect variation if scores across MC parameterizations.

Model	CP		CPS		PM		D	
Grid-point	0.24	(0.17)	0.26	(0.18)	0.22	(0.16)	24.6	(10.2)
Traditional	0.25	(0.18)	0.25	(0.18)	0.17	(0.15)	27.1	(10.1)

Table 5 Determinants of the differences in the prediction metrics for the two models broken down by the data generation parameters. The dependent variable is mean difference prediction metric between models (Grid score – Traditional score). A positive coefficient indicates that this MC variable increases the score relatively more for the grid model than the traditional model. Since high scores are better for CP, CPS, and PM, a positive coefficient means that variable tend to favor the grid model. For the distance score the opposite is true.

Monte Carlo Variable	CP Scores		CPS Scores		PM Scores		Distance Scores	
Constant	-0.042	***	0.015	***	0.015	***	-1.372	***
Clump 128	0.003		-0.004	***	0.002		1.621	***
Clump 256	0.015	***	-0.008	***	0.000		2.290	***
Diffusion 0.075	0.013	***	0.001		0.016	***	-0.530	**
Diffusion 0.10	0.016	***	0.003		0.013	***	-0.372	
Diffusion 0.20	0.018	***	-0.002		0.012	***	0.296	
GEV Scale 30	0.022	***	-0.007	***	0.008	***	1.457	***
Distance Scale 3	0.011	***	0.005	***	0.017	***	-2.196	***
Std. Dev. Exp. Rev. 20	0.002		0.001		0.003	*	0.018	
Std. Dev. Act. Rev. 20	-0.003		0.000		-0.002		0.124	
Std. Dev. Act. Rev. 40	-0.011	***	-0.001		-0.003		0.428	**
Agg. Factor 64	0.006	***	0.004	***	0.010	***	-1.888	***
Agg. Factor 256	0.016	***	0.009	***	0.028	***	-3.530	***
R-square	0.24		0.16		0.19		0.51	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Table 6. Fleets and Fleet Groups Modeled

Fleet	Vessel Count	State	Major Ports
1	7	CA	Moss Landing & San Francisco
2	6	CA	Fort Bragg
3	9	CA	Eureka
4	6	OR	Crescent City & Brookings
5	12	OR	Charleston
6	8	OR	Newport
7	18	OR	Astoria
8	5	WA	Ilwaco & Westport

Table 7. Independent Variables for RUM Model

Variable Name	Description
$Distance$	Distance (in miles) to Tow Choice Location
$Distance_{First}$	Distance (in miles) to Tow Choice Location for 1st Tow of Trip
$Revenue$	Expected Revenues (in \$100)
$Revenue_{First}$	Expected Revenues (in \$100) for 1st Tow of Trip
$Dum_{Missing}$	(=1) if no observations in support of Expected Revenue calculation
$Habit_{30}$	(=1) if vessel has previously fished within 5 miles of site within 30 days
$Habit_{Year}$	(=1) if vessel has previously fished within 5 miles of site in 30 day of preceding year

Table 8: Parameter estimates for the eight fleets estimated within the Pacific groundfish fishery. The top panel contains the results from the grid-point model. The bottom panel contains the results from the traditional choice set model.

	Grid-point Model															
Parameter	Fleet 1		Fleet 2		Fleet 3		Fleet 4		Fleet 5		Fleet 6		Fleet 7		Fleet 8	
Distance	-0.119	***	-0.084	***	-0.077	***	-0.056	***	-0.086	***	-0.105	***	-0.112	***	-0.098	***
DistanceFirst	-0.032	***	-0.023	***	-0.053	***	-0.063	***	-0.051	***	-0.027	***	-0.017	***	-0.024	***
Revenue	0.027	***	0.004	*	0.004	*	0.012	***	0.009	***	0.009	**	0.017	***	0.008	**
Revenue_First	0.054	***	0.017	***	0.012	***	0.012	**	0.033	***	0.025	***	0.037	***	0.028	***
Dum_Missing	0.168		-0.756	***	-0.789	***	-0.702	***	-0.478	***	-0.984	***	-0.934	***	-0.232	
Habit_30	2.717	***	1.035	***	1.404	***	1.110	***	1.633	***	1.033	***	1.546	***	1.427	***
Habit_Year	1.496	***	0.774	***	1.000	***	0.745	***	1.149	***	0.996	***	0.842	***	1.082	***
	Traditional Model															
Parameter	Fleet 1		Fleet 2		Fleet 3		Fleet 4		Fleet 5		Fleet 6		Fleet 7		Fleet 8	
Distance	-0.133	***	-0.076	***	-0.074	***	-0.050	***	-0.092	***	-0.096	***	-0.113	***	-0.088	***
DistanceFirst	-0.043	***	-0.030	***	-0.052	***	-0.056	***	-0.067	***	-0.029	***	-0.024	***	-0.016	***
Revenue	0.006		0.001		0.000		0.001		-0.004	*	0.002		0.003	**	0.007	***
Revenue_First	0.014	*	0.011	***	0.014	***	0.007		0.025	***	0.020	***	0.017	***	0.022	***
Dum_Missing	-1.314	***	-1.032	***	-1.410	***	-1.034	***	-1.560	***	-1.035	***	-1.572	***	-0.982	***
Habit_30	0.551	***	0.027		0.319	***	0.266	**	0.193	**	0.350	**	0.526	***	0.262	**
Habit_Year	1.280	***	0.843	***	0.459	***	1.109	***	0.594	***	0.581	***	0.521	***	0.649	***
Observations	1451		1185		1496		895		2148		921		7084		1543	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Table 9: Predictive accuracy measures for the Pacific groundfish application for tows other than the first tow of a trip

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
				Correct Prediction (CP)				
Grid-point	0.58	0.23	0.22	0.14	0.27	0.23	0.34	0.30
Traditional	0.21	0.10	0.12	0.15	0.10	0.17	0.06	0.20
				Correct Prediction Summed (CPS)				
Grid-point	0.60	0.25	0.26	0.19	0.31	0.24	0.37	0.25
Traditional	0.21	0.10	0.12	0.15	0.10	0.17	0.06	0.20
				Probability Mass (PM)				
Grid-point	0.46	0.12	0.15	0.11	0.18	0.15	0.24	0.17
Traditional	0.15	0.04	0.06	0.07	0.06	0.09	0.04	0.10
				Distance (D)				
Grid-point	9.9	15.0	14.2	18.2	13.2	14.5	13.6	12.4
Traditional	29.9	40.1	38.7	25.8	27.6	28.0	56.4	33.6

Table 10: Predictive accuracy measures for the Pacific groundfish application on the first tow in the trip

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
Correct Prediction (CP)								
Grid-point	0.43	0.11	0.23	0.21	0.27	0.14	0.13	0.22
Traditional	0.21	0.14	0.13	0.17	0.12	0.26	0.07	0.21
Correct Prediction Summed (CPS)								
Grid-point	0.50	0.11	0.25	0.25	0.34	0.16	0.17	0.19
Traditional	0.21	0.14	0.13	0.17	0.12	0.26	0.07	0.21
Probability Mass (PM)								
Grid-point	0.29	0.06	0.12	0.11	0.13	0.07	0.07	0.08
Traditional	0.15	0.05	0.07	0.09	0.07	0.11	0.05	0.10
Distance (D)								
Grid-point	17.5	28.1	15.9	17.7	15.0	32.3	36.5	32.6
Traditional	27.9	31.5	36.9	21.9	22.2	20.7	51.6	27.2

Table 11: Welfare estimates for closures for the Pacific groundfish for tows other than the first tow

