# Spatial Search in Commercial Fishing: A Discrete Choice Dynamic Programming Approach

Martin D. Smith*
Assistant Professor of Environmental Economics
Nicholas School of the Environment and Earth Sciences
Duke University
Box 90328
Durham, NC 27708
ph: (919) 613-8028 fax: (919) 684-8741
marsmith@duke.edu

and

Bill Provencher
Associate Professor
Department of Agricultural and Applied Economics
University of Wisconsin - Madison
provencher@aae.wisc.edu

* Indicates the contact author.

Selected Paper
American Agricultural Economics Association Annual Meeting
July 2003, Montreal, Quebec, Canada

JEL Codes: Q22

Abstract

We specify a discrete choice dynamic programming model of commercial fishing participation and location choices. This approach allows us to examine how fishermen collect information about resource abundance and whether their behavior is forward-looking.

**Spatial Search in Commercial Fishing:**
**A Discrete Choice Dynamic Programming Approach**

## I. Introduction

In recent years, empirical models of space have become popular in resource economics. Many studies now use spatial econometrics to explore unobserved attributes of the resource base or unobserved attributes of decision-makers that are spatially autocorrelated. The foundation of the spatial econometric approach is the existence of continuous unobserved gradients that are economic, physical, biological, or some combination thereof. However, many renewable resources do not fit this description. Biological resources are often "patchy" and occur in discrete clumps [9]. In a patchy setting, breaks in habitat generate discontinuities such that the spatial structure may be very different from the continuous underpinnings of a geospatial model or a spatial econometric approach. Thus, a patchy resource warrants a different sort of empirical model. For this reason, discrete choice analysis has been a common approach to modeling spatial location choice in commercial fisheries [1,2,3,5,6,7,11,12,13,14].

The results of discrete choice fisheries models are quite consistent across studies but leave substantial variation in spatial behavior unexplained. Studies generally find that higher expected revenues increase the probability of visiting a site, greater travel distances decrease this probability, and site-specific constants have considerable explanatory power. Several authors also find evidence of temporally autocorrelated spatial behavior [5,11,12]. This is not surprising given that individual harvesters inevitably have heterogeneous information sets about resource abundance and quality based on their own experiences. The existing literature focuses on expected revenues modeled as common knowledge. To date, no one has estimated a structural model that accounts for harvester habit persistence based on private information acquisition, nor

has anyone structurally accounted for the apparent unobserved spatial heterogeneity in the resource base as indicated by location-specific constants. In this paper, we address both of these shortcomings with an explicit model of how individual-specific information sets affect participation decisions and location choices. We also allow for active and purposeful information gathering by estimating a discrete choice dynamic programming (DCDP) model of spatial search.

Similar approaches have been applied in the marketing literature in which consumers sample goods to learn about their characteristics [4]. Consumers have some common knowledge or public information about goods but do not know all characteristics of goods with certainty in the product "space." Often consumers must purchase goods to learn more about them and thus generate private information. Similarly, resource harvesters have some public information about resource abundance and quality but inevitably gain private information about specific locations by visiting them. The two problems are isomorphic in that consumers receive some utility from consuming goods as they learn about their characteristics, just as harvesters catch some fish as they gather information. However, a significant complication exists in renewable resource harvesting that is absent in the product marketing context. In particular, spatially explicit information about the resource stock and quality decays over time [10]. This decay, which may be slow or rapid, is potentially due to migration of organisms, oceanographic factors, and the harvest activities of the fleet. Thus, harvesters face an interesting dilemma: is it worth devoting effort explicitly to spatial search or are harvesters better off simply incorporating information that they gain from whatever locations they would choose anyway? In essence, we are trying to uncover whether harvesters are forward-looking about information gathering. More specifically, is there evidence that harvesters forego short-run gains to pursue a long-run search strategy?

To assess the importance of private information and spatial search strategies, we analyze logbook and landings ticket data from the California red sea urchin fishery. This fishery is an ideal case study because urchin harvesters are owner-operators whom we observe on a daily basis over a long time period (some up to 10 years). Moreover, adult sea urchins occur in patches and are sedentary. Thus, private information is not necessarily so short-lived that forward-looking search is futile.

To model information sets, we use an updating framework in which expected revenues are distributed normally. This parametric assumption for trip revenues is analytically tractable and implicitly accounts for abundance and quality-based price effects. At the beginning of the sample period, the prior belief for each individual is the public information (backward-looking revenues averaged across the entire fleet) in each location. Then in each subsequent period, new public information combines with private information to update each harvester's prior on each patch and produce an individual-specific posterior estimate of revenue distributions in each location. The posteriors enter into the harvester utility functions, and in turn, affect location choices. The estimation strategy follows Provencher and Bishop [8], which applied DCDP to recreation behavior. This allows us to consider explicitly the possibility of forward-looking search behavior.

In section II, we develop a dynamic programming model for discrete fishing decisions that accounts for learning about resource abundance. In Section III, we discuss the California sea urchin fishery background, our unique data set, and the empirical model. Finally, in Section IV we discuss our estimation strategy.

## II. A Model of Fishing Search Behavior

### A. The information problem

For now we assume that net revenues associated with a trip to site j are distributed normally with mean $\mu_j$ and variance $\sigma_j^2$. We assume that the time scale of change in the fishery is sufficiently long that the distribution of revenue is constant over the observation period of the study. Suppose a boat does not know the distribution of revenues at site *j*. Let $S_j(t-1)$ denote the total trips taken to site *j* by all boats, up until time *t-1*, and let $\mathbf{r}_j$(t-1) denote an associated vector of trip revenues (it has dimension $S_j(t-1)$), with revenues from the $s^{th}$ trip denoted by $r_{js}(t-1)$. The relevant likelihood function for the mean revenue associated with a trip to site *j* at time *t* is then,

$$L_{jt}\left(\sigma_j, \mu_j \mid \mathbf{r}_j(t-1)\right) = \prod_{S_j(t-1)} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{\left(r_{js}(t-1)-\mu_j\right)^2}{2\sigma_j^2}\right), \qquad (1.1)$$

The relevant maximum likelihood statistics are simply,

$$\overline{r}_j(t-1) = \frac{\sum_{S_j(t-1))} r_{js}(t-1)}{S_j(t-1)}, \quad s_j^2 = \frac{\sum_{S_j(t-1)} \left(r_{js}(t-1)-\overline{r}_j(t-1)\right)^2}{S_j(t-1)} \qquad (1.2)$$

One might imagine that any boat *k* uses the statistics in (1.2) in its decision about where to fish; each trip adds information about a site, and this additional information is relevant to *future* trip

decisions, and so this information-gathering aspect of a trip makes the decision problem dynamic. Yet if these were the statistics actually used by boats in their calculations of the distribution of revenues, the search motive would be quite small so long as $S_j\left(t-1\right)$ is reasonably large; each trip influences the estimate of the distribution of revenues to such a small degree that the dynamic aspects of the trip decision are slight, and probably undetectable by an outside observer.

It seems reasonable, though, that in evaluating its distribution of revenues associated with site *j*, boat *k* will modify its evaluation of the distribution of net revenues to reflect private information from its past trips to site *j*. A useful way to present this modification is via a weighting scheme in which boat *k* differentially weights observations of net revenue from its own history of trips to site j and observations of net revenue from trips taken to site j by other boats. Let $S_j^k\left(t-1\right)$ denote all trips taken by boat *k* to site *j* up until time t-1, and let $S_j^{-k}\left(t-1\right)$ denote the total number of trips taken by all other boats to site *j* up until time t-1. We partition the vector of revenues associated with trips in a similar fashion. Suppose the boat applies the following weighting scheme:

$$L_{jt}\left(\sigma_j,\mu_j \mid \mathbf{r}_j^{-k}(t-1),\mathbf{r}_j^k(t-1)\right)=\left[\prod_{S_j^{-k}(t-1)}\frac{1}{\sqrt{2\pi\sigma_j^2}}\exp\left(-\frac{\left(r_{js}^{-k}\left(t-1\right)-\mu_j\right)^2}{2\sigma_j^2}\right)\right]^{\left(1-\omega^k\right)} \cdot\left[\prod_{S_j^{k}(t-1)}\frac{1}{\sqrt{2\pi\sigma_j^2}}\exp\left(-\frac{\left(r_{js}^{k}\left(t-1\right)-\mu_j\right)^2}{2\sigma_j^2}\right)\right]^{\omega^k} \tag{1.3}$$

where $\omega^k$ is the weight the boat assigns to the likelihood function associated with its own past revenues. In this case the relevant statistics are (where we drop the time index on revenues to reduce the notational clutter):