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
			50-100 Fathom Closure - Grid Model					
Lower Bound	32.3	41.2	-22.0	2.4	48.3	11.9	28.7	24.7
Mean	33.9	65.5	43.4	2.5	52.9	21.9	29.2	43.6
Upper Bound	35.4	89.8	108.9	2.7	57.5	31.8	29.7	62.5
			50-100 Fathom Closure - Traditional Model					
Lower Bound	71.8	-479.0	-59.1	-4.9	-219.5	-3371.4	31.3	17.9
Mean	158.9	-91.6	0.0	6.9	-120.9	3400.4	67.3	19.7
Upper Bound	246.0	295.8	59.1	18.7	-22.3	10172.2	103.2	21.5
			200-250 Fathom Closure - Grid Model					
Lower Bound	0.6	22.6	-43.3	9.0	11.1	17.9	11.4	7.2
Mean	0.7	34.8	85.4	9.5	12.1	32.5	11.6	12.7
Upper Bound	0.7	47.0	214.1	10.0	13.0	47.1	11.8	18.2
			200-250 Fathom Closure - Traditional Model					
Lower Bound	6.1	-566.1	-284.5	-53.4	-204.3	-8996.6	26.9	31.2
Mean	13.3	-106.1	-3.9	51.1	-111.4	9075.2	58.2	34.2
Upper Bound	20.5	354.0	276.6	155.6	-18.5	27147.1	89.5	37.2

Table 12: Welfare estimates for closures for the Pacific groundfish on the first tow of trip

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
	50-100 Fathom Closure - Grid Model							
Lower Bound	7.0	15.6	6.8	2.7	11.4	7.3	10.1	9.9
Mean	7.3	16.5	7.7	6.3	11.8	7.8	10.3	10.6
Upper Bound	7.6	17.4	8.5	10.0	12.2	8.3	10.5	11.2
	50-100 Fathom Closure - Traditional Model							
Lower Bound	19.6	13.4	3.4	2.3	6.8	4.3	13.7	5.4
Mean	44.0	15.5	3.7	5.7	7.0	4.8	14.3	6.9
Upper Bound	68.4	17.5	3.9	9.2	7.2	5.3	14.9	8.3
	200-250 Fathom Closure - Grid Model							
Lower Bound	0.2	7.0	11.1	5.5	2.5	10.8	4.2	3.8
Mean	0.2	7.3	12.3	11.0	2.6	11.4	4.3	4.0
Upper Bound	0.2	7.7	13.5	16.4	2.6	12.0	4.3	4.3
	200-250 Fathom Closure - Traditional Model							
Lower Bound	1.5	16.5	16.5	17.0	6.9	11.5	13.5	8.1
Mean	3.5	19.0	17.6	43.2	7.1	12.9	14.1	10.0
Upper Bound	5.4	21.5	18.7	69.5	7.3	14.2	14.6	12.0

Supplementary Online Appendix A: Further Results from Monte Carlo Analysis

Bias in Parameter Estimates

Both the grid and traditional models tend to have negative bias in estimating the expected revenue parameter (Figure A1). The parameter is positive and negative bias shrinks it toward zero. Negative bias increases with a higher aggregation level (fewer, larger discrete areas) but the grid model is unaffected by this (Table A1). Both models are able to estimate the "true" parameter for distance fairly accurately across a range of MC model parameterizations (Figure A1). The traditional model tends to exhibit positive bias in the distance parameter with higher aggregation levels while the bias of the distance parameter is negative for the grid model (Figure A1). Note that the distance parameter is negative so positive bias shrinks it toward zero (i.e. reducing its absolute value). Higher aggregation factors do not affect the grid-point model but increases positive bias for the traditional model.

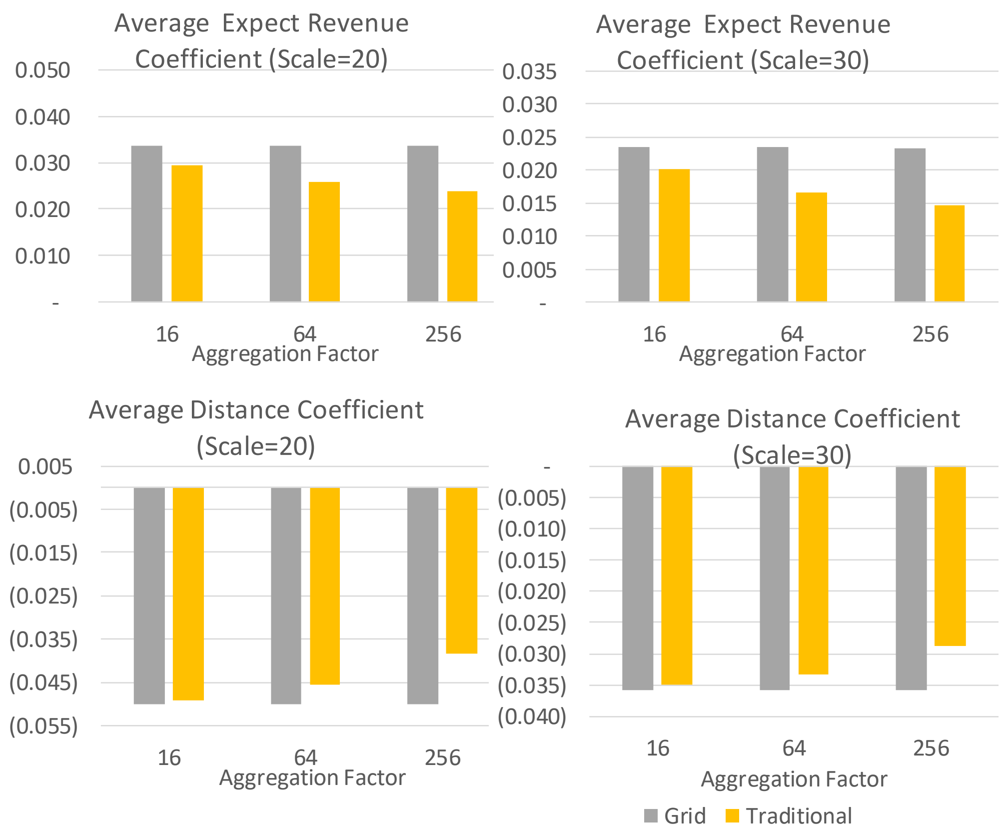

Figure A1: Average expected revenue and distance coefficients for grid and traditional models for different aggregation factors and scale parameters. "True" parameters values are 0.05 (scale=20) and 0.033 (scale=30) for the revenue coefficient and -0.05 (scale=20) and -0.033 (scale=30) for the distance coefficient.

The bias in the expected revenue parameter varies substantially across replications with different randomly generated fish surfaces even for fixed set of MC variables. This is demonstrated by considering bias from 100 replicates with a base case set of MC variables. Bias tends to move in synchrony for the different models preserving relative bias across models even as absolute bias varies (Figure A2).

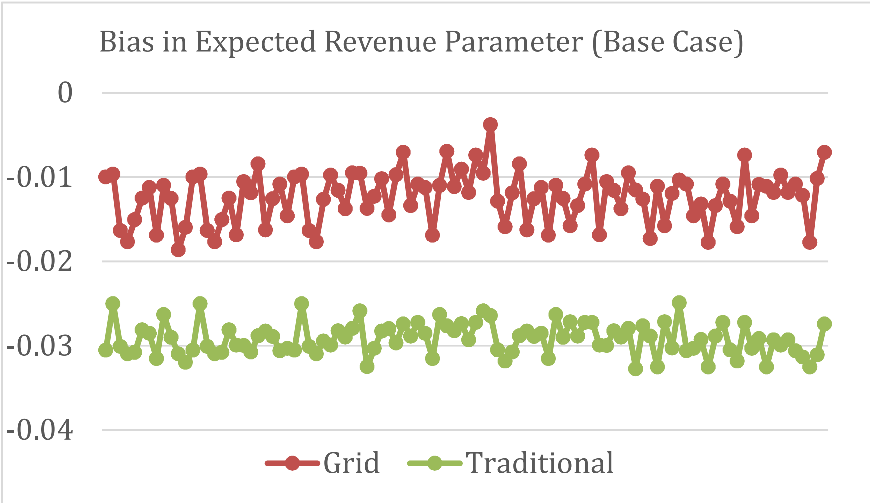

Figure A2: Bias in expected revenue parameter for grid-point and traditional models over 100 replications with a single set of MCl parameters. (Clumps=128, Diffusion=0.750, GEV scale=20, Distance Scale=1, Std. Dev. Exp. Revenue=10, Std.Dev. Actual Revenue=20, Aggregation Factor=64)