$$\begin{aligned} \overline{r}_j &= \omega^k \left[ \frac{\sum_{S_j^k(t-1))} r_{js}^k}{S_j^k(t-1)} \right] + \left(1-\omega^k\right) \left[ \frac{\sum_{S_j^{-k}(t-1))} r_{js}^{-k}}{S_j^{-k}(t-1)} \right], \\ s_j^2 &= \omega^k \left[ \frac{\sum_{S_j^k(t-1)} \left(r_{js}^k - \overline{r}_j^k\right)^2}{S_j^k(t-1)} \right] + \left(1-\omega^k\right) \left[ \frac{\sum_{S_j^{-k}(t-1)} \left(r_{js}^{-k} - \overline{r}_j^{-k}\right)^2}{S_j^{-k}(t-1)} \right] \end{aligned} . \qquad (1.4)$$

Note that the means used to estimate the variance of net revenues are groupwise means. Note too that the weighting parameter is individual-specific, but one could assert that this parameter is the same for all boats.

**B. Illustration of the decision problem**

Now consider the decision problem faced by a boat. We wish to demonstrate two points using a simple two-period, two-site model. The first is that, for the usual random utility model in which revenues enter linearly, only the calculation of mean revenues $\overline{r}_j$ is relevant.[1] And the second is that the decision problem is dynamic because boats understand that information on mean revenues collected today influence expected future returns.

We consider a boat with three choices: to fish at either site 1 or site 2, or to stay in port. Setting the baseline systematic portion of utility associated with staying in port equal to zero, we denote the expected utility associated with a trip to site *j* by the boat on day *t* as

[1] Note that for a repeated choice model such as the one considered here, it seems reasonable to argue that only expected net revenues enter the utility function.

$$U_j = \beta x_j(t) + \bar{r}_j\ (t-1), \tag{1.5}$$

where $x_j(t)$ denotes exogenous variables relevant to the day $t$ decision. Note that we have not appended the usual variable contemporaneously observed by the boat but unobserved by the analyst; at this point we are simply illuminating the essential nature of the search problem. In the last period of our model (period T), the boat merely maximizes current expected utility:

$$V_1\left(\bar{\mathbf{r}}(T)\right) = \max\left[0,\ \beta x_1(T) + \bar{r}_1(T-1),\ \beta x_2(T) + \bar{r}_2(T-1)\right]. \tag{1.6}$$

The dynamic decision problem faced by the boat in period T-1 is then,

$$V_{T-1}\left(\bar{\mathbf{r}}_{k,T-1}, \mathbf{s}_{k,T-1}\right) = \max\begin{bmatrix} 0 + E_{\bar{\mathbf{r}}(T-2),\mathbf{s}^2(T-2)}\left\{V_T\left(\bar{\mathbf{r}}(T-1)\right)\right\}, \\ \beta x(T-1) + \bar{r}_1(T-2) + E_{\bar{\mathbf{r}}(T-2),\mathbf{s}^2(T-2)}\left\{V_T\left(\bar{\mathbf{r}}(T-1)\right)\right\}, \\ \beta x(T-1) + \bar{r}_2(T-2) + E_{\bar{\mathbf{r}}(T-2),\mathbf{s}^2(T-2)}\left\{V_T\left(\bar{\mathbf{r}}(T-1)\right)\right\} \end{bmatrix} \tag{1.7}$$

$$s.t. \qquad (1.4)$$

The subscripts on the expectation operator in (1.7) indicate that the expectation is conditional on current information held by boat, as embodied in the vectors of sufficient statistics $\bar{\mathbf{r}}(T-2)$ and $\mathbf{s}^2(T-2)$. To grasp the dynamic nature of the problem, suppose that the current (time T-1) expected utility of all three choices is zero, and that the site-specific values of **x** do not change over time. Which option would be preferred? The answer turns on the effect of a trip on the expected value of future returns. Under the no-trip choice at time T-1, the value of a trip at time T is known, and is zero by the assumptions just stated:[2]

[2] To reduced notational clutter, for the remainder of the discussion we suppress the subscripts indicating that the expectation is conditional on the state variables.