The determinants of the revenue parameter bias are shown in Table A1, based on regression of bias on dummy variables for the parameters varied in the Monte Carlo experimental design. Separate regression of bias are run for the two levels of the GEV scale parameters since absolute values of parameters and consequently bias differ on average. Since bias in the revenue parameter is negative (it underestimates the true parameter) a negative coefficient in these regressions indicates a larger absolute bias. Bias becomes more negative (larger in absolute value) as the number of clumps (CL) increases across both the estimators, and the magnitude of the change in bias is consistent across the estimators as well. A higher number of clumps results in a patchier resource surface. The highest diffusion rate, $\delta < 0.20$, increases the absolute bias across both estimators. The impact is significantly smaller on the traditional choice set model which masks fine-scale heterogeneity in the resource. The distance scale parameter, which scales the relative distance between points increases negative bias for both models, particularly when GEV scale is 30. The standard deviation in expected revenues has a marginal impact on some of the estimators but the magnitude of the effects is also small. Increasing the standard deviation in the actual revenues (which creates a gap between the expected revenues that drove the choice and the realized revenues that are used to develop expectations for choices in the RUM model) increases the absolute bias in the revenue parameter across both estimators. However, it has a greater impact on the grid-point estimator. The aggregation factor only impacts the bias in the traditional choice set model, increasing the absolute bias as the grids become coarser. The grid size is not relevant to parameter estimates for the grid model though it does affect the prediction metrics since prediction for both models is done for the coarser areas of the traditional model.

Table A1: Factors influencing the bias for the revenue parameter, β_1. The dependent variable is the mean bias observed in each of the 864 Monte Carlos (per scale parameter) with the data generation parameters serving as independent variables including choice sample size of 100 vs. the base case of 50. The left column contains the results for the GEV scale parameter of 20 ($\zeta = 20$) and the right column possesses the results for the GEV scale parameter of 30 ($\zeta = 30$). A negative coefficient indicates the positive revenue coefficient moves closer to zero (greater absolute bias).

Monte Carlo Variable	Grid Point				Traditional			
	$\zeta = 20$		$\zeta = 30$		$\zeta = 20$		$\zeta = 30$	
Constant	0.006	***	0.004	***	-0.003	***	-0.002	***
Clump 128	-0.009	***	-0.006	***	-0.015	***	-0.011	***
Clump 256	-0.018	***	-0.011	***	-0.018	***	-0.012	***
Diffusion 0.075	0.000		0.001		0.000		0.000	
Diffusion 0.10	0.000		0.000		0.001		-0.001	
Diffusion 0.20	-0.011	***	-0.007	***	-0.004	***	-0.002	**
Distance Scale 3	-0.001		-0.008	***	-0.001	***	-0.008	***
Std. Dev. Exp. Rev. 20	-0.001	***	0.000		-0.001		0.000	
Std. Dev. Act. Rev. 20	-0.008	***	-0.005	***	-0.004	***	-0.002	***
Std. Dev. Act. Rev. 40	-0.021	***	-0.013	***	-0.012	***	-0.007	***
Agg. Factor 64	0.000		0.000		-0.003	***	-0.004	***
Agg. Factor 256	0.000		0.000		-0.005	***	-0.005	***
R-square	0.92		0.65		0.84		0.55	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

The bias determinants for the distance parameter, β_2, are shown in Table A2 and broken down by the two GEV scale parameters. Recall that average bias of the distance coefficient is close to zero or slightly negative for the grid model on average and positive for the traditional model on average. Thus a positive coefficient for the grid model indicates improved performance (lower absolute bias) while the opposite is true for the traditional model. However, bias for both models can be either positive or negative for some combinations of variables in contrast to bias the revenue parameter which is almost always negative for both models. The effects of the higher clumps (CL) value is to increase the negative parameter bias for the grid model, but for the traditional model the effect varies with the scale parameter. The diffusion rates also increases negative bias for the grid model but reduces the positive bias of the traditional model for the highest diffusion rate. The higher distance scale parameter reduces the negative bias for the grid model but increases the positive bias for the traditional model. The standard deviation in expected revenues has a very small effect on both models. The higher standard deviation in the actual revenues (40) increases the negative bias for the grid model but actually reduces absolute bias for the traditional model. Again the magnitude is small. Increasing the aggregation factor increases the bias (shrinks the negative coefficient further toward zero) in the traditional choice set model but has no impact on the grid model.

Table A2: Factors influencing the bias for the distance parameter, β_2. The dependent variable is the mean bias observed in each of the 864 Monte Carlos (per scale parameter) with the data generation parameters serving as independent variables including choice sample size of 100 vs. the base case of 50.. The left column contains the results for the GEV scale parameter of 20 ($\zeta = 20$) and the right column possesses the results for the GEV scale parameter of 30 ($\zeta = 30$). A positive coefficient indicates the negative distance coefficient moves closer to zero which decreases absolute bias for the grid model but increases it for the traditional model.

Monte Carlo Variable	Grid Point				Traditional			
	$\zeta = 20$		$\zeta = 30$		$\zeta = 20$		$\zeta = 30$	
Constant	0.007	***	-0.001	***	0.002	***	-0.004	***
Clump 128	-0.009	***	-0.001	***	-0.006	***	0.000	
Clump 256	-0.009	***	-0.001	***	-0.005	***	0.001	***
Diffusion 0.075	-0.001	***	-0.003	***	0.000		-0.002	***
Diffusion 0.10	-0.003	***	-0.004	***	0.000		-0.002	***
Diffusion 0.20	-0.006	***	-0.004	***	-0.001	*	-0.001	***
Distance Scale 3	0.005	***	0.001	***	0.008	***	0.005	***
Std. Dev. Exp. Rev. 20	0.000		0.000	**	0.001	*	0.000	***
Std. Dev. Act. Rev. 20	0.000		0.000		-0.001		0.000	
Std. Dev. Act. Rev. 40	-0.002	***	-0.001	***	-0.002	***	-0.001	***
Agg. Factor 64	0.000		0.000		0.003	***	0.002	***
Agg. Factor 256	0.000		0.000		0.011	***	0.006	***
R-square	0.31		0.77		0.82		0.88	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Choice Prediction

Prediction performance for both models improve as the aggregation factor is increased simply because there are fewer choices to predict among (Figure A3 and Tables A3-A6). In contrast, the prediction accuracy of both models declines as the number of clumps is increased. Prediction performance also declines as the diffusion rate is increased and with the higher GEV scale variable. The effect of GEV scale is intuitive since the higher scale variable tends to tie revenues less closely to the areas with higher expected revenue making it harder to predict choices based on expected revenue. Higher error in expected revenue, while it does negatively impact prediction performance, has a relatively small effect. Overall this suggests that these models are both relatively robust to noise in specifications of expected revenue in terms of their ability to predict choices.

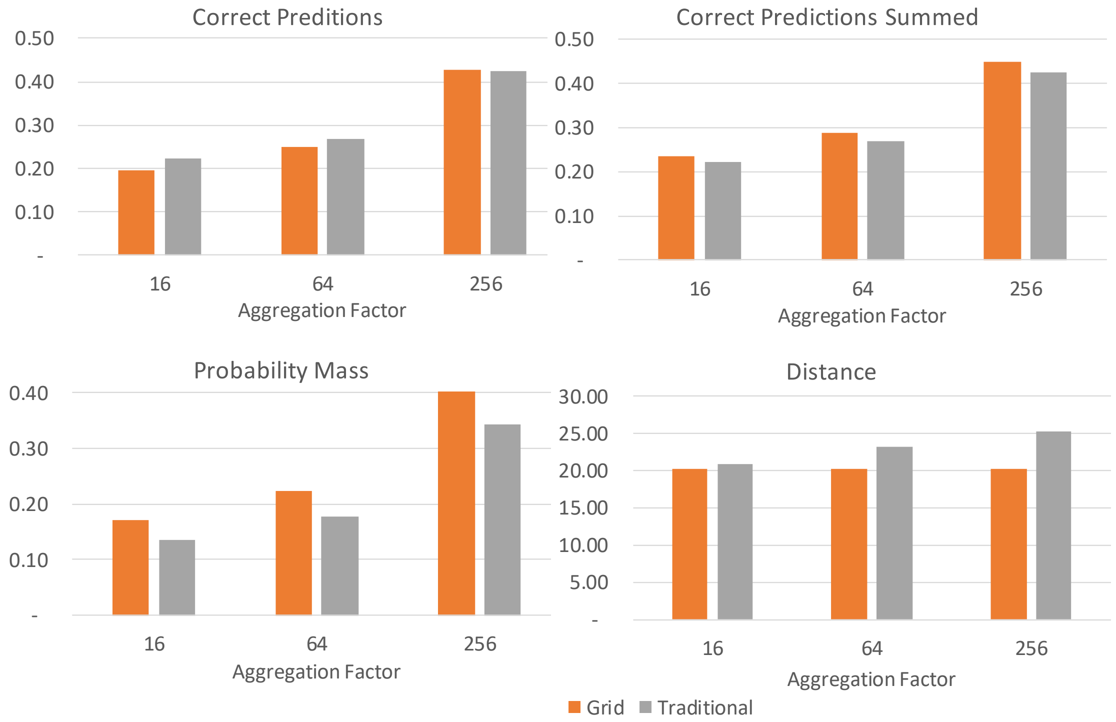

Figure A3: Average results for the different prediction metrics with different aggregation factors with results averaged over all other MC variables and replications.