$$
\begin{aligned}
E\left\{V_T\left(\bar{\mathbf{r}}\left(T-1\right)\right)\right\} &= E\left\{\max\left[0,\ \beta x\left(T-1\right)+\bar{r}_1\left(T-2\right),\ \ \beta x_2\left(T-1\right)+\bar{r}_2\left(T-2\right)\right]\right\} \\
&= \max\left[0, \beta x_1\left(T-1\right)+\frac{\sum_{S_1(T-2)} r_{1s}\left(T-2\right)}{S_1\left(T-2\right)},\ \ \beta x_2\left(T-1\right)+\frac{\sum_{S_2(T-2)} r_{2s}\left(T-2\right)}{S_2\left(T-2\right)}\right] \quad ,(1.8) \\
&= \max\left[0,0,0\right]=0
\end{aligned}
$$

where to simplify the illustration we assume that the boat uses the statistics in (1.2). By contrast, given the decision to take a trip to site 1 at time T-1, with net revenues denoted by $r_{1(T-1)}$ the expected value function at time T is,

$$
\begin{aligned}
E\left\{V_T\left(\bar{\mathbf{r}}\left(T-1\right)\right)\right\} &= E\left\{\max\left[0, \beta x_1+\bar{r}_1\left(T-1\right),\ \ \beta x_2+\bar{r}_2\left(T-2\right)\right]\right\} \\
&= E\left\{\max\left[0, \beta x_1+\frac{\left(\sum_{S_1(T-2)} r_{1s}\left(T-2\right)\right)+r_{1(T-1)}}{S_1\left(T-2\right)+1},\ \ \beta x_2+\frac{\sum_{S_2(T-2)} r_{2s}\left(T-2\right)}{S_2\left(T-2\right)}\right]\right\} \\
&= E\left\{\max\left[0, \beta x_1+\frac{\left(\sum_{S_1(T-2)} r_{1s}\left(T-2\right)\right)+r_{1(T-1)}}{S_1\left(T-2\right)+1},\ \ 0\right]\right\} \\
&= 0\cdot\Pr\left(r_{1(T-1)}<\bar{r}_1\left(T-2\right)\right)+\left[\int_{\bar{r}_1(T-2)}^{\infty}\left(\bar{r}_1\left(T-1\right)-\bar{r}_1\left(T-2\right)\right)\cdot\hat{g}\left(r_{1(T-1)};\bar{r}_1\left(T-2\right),s_1^2\left(T-2\right)\right)\cdot dr_{1(T-1)}\right]\cdot\Pr \\
&= \left[\int_{\bar{r}_1(T-2)}^{\infty}\left(\bar{r}_1\left(T-1\right)-\bar{r}_1\left(T-2\right)\right)\cdot\hat{g}\left(r_{1(T-1)};\bar{r}_1\left(T-2\right),s_1^2\left(T-2\right)\right)\cdot dr_{1(T-1)}\right]\cdot\frac{1}{2}
\end{aligned}
$$

(1.9)

Here is the logic behind this expression. There is a .5 probability that a trip to site 1 in period T will generate net revenue lower than the mean net revenues generated from previous trips to site

1, $r_{1(T-1)} < \overline{r}_1(T-2)$. In this case, the new estimate of mean revenues to site 1 are lower than before, $\overline{r}_{k1T} < \overline{r}_{k1,T-1}$,and a repeat trip is not made to site 1 in period T, because the updated expected utility from a trip to site 1 is less than zero (recall that by assumption, at time T-1 the expected utility of a trip is zero); the expected utility in period T is then zero (which is the expected utility from either staying in port or going to site 2). On the other hand, there is a .5 probability that the net revenue from a trip to site 1 at time T-1 is *greater* than the mean of net revenues generated from previous trips to site 1. In this case the expected utility from a trip to site 1 at time T is greater than zero, reflecting the updated information about net revenues engendered by the trip to site 1 at time T-1. The increase in the utility is merely the difference between the expected net revenues at time T and the expected net revenues at time T-1, $\overline{r}_{k1T} - \overline{r}_{k1,T-1}$; this difference is positive because net revenue from a trip to site 1 at time T-1 is "greater than average". To calculate the expected gain from a trip to site 1 at time T, conditional on finding that a trip to site 1 at time T-1 yields net revenue greater than expected, we integrate the net revenue difference from $\overline{r}_{k1,T-1}$ to infinity, weighted by the prior normal probability distribution, $N\left(\overline{r}_{k1,T-1}, s^2_{k1,T-1}\right)$.