Table A3: Determinants of Correct Predicted (CP) score broken down by the data generation parameters.

Monte Carlo Variable	Grid Point		Traditional	
Constant	0.311	***	0.352	***
Clump 128	-0.152	***	-0.156	***
Clump 256	-0.197	***	-0.212	***
Diffusion 0.075	0.035	***	0.022	**
Diffusion 0.10	0.003		-0.014	
Diffusion 0.20	-0.067	***	-0.085	***
GEV Scale 30	-0.163	***	-0.185	***
Distance Scale 3	0.066	***	0.055	***
Std. Dev. Exp. Rev. 20	-0.005		-0.007	
Std. Dev. Act. Rev. 20	-0.006		-0.003	
Std. Dev. Act. Rev. 40	-0.022	***	-0.011	
Agg. Factor 64	0.045	***	0.039	***
Agg. Factor 256	0.202	***	0.186	***
R-square	0.711		0.69	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Table A4: Determinants of Correct Predicted Summed (CPS) score broken down by the data generation parameters.

Monte Carlo Variable	Grid Point		Traditional	
Constant	0.368	***	0.352	***
Clump 128	-0.160	***	-0.156	***
Clump 256	-0.220	***	-0.212	***
Diffusion 0.075	0.023	**	0.022	**
Diffusion 0.10	-0.011		-0.014	
Diffusion 0.20	-0.087	***	-0.085	***
GEV Scale 30	-0.193	***	-0.185	***
Distance Scale 3	0.060	***	0.055	***
Std. Dev. Exp. Rev. 20	-0.006		-0.007	
Std. Dev. Act. Rev. 20	-0.003		-0.003	
Std. Dev. Act. Rev. 40	-0.012		-0.011	
Agg. Factor 64	0.043	***	0.039	***
Agg. Factor 256	0.195	***	0.186	***
R-square	0.90		0.88	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Table A5: Determinants of Probability Mass (PM) score broken down by the data generation parameters.

Monte Carlo Variable	Grid Point		Traditional	
Constant	0.247	***	0.232	***
Clump 128	-0.119	***	-0.121	***
Clump 256	-0.166	***	-0.165	***
Diffusion 0.075	0.036	***	0.021	**
Diffusion 0.10	0.015	*	0.001	
Diffusion 0.20	-0.042	***	-0.054	***
GEV Scale 30	-0.139	***	-0.147	***
Distance Scale 3	0.073	***	0.056	***
Std. Dev. Exp. Rev. 20	-0.002		-0.006	
Std. Dev. Act. Rev. 20	-0.006		-0.004	
Std. Dev. Act. Rev. 40	-0.017	**	-0.014	*
Agg. Factor 64	0.045	***	0.035	***
Agg. Factor 256	0.206	***	0.178	***
R-square	0.75		0.68	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Table A6: Determinants of Distance (D) score broken down by the data generation parameters.

Monte Carlo Variable	Grid Point		Traditional	
Constant	10.721	***	12.093	***
Clump 128	6.231	***	4.609	***
Clump 256	8.271	***	5.982	***
Diffusion 0.075	-3.986	***	-3.455	***
Diffusion 0.10	-3.749	***	-3.377	***
Diffusion 0.20	-1.166	*	-1.462	***
GEV Scale 30	10.124	***	8.668	***
Distance Scale 3	13.276	***	15.471	***
Std. Dev. Exp. Rev. 20	0.185		0.167	
Std. Dev. Act. Rev. 20	0.058		-0.066	
Std. Dev. Act. Rev. 40	0.290		-0.137	
Agg. Factor 64	0.004		1.893	***
Agg. Factor 256	-0.001		3.529	***
R-square	0.65		0.77	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Supplementary Online Appendix B: Model Comparison of Grid RUM Models with Alternative Distances and Time Lags Used to Create Expected Revenue

This contains results for the multiple runs of the empirical RUM Grid models with varying radius and time lag window used to construct expected revenue. Results support the use of the smallest radius (3 miles) and a 30 day time window.

Table B1: Parameter Settings for Models Considered

Model	Radius	Time Window (days)
1	5	30[1]
2	5	14
3	3	30
4	3	14
5	10	30
6	10	14

Table B2: Log-likelihoods for Alternative Models

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 8	Fleet 8
Model 1	-2,725	-3,220	-3,425	-2,375	-5,055	-2,057	-11,716	-3,123
Model 2	-2,876	-3,270	-3,581	-2,407	-5,335	-2,090	-12,049	-3,184
Model 3	**-2,383**	**-3,100**	**-3,298**	**-2,283**	**-4,943**	**-1,992**	**-11,064**	**-3,051**
Model 4	-2,595	-3,209	-3,517	-2,388	-5,220	-2,093	-11,662	-3,098
Model 5	-2,906	-3,320	-3,541	-2,493	-5,191	-2,100	-12,534	-3,266
Model 6	-3,091	-3,366	-3,635	-2,504	-5,388	-2,152	-12,930	-3,301

Supplementary Online Appendix C: Charts of Traditional Areas with Grid Points Overlaid