The upshot is that given current expected utility is the same for staying in port and going to site 1, and that exogenous variables *x* are not changing over time, the boat will choose to take a trip to site 1, because there is a positive payoff to learning; in particular, the boat may find that it is underestimating the mean net revenues associated with site 1. A similar situation applies to site 2, and it is not difficult to understand how the choice will be made between visiting sites 1 and 2 at time T-1: the site with greater variance will be the preferred site, because this is the site for which the boat has the most to learn from a trip. Indeed, from the last two lines of (1.9) it is

easy to see why the standard error is included as a state variable: it is necessary in the calculation of the expected future gain associated with a current trip. Moreover, it is not too difficult to imagine that if we extend this model to more than two periods, the information updated by a trip is not only the mean revenue, but the variance in revenues as well, and it follows that this influences the current trip decision. Nonetheless, this example illustrates that updating the estimate of the variance is of second-order importance, and this has implications for developing an estimable model of the boat's decision process, as discussed below.

In this simple model we have two state variables associated with each site (the mean and standard error of net revenues). If we distinguish the fleet mean and standard error from the boat's private mean and standard error, we end up with four state variables per site, although only the boat's private mean and standard error have implications for what Bellman calls the "curse of dimensionality", because only private variables are affected by the boat's current trip decision (here we assume that the net revenue data for the fleet is sufficiently large that the boat's own data has an imperceptible effect on them, and so fleet state variables are treated as exogenous to the decision problem). Of course, the mean and variance of the boat's revenues are calculated statistics, and so in a practical application the dimensionality of the problem depends on the underlying "primitive" state variables –that is, the variables necessary to calculate these statistics. These are (1) the number of trips, (2) the aggregate net revenue across these trips, and (3) squared aggregate net revenues. This suggests that the model is practical only if the choice set is small, say 2 or at most 3 sites. There are two alternatives that would ease the burden of the dimensionality problem.

The first is to abandon the formal structure of the boat's forecasting problem, and to instead include in the site-specific utility function an unobserved state variable that evolves

according to a transition equation (to be estimated) affected by the trip decision. In this approach each site is associated with at most a single state variable. Alternatively, we might capture the dynamics with several state variables that enter the utility functions of all sites. The point here is that such a formulation allows us to manage the curse of dimensionality while still accounting for decision dynamics, albeit without a clear interpretation of the dynamics.

The second approach –one that we employ in this paper–is to simply assert that the agent knows the variance in revenues. Insofar as we do not actually observe $\sigma^2$, we can represent this variable in our modeling as the sample variance $s^2$ as calculated from the entire sample (i.e., from the trips taken by the entire fleet). In this case the only state variable relevant to the agent is the mean revenue –the agent takes trips to learn only about the true mean of net revenues, not to learn about the variance of revenues –and every trip taken by the agent reduces the standard error around the mean of net revenues (because this variance is $\sigma^2 / N$). Note that in the simple two-period example considered above, it is *this* problem that concerns the boat; the issue of a better estimate of the variance itself does not "kick in" until the model reaches three periods, which is why we refer to the matter of improving estimates of $\sigma^2$ as "second-order". In a sense, then, this model proposes that the typical fisher accepts that variability of revenues is the same for all boats, but searches to determine whether, when it comes to generating revenues from patch *j*, his boat is "better than average". It follows of course that a boat may find itself better than average in some patches and worse than average in others.

We now construct the estimable dynamic structural model for such a problem. The boat's expected utility at time t from a trip to site j is (with a slight change in notation):

$$U_{jt} = \beta x_{jt} + \overline{r}_{jt} + \varepsilon_{jt} \tag{1.10}$$

where the mean net revenue is calculated from the weighting of past trips, as in (1.4), and the disturbance term is iid Gumbel distributed. The boat's dynamic decision problem is then to choose the site (patch) to solve

$$V_t\left(\overline{\mathbf{r}}_t\right) = \max\left[\left\{\beta x_{jt} + \overline{r}_{jt} + \varepsilon_{jt} + \delta E_{r_j,\varepsilon|\mathbf{R}_t,\mathbf{N}_t}\left\{V_{t+1}\left(\overline{\mathbf{r}}_{t+1}\right);\boldsymbol{\sigma}^2\right\}\right\}_J\right] \tag{1.11}$$