Figure C1: Traditional Areas and Grid Points – Full Extent

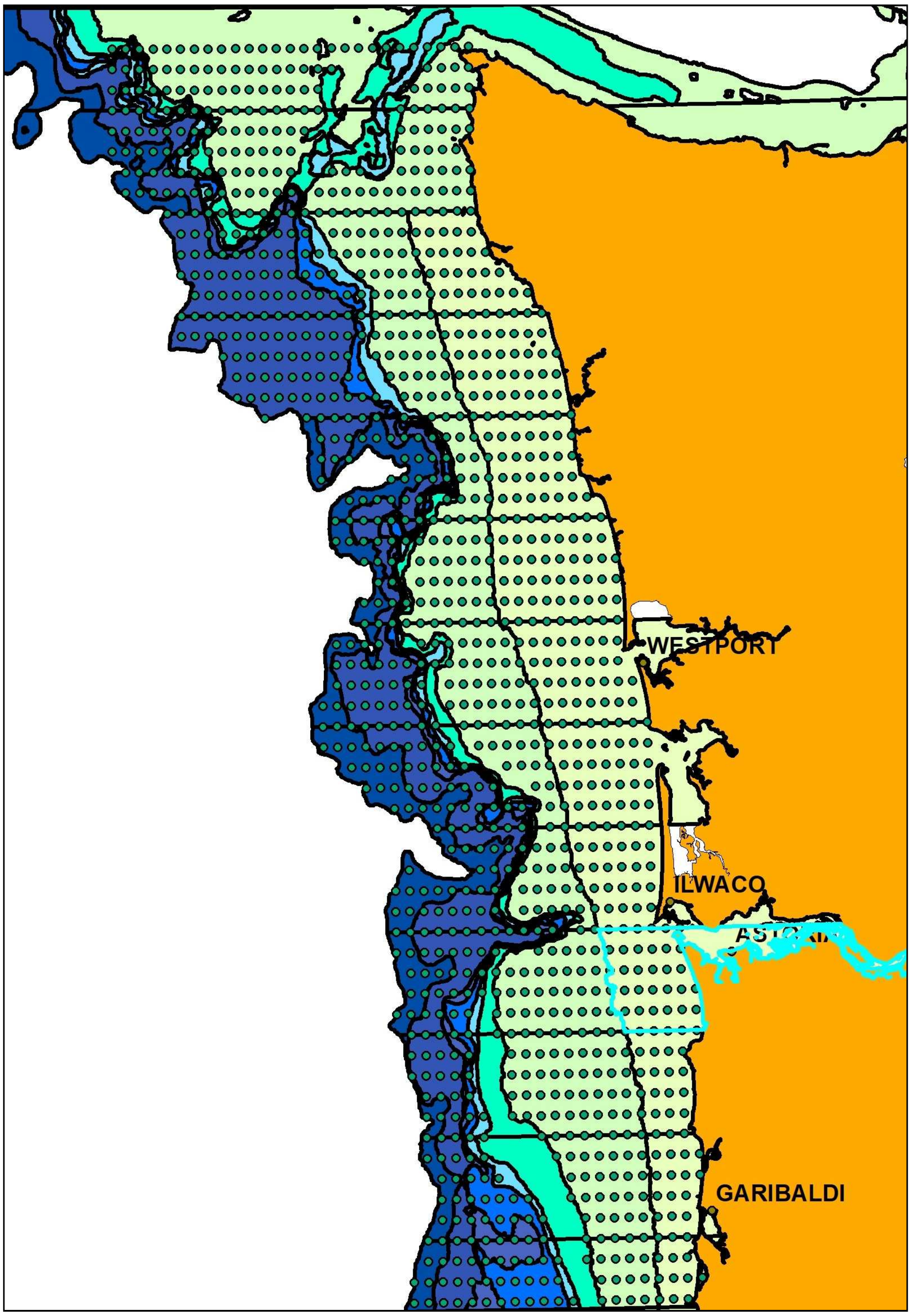

Figure C2: Traditional Areas and Grid Points – Northern Ports

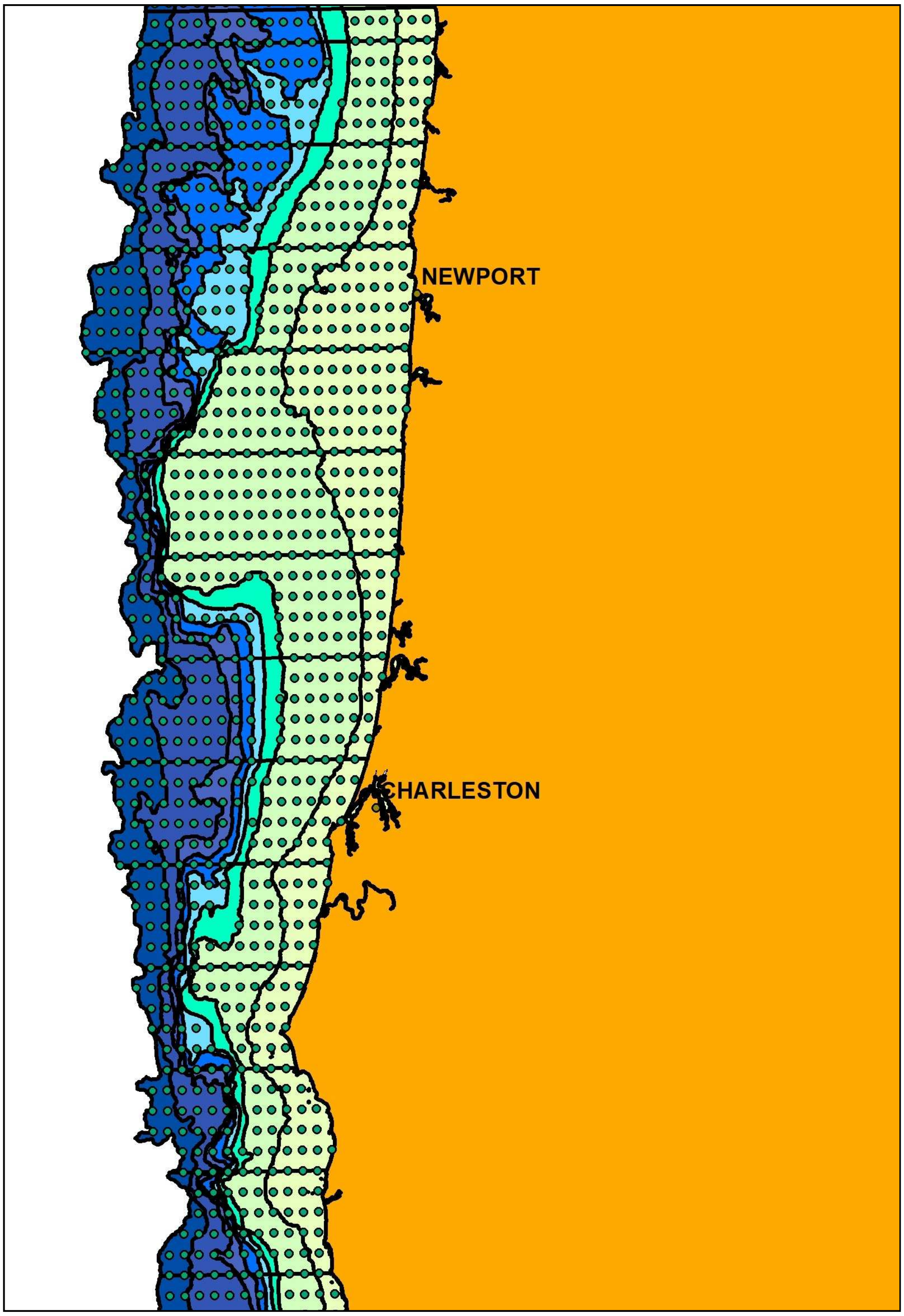

Figure C3: Traditional Areas and Grid Points – North Central Ports

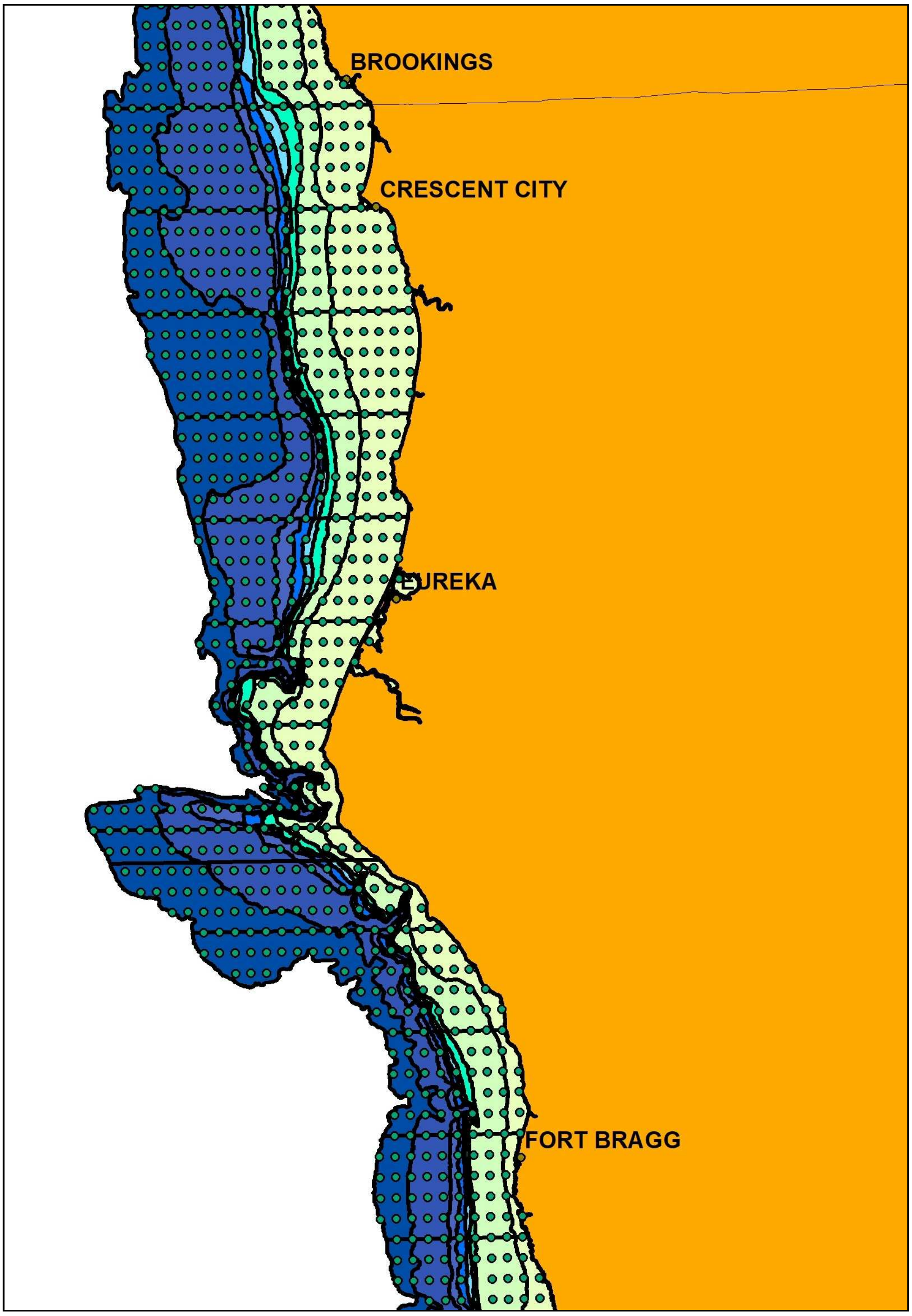

Figure C3: Traditional Areas and Grid Points – South Central Ports

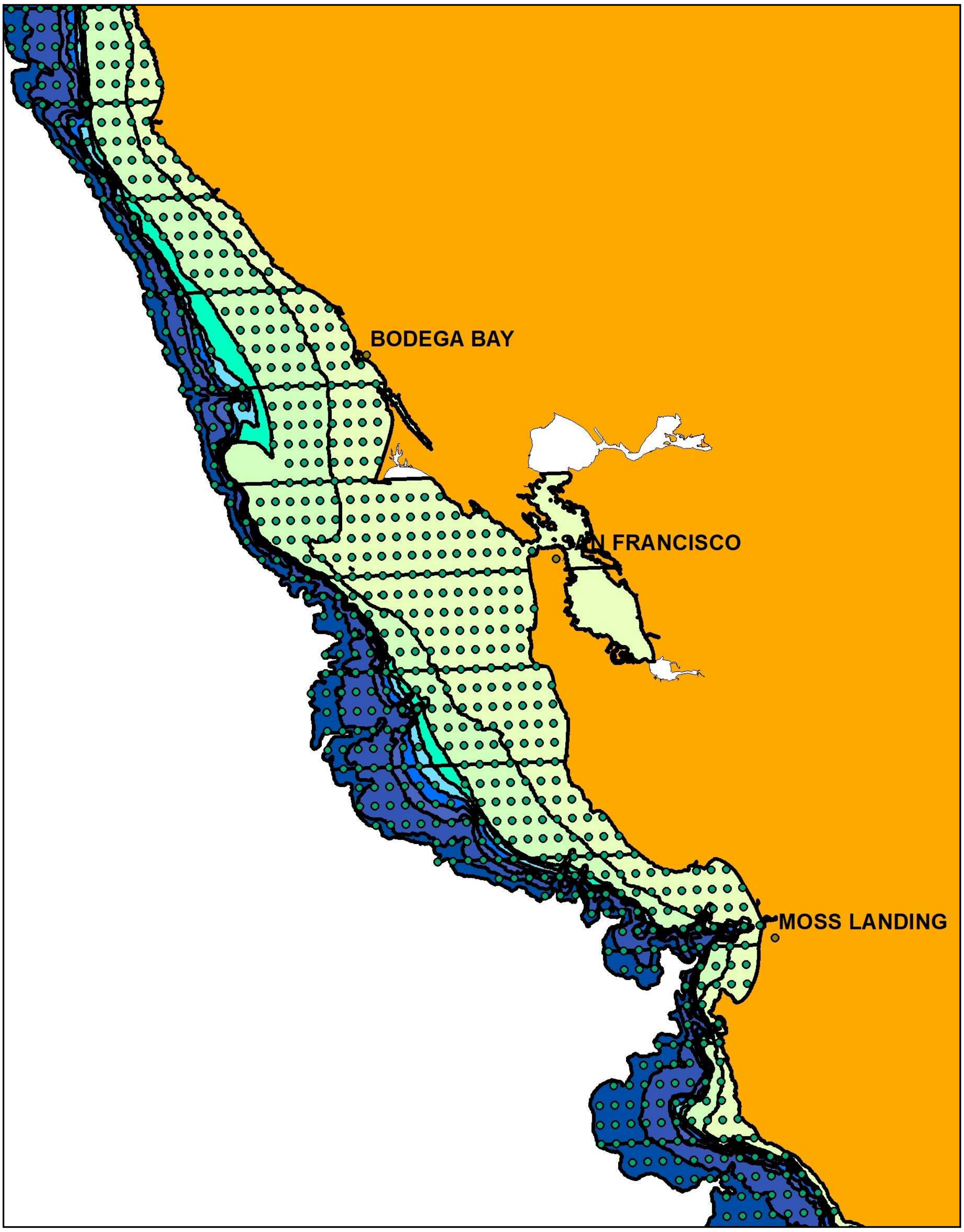

Figure C3: Traditional Areas and Grid Points – Southern Ports

Supplementary Online Appendix D: Results From Models Without Habit Variables and Mixed Logit Models with Normally Distributed Random Parameters

Tables D1-D10 provide RUM model coefficient estimates, predictive accuracy measures, and welfare estimates for RUM models with a standard conditional logit estimation but with "habit" variables removed from the utility function and for RUM models estimated with a mixed logit model with normally distributed random parameters.

For models run without habit variables, coefficients on distance and expected revenue generally exhibit only small changes in magnitude and levels of statistical significance are unaffected in all but a few cases. Likelihood ratio tests reject eliminating the habit variables, though as we note below they do support allowing for heterogeneity in the utility function for some parameters. Predictive accuracy generally declines with elimination of habit variables as expected but the degree of decline varies by fleet. Decline in prediction power without habit variables is much greater when predicting location of the first tow than for subsequent tows suggesting that considering past choices is less important for predicting behavior within the trip.

Welfare measures also change with elimination of habit variables. Changes are relatively small for grid model welfare estimates other than for models where expected revenue parameters are not close to zero and not significant. Welfare estimates for traditional models exhibit much larger changes which is consistent with the much higher volatility and wider bounds of those estimates due to the revenue coefficients being close to zero and often insignificant. For both grid and traditional welfare estimates changes are much smaller for first tows. This is due to the fact that expected revenue parameters tend to be larger and more significant and more robust to model specification.

For the mixed logit models we again find that mean coefficients for distance and expected revenue generally exhibit only small changes in magnitude and levels of statistical significance are mostly unaffected. In many, but not all, cases the expected revenue coefficients from the mixed logit model are smaller relative to the base model but this effect is not consistent across fleets. For most fleets the standard deviation of the coefficients on expected revenue and distance for the first tows is not statistically significant suggesting there in not significant heterogeneity in these coefficients. However, the coefficients for tows other than the first tow often do exhibit statistically significant variation suggesting heterogeneity across vessels. Likelihood ratio tests support the mixed logit model over the base model in terms of overall fit. However, prediction scores for the mixed logit models are slightly lower than for the base model in almost all cases. It is not clear why this would be the case but the result is consistent. Welfare estimates for our policy experiments modeling the cost of closures tend to be more erratic with wider bounds as compared to the base model. Welfare estimates for tows other than the first tow, which were almost always positive for the base conditional logit grid model, are sometimes negative for the mixed logit model and have wider bounds. This is consistent with the smaller coefficient (closer to zero) which can cause welfare estimates to be negative when draws of the revenue coefficient are negative in some policy simulations. Given that our main focus is on comparing the grid and traditional choice formulation opt to use our simpler conditional logit model with habit variables in the main paper and present the mixed logit results in an appendix. In particular, with the prediction scores and policy experiments and welfare estimates this base model provides a more clear illustration of how negative bias in the revenue coefficients can impact ability to predict choice and welfare impacts. We note that Smith (2005) cautions more against modeling preference heterogeneity in isolation (e.g. with random parameters) than against modeling state dependence in isolation.