subject to the evolution of means as described above. The parameter $\delta$ is a discount factor which we use in the estimation to test whether anglers are forward-looking; failure to reject the hypothesis $\delta = 0$ suggests behavior is static. The subscript on the expectation operator indicates that the expectation of the seasonal value of fishing at time t+1 is taken over the random variable $\varepsilon$ and current trip revenues $r_j$, conditional on the current state of total revenues across all sites, $\mathbf{R}_t$ and total trips across all sites, $\mathbf{N}_t$. In particular, the observation of current trip revenues for site *j* affects the estimate of mean revenues tomorrow at site *j*, $\overline{r}_{j,t+1}$. For notational convenience, we define

$$E_{r_j,\varepsilon|\mathbf{R}_t,\mathbf{N}_t}\left\{V_{t+1}\left(\overline{\mathbf{r}}_{t+1}\right);\boldsymbol{\sigma}^2\right\} = v_{t+1}\left(j,\mathbf{R}_t,\mathbf{N}_t,\boldsymbol{\sigma}^2\right)\ ; \tag{1.12}$$

expected future returns depend upon the current trip decision *j* and the current states of aggregate revenues and total trips. We restate the boat's problem,

$$V_t\left(\overline{\mathbf{r}}_t\right) = V_t\left(\mathbf{R}_t,\mathbf{N}_t\right) = \max_j\left\{\beta x_{jt} + \overline{r}_{jt} + \varepsilon_{jt} + \delta v_{t+1}\left(j,\mathbf{R}_t,\mathbf{N}_t,\boldsymbol{\sigma}^2\right)\right\}\ . \tag{1.13}$$

Shifting (1.13) forward one period, and assuming that the disturbance terms are iid Gumbel-distributed with location parameter equal to zero and scale parameter $\mu$, we have,

$$E_\varepsilon\left\{V_{t+1}\left(\mathbf{R}_{t+1},\mathbf{N}_{t+1}\right);\boldsymbol{\sigma}^2\right\} = \frac{1}{\mu}\ln\left\{\sum_{j=0}^{J}\exp\mu\left[\beta x_{j,t+1} + \overline{r}_{j,t+1}\right] + \delta v_{t+2}\left(j,\mathbf{R}_{t+1},\mathbf{N}_{t+1},\boldsymbol{\sigma}^2\right)\right\} \tag{1.14}$$

substituting (1.14) into (1.12) yields,

$$v_{t+1}\left(j,\mathbf{R}_t,\mathbf{N}_t,\boldsymbol{\sigma}^2\right)=E_{r_j|\mathbf{R}_{t+1},\mathbf{N}_{t+1}}\left\{\frac{1}{\mu}\ln\left\{\sum_{j=0}^{J}\exp\mu\left[\beta x_{j,t+1}+\overline{r}_{j,t+1}\right]+\delta v_{t+2}\left(j,\mathbf{R}_{t+1},\mathbf{N}_{t+1},\boldsymbol{\sigma}^2\right)\right\}\right\} \quad (1.15)$$

which states that, given we know $v_{t+2}(\cdot)$ from the previous stage of the recursion, calculation of $v_{t+1}(\cdot)$ is a fairly straightforward affair involving integration over the random variable $r_j$, which has distribution $N\left(\overline{r}_{j,t+1},\sigma_j^2/N_{jt}\right)$.

Solution of the dynamic problem represented by (1.13)-(1.15) is the greatest obstacle to the estimation problem. With $v_t(\cdot)$ known, maximum likelihood estimation of the dynamic problem is analogous to that for static multinomial logit models. The probability that a boat visits site *k* at time *t* is given by:

$$\Pr\left(k\,|\,\mathbf{R}_t,\mathbf{N}_t;\boldsymbol{\beta},\boldsymbol{\sigma}^2,\mu,\delta\right)=\frac{\exp\mu\left[\beta x_{k,t+1}+\overline{r}_{k,t+1}\right]+\delta v_{t+2}\left(k,\mathbf{R}_{t+1},\mathbf{N}_{t+1},\boldsymbol{\sigma}^2\right)}{\sum_{j=0}^{J}\exp\mu\left[\beta x_{j,t+1}+\overline{r}_{j,t+1}\right]+\delta v_{t+2}\left(j,\mathbf{R}_{t+1},\mathbf{N}_{t+1},\boldsymbol{\sigma}^2\right)} \quad (1.16)$$

Letting $y_{mt}$ denote the trip decision of boat *n* on day *t*, with *m*=1,…*M*, the likelihood of the sample is

$$\ln L=\sum_{M}\sum_{T}\Pr\left(y_{mt}\,|\,\mathbf{R}_{mt},\mathbf{N}_{mt};\boldsymbol{\beta},\boldsymbol{\sigma}^2,\mu,\delta\right) \quad (1.17)$$

Finally, note that we can introduce heterogeneity not only in preferences, but also in forward-looking behavior using either latent class analysis or mixed logit.