Table D1a: Parameter estimates from RUM model with no "Habit" variables for the eight fleets estimated within the Pacific groundfish fishery. The top panel contains the results from the grid-point model. The bottom panel contains the results from the traditional choice set model.

	Grid-point Model															
Parameter	Fleet 1		Fleet 2		Fleet 3		Fleet 4		Fleet 5		Fleet 6		Fleet 7		Fleet 8	
Distance	-0.147	***	-0.087	***	-0.082	***	-0.058	***	-0.100	***	-0.108	***	-0.125	***	-0.103	***
DistanceFirst	-0.043	***	-0.024	***	-0.055	***	-0.064	***	-0.052	***	-0.029	***	-0.019	***	-0.024	***
Revenue	0.016	***	0.006	***	0.005	**	0.008	**	0.008	***	0.006	*	0.016	***	0.013	***
Revenue_First	0.036	***	0.017	***	0.011	***	0.005		0.028	***	0.024	***	0.033	***	0.023	***
Dum_Missing	-2.234	***	-1.310	***	-1.574	***	-1.563	***	-1.476	***	-1.747	***	-1.742	***	-1.501	***
	Traditional Model															
Parameter	Fleet 1		Fleet 2		Fleet 3		Fleet 4		Fleet 5		Fleet 6		Fleet 7		Fleet 8	
Distance	-0.140	***	-0.077	***	-0.075	***	-0.050	***	-0.093	***	-0.097	***	-0.117	***	-0.089	***
DistanceFirst	-0.043	***	-0.030	***	-0.054	***	-0.056	***	-0.066	***	-0.028	***	-0.026	***	-0.016	***
Revenue	-0.002		0.001		0.000		0.000		-0.004	*	0.004		0.006	***	0.008	***
Revenue_First	0.009		0.012	***	0.014	***	0.007		0.024	***	0.021	***	0.017	***	0.022	***
Dum_Missing	-2.495	***	-1.480	***	-1.845	***	-1.819	***	-1.993	***	-1.636	***	-2.096	***	-1.533	***
Observations	1451		1185		1496		895		2148		921		7084		1543	

* indicates statistically significant at the 90% level; ** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level.

Table D1b: Likelihood scores and likelihood ratio tests comparing models with and without habit variables.

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
Grid No Habit	-2910	-3202	-3539	-2376	-5464	-2055	-12216	-3188
Grid Base Model	-2348***	-3109***	-3303***	-2293***	-4927***	-1971***	-11136***	-3000***
Traditional No Habit	-2892	-3669	-4344	-2703	-6198	-2418	-18705	-4203
Traditional Base Model	-2753***	-3625***	-4313***	-2658***	-6141***	-2401***	-18234***	-4156***

*** indicates statistically significant at the 99% level for Likelihood Ratio Tests for base model vs restricted model with no habit variables

Table D2a: Parameter estimates from mixed logit estimation of RUM model with normally distributed random parameters for the eight fleets estimated within the Pacific groundfish fishery. The top panel contains the results from the grid-point model. The bottom panel contains the results from the traditional choice set model.

	Grid-point Model															
Parameter	Fleet 1		Fleet 2		Fleet 3		Fleet 4		Fleet 5		Fleet 6		Fleet 7		Fleet 8	
Distance	-0.138	***	-0.085	***	-0.081	***	-0.055	***	-0.094	***	-0.106	***	-0.128	***	-0.114	***
Distance SD	0.052	***	-0.019	***	-0.019	***	0.002		0.036	***	-0.013		0.049	***	0.042	***
DistanceFirst	-0.033	***	-0.023	***	-0.058	***	-0.066	***	-0.051	***	-0.028	***	-0.017	***	-0.027	***
DistanceFirst SD	0.000		-0.001		0.017	**	-0.001		0.001		-0.001		0.000		-0.017	**
Revenue	0.036	***	0.008	***	0.005	*	0.010	***	0.009	***	0.008	*	0.017	***	0.016	***
Revenue SD	-0.028	*	-0.004		0.008		-0.001		0.009		0.009		0.016	***	0.031	***
Revenue_First	0.056	***	0.021	***	0.011	***	0.010	*	0.034	***	0.025	***	0.037	***	0.027	***
Revenue_First SD	-0.003		0.000		-0.004		0.001		-0.003		-0.001		0.002		0.009	
Dum_Missing	0.262		-0.669	***	-0.844	***	-0.753	***	-0.515	***	-1.144	***	-1.030	***	-0.074	
Dum_Missing SD	0.252		0.524	***	0.718	***	0.594	***	0.631	***	0.197		-0.614	***	-0.530	*
Habit_30	3.115	***	1.005	***	1.562	***	1.083	***	1.689	***	0.866	***	1.637	***	1.478	***
Habit_30 SD	1.383	***	1.045	***	1.056	***	0.835	***	0.737	***	-0.698	***	-0.898	***	-0.764	***
Habit_Year	1.705	***	0.699	***	0.903	***	0.741	***	1.120	***	1.091	***	0.938	***	1.116	***
Habit_Year SD	0.854	***	0.811	***	0.978	***	0.504	**	1.110	***	0.765	***	1.040	***	0.464	***

Table D2b: Likelihood scores and likelihood ratio tests comparing conditional logit and mixed logit models.

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
Grid Mixed Logit	-2310***	-3072***	-3255***	-2296	-4844***	-1985	-10839***	-2941***
Grid Base Model	-2348	-3109	-3303	-2293	-4927	-1971	-11136	-3000

Traditional Mixed Logit	-2715***	-3607***	-4281***	-2649**	-6037***	-2377***	-17587***	-4115***
Traditional Base Model	-2753	-3625	-4313	-2658	-6141	-2401	-18234	-4156

** indicates statistically significant at the 95% level; *** indicates statistically significant at the 99% level for Likelihood Ratio Tests for mixed logit model over base model

Table D3: Predictive accuracy measures for tows other than the first tows for the Pacific groundfish RUM model with no "Habit" variables

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
				Correct Prediction (CP)				
Grid-point	0.50	0.20	0.19	0.13	0.24	0.22	0.34	0.29
Traditional	0.19	0.10	0.11	0.14	0.10	0.16	0.05	0.19
				Correct Prediction Summed (CPS)				
Grid-point	0.52	0.24	0.19	0.17	0.27	0.20	0.34	0.23
Traditional	0.19	0.10	0.11	0.14	0.10	0.16	0.05	0.19
				Probability Mass (PM)				
Grid-point	0.36	0.11	0.12	0.10	0.13	0.13	0.20	0.16
Traditional	0.13	0.04	0.06	0.07	0.06	0.09	0.03	0.09
				Distance (D)				
Grid-point	11.5	15.0	14.3	18.2	13.9	14.5	13.9	12.8
Traditional	30.3	40.3	38.7	25.9	27.8	28.1	56.7	33.7

Table D3: Predictive accuracy measures for first tows for the Pacific groundfish RUM model with no "Habit" variables

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
				Correct Prediction (CP)				
Grid-point	0.16	0.04	0.11	0.20	0.16	0.12	0.05	0.14
Traditional	0.20	0.13	0.12	0.15	0.12	0.24	0.06	0.22
				Correct Prediction Summed (CPS)				
Grid-point	0.21	0.09	0.13	0.25	0.23	0.12	0.08	0.14
Traditional	0.20	0.13	0.12	0.15	0.12	0.24	0.06	0.22
				Probability Mass (PM)				
Grid-point	0.14	0.05	0.09	0.10	0.08	0.05	0.05	0.06
Traditional	0.13	0.05	0.07	0.08	0.07	0.11	0.04	0.10
				Distance (D)				
Grid-point	25.5	30.1	17.4	17.7	20.0	34.1	43.3	37.2
Traditional	28.5	31.7	36.9	22.5	22.3	20.7	52.0	27.4

Table D5: Predictive accuracy measures for tows other than the first tows for the Pacific groundfish mixed logit estimation of RUM model with normally distributed random parameters

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
				Correct Prediction (CP)				
Grid-point	0.55	0.20	0.21	0.12	0.25	0.23	0.33	0.27
Traditional	0.19	0.09	0.11	0.14	0.10	0.16	0.05	0.18
				Correct Prediction Summed (CPS)				
Grid-point	0.59	0.22	0.24	0.17	0.29	0.23	0.36	0.25
Traditional	0.19	0.09	0.11	0.14	0.10	0.16	0.05	0.18
				Probability Mass (PM)				
Grid-point	0.46	0.12	0.15	0.10	0.18	0.15	0.24	0.17
Traditional	0.15	0.04	0.06	0.07	0.06	0.09	0.04	0.10
				Distance (D)				
Grid-point	11.0	15.8	14.5	21.3	13.8	15.4	14.9	13.8
Traditional	30.3	40.3	38.8	26.0	28.2	28.5	59.2	34.2

Table D6: Predictive accuracy measures for first tows for the Pacific groundfish mixed logit estimation of RUM model with normally distributed random parameters

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
				Correct Prediction (CP)				
Grid-point	0.40	0.09	0.20	0.18	0.26	0.14	0.13	0.18
Traditional	0.20	0.12	0.12	0.17	0.11	0.24	0.07	0.20
				Correct Prediction Summed (CPS)				
Grid-point	0.47	0.10	0.23	0.22	0.30	0.16	0.17	0.17
Traditional	0.20	0.12	0.12	0.17	0.11	0.24	0.07	0.20
				Probability Mass (PM)				
Grid-point	0.30	0.06	0.13	0.11	0.14	0.07	0.08	0.08
Traditional	0.15	0.05	0.07	0.08	0.07	0.12	0.05	0.10
				Distance (D)				
Grid-point	18.2	30.5	16.4	32.9	15.7	37.8	36.9	37.4
Traditional	28.6	32.0	37.0	22.5	22.9	21.2	54.3	27.9

Table D7: Welfare measures for first tows for the Pacific groundfish RUM model with no "Habit" variables

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
50-100 Fathom Closure - Grid Model								
Lower Bound	37.9	4.2	11.9	-25.4	40.8	-30.9	27.6	29.3
Mean	40.7	29.3	13.3	-4.6	54.1	24.0	28.1	30.8
Upper Bound	43.5	54.4	14.8	16.1	67.4	78.9	28.6	32.3
50-100 Fathom Closure - Traditional Model								
Lower Bound	-260.4	-1020.6	-478.2	-2.6	-184.3	-11.4	35.0	14.9
Mean	209.0	-228.3	-167.3	43.4	-91.3	10.1	36.9	16.1
Upper Bound	678.4	564.0	143.5	89.5	1.8	31.6	38.7	17.4
200-250 Fathom Closure - Grid Model								
Lower Bound	1.2	3.5	18.4	-78.5	12.8	-42.4	10.1	7.9
Mean	1.3	16.4	20.6	-13.2	16.8	33.4	10.3	8.3
Upper Bound	1.3	29.4	22.9	52.1	20.7	109.2	10.5	8.7
200-250 Fathom Closure - Traditional Model								
Lower Bound	-32.1	-1171.5	-2364.1	-14.3	-182.7	-28.3	33.0	26.0
Mean	24.9	-259.4	-826.6	227.8	-89.7	26.0	34.7	28.1
Upper Bound	81.9	652.6	710.9	470.0	3.2	80.3	36.5	30.3

Table D8: Welfare measures for non-first tows for the Pacific groundfish RUM model with no "Habit" variables

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
50-100 Fathom Closure - Grid Model								
Lower Bound	6.5	15.2	7.3	-0.1	12.7	8.0	11.1	13.9
Mean	6.8	16.1	7.9	18.2	13.2	8.6	11.3	14.9
Upper Bound	7.1	17.1	8.6	36.4	13.6	9.2	11.5	15.9
50-100 Fathom Closure - Traditional Model								
Lower Bound	-494.2	13.3	3.6	-0.2	7.1	4.1	11.3	5.6
Mean	-132.0	14.5	3.8	4.8	7.4	4.4	11.8	5.9
Upper Bound	230.2	15.7	4.0	9.9	7.7	4.6	12.4	6.2
200-250 Fathom Closure - Grid Model								
Lower Bound	0.3	6.9	9.4	1.5	2.9	11.2	4.3	4.6

Mean	0.4	7.3	10.1	28.6	3.0	11.9	4.3	4.9
Upper Bound	0.4	7.6	10.7	55.6	3.1	12.7	4.4	5.2
	200-250 Fathom Closure - Traditional Model							
Lower Bound	-61.5	15.8	17.3	0.1	7.6	10.8	12.0	8.3
Mean	-16.6	17.2	18.4	26.3	7.9	11.5	12.6	8.6
Upper Bound	28.4	18.6	19.5	52.6	8.2	12.2	13.2	9.0

Table D9: Welfare measures for first tows for the Pacific groundfish mixed logit estimation of RUM model with normally distributed random parameters

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
50-100 Fathom Closure - Grid Model								
Lower Bound	1.5	19.0	-15.2	1.5	-62.9	3.3	-93.2	-6.0
Mean	29.9	38.6	-5.1	3.2	9.1	20.2	117.4	42.1
Upper Bound	58.3	58.2	5.0	4.9	81.2	37.1	327.9	90.1
50-100 Fathom Closure - Traditional Model								
Lower Bound	-513.6	-64.8	-99.3	-46.5	-37.6	-19.5	-16.1	-16.7
Mean	-135.1	-24.2	-38.6	-15.8	-7.0	-9.2	1.9	24.3
Upper Bound	243.3	16.4	22.0	14.9	23.6	1.1	19.9	65.2
200-250 Fathom Closure - Grid Model								
Lower Bound	0.0	10.7	-37.9	4.5	0.3	-28.6	-5.3	-7.4
Mean	0.8	17.3	-12.2	11.9	14.0	1.0	9.3	8.0
Upper Bound	1.6	23.9	13.5	19.2	27.7	30.5	23.9	23.4
200-250 Fathom Closure - Traditional Model								
Lower Bound	-55.4	-52.7	-531.0	-88.9	-49.7	-116.8	-6.4	-28.7
Mean	-17.5	-6.3	-226.1	-33.5	32.6	-40.5	0.4	39.6
Upper Bound	20.3	40.0	78.8	21.9	115.0	35.7	7.3	107.8

Table D10: Welfare measures for non-first tows for the Pacific groundfish mixed logit estimation of RUM model with normally distributed random parameters

Model	Fleet 1	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 6	Fleet 7	Fleet 8
50-100 Fathom Closure - Grid Model								
Lower Bound	9.6	10.3	-0.7	2.8	11.2	4.0	7.9	-3.6
Mean	10.6	13.5	4.5	9.8	14.5	10.2	11.9	6.2
Upper Bound	11.5	16.7	9.7	16.8	17.8	16.4	15.8	16.1
50-100 Fathom Closure - Traditional Model								
Lower Bound	-728.6	7.8	2.0	0.9	-2.0	0.3	10.5	-10.7
Mean	902.5	14.6	3.0	4.6	5.0	3.4	19.1	20.4
Upper Bound	2533.7	21.3	4.0	8.4	12.0	6.4	27.8	51.6
200-250 Fathom Closure - Grid Model								
Lower Bound	0.2	4.5	3.1	4.2	2.4	4.7	1.0	1.3
Mean	0.2	5.7	10.7	14.9	3.1	10.6	3.7	4.8
Upper Bound	0.2	6.9	18.3	25.7	3.8	16.4	6.5	8.3
200-250 Fathom Closure - Traditional Model								

Lower Bound	-99.3	7.3	10.6	-4.4	-3.7	6.9	13.2	-12.3
Mean	58.7	23.0	14.1	21.0	6.7	12.1	15.3	24.7
Upper Bound	216.7	38.7	17.5	46.4	17.1	17.3	17.5	61.7