## III. Empirical Setting and Data

This work draws on a data set that was constructed to trace participation and location choice behavior in the northern California red sea urchin fishery. Sea urchins are harvested for their roe, which is a delicacy in Japanese cuisine called 'uni.' The complete data set tracks the daily decisions of about 1000 harvesters over a ten-year period. The data set used in this paper

consists of daily observations on individual California sea urchin divers from 1988 to 1990, including departing port location, diving location, and revenues. The focus is on the northern California fishery, since regulators have shown more concern about its potential collapse than the southern California fishery, and as a result, future spatial management in the north is more likely.

The sea urchin fishery is an ideal setting in which to explore sequences of discrete participation and location choices using a dynamic model. As a dive fishery composed mostly of owner-operators, fishing equipment and skills are not easily substituted into other fisheries. Moreover, there is virtually no variation in observable characteristics because vessels and gear are nearly uniform across divers. In this fishery, harvesters make day trips from each of four northern California ports to locations offshore in waters up to 60 feet deep. Thus, choice occasions are easy to define, avoiding complications of multi-day trips that might occur in other fisheries. Connected through a hookah to an air compressor on the vessel, harvesters dive for the urchins, scrape them from the bottom using hand-held rakes, collect them in mesh bags, and then deliver the urchins to processing facilities at the port of landing. Urchins are processed immediately, packaged, and shipped to the Tokyo Central Wholesale Market for sale in the fresh market. In earlier work, the data were divided into eleven geographically distinct harvest zones or patches [11,12,13,14]. The patches are not of equal size, but instead they reflect spatial breaks in harvest activity that suggest natural divisions between patches. With the exception of patch 0 (the Farallon Islands), all patches are contiguous along the northern California coast, beginning in Half Moon Bay and stretching north to the Oregon border. Thus, the relevant spatial choices can be thought of as occurring in one dimension rather than in two.

To reduce the state space to a manageable size, we focus on divers fishing from only the Fort Bragg port and divide the locations choices into four total patches. Three of these

patches are patches from earlier work of Smith and Wilen, specifically the ones closest to Fort Bragg. The fourth patch represents all other spatial locations so that the choice set completely partitions the set of possible actions. To create a manageable sized data set, we first selected only records prior to 1991, which gives us 3 years of data. This cuts down the size of the data set and focuses on the period in which we see a lot more mobility. For fleet-wide totals, we include all observations in this time period. For individuals, we choose only individuals with at least 20 dives over the 3-year period. This leaves us with 69 individual divers. A unique new diver index (from 1 to 69) combined with a unique open season day index identifies individual-specific choice occasions. For each diver, we identify their first fishing observation as their first choice occasion. We also exclude choice occasions after their last observation in the full data set. For most individuals, this occurs after 1990, so no choice occasions are dropped.

Open season days are restricted by partial season closures. In 1989, the season was shut down for nine months. In 1990, fishing was not permitted in the month of July. In later years, partial season closures were expanded significantly, but we focus on just 1988-90. There are a total of #### observations. Of these, 3127 are choice occasions on which a trip occurred. That averages out to about 45 trips per person, but there is substantial heterogeneity in participation rates.

The main focus of our paper is on fishing search behavior and how fishers update spatially explicit information on returns. In our model we treat gross fishing revenue as stochastic but not independent across time and individuals. Stochastic revenues are driven by abundance uncertainty and price uncertainty, the latter of which may be spatially explicit as a result of quality differences over space. The other feature of spatially explicit returns is the cost of travel. For patches 1-3, the corresponding distances are: 0.18, 0.00, and 0.22 degrees of

latitude (there are 60 nautical miles per degree). For patch 4, since this location represents all other fishing locations, we compute a trip weighted average travel distance, which is 0.46 degrees of latitude. These distances (DISTANCE) do not vary over time, so the time index in (1.10) can be dropped from $x_{jt}$. Travel costs are simply a fixed marginal cost of travel multiplied by distance. While theoretically these costs could vary across individuals and time, we assume that they do not. Thus, we can estimate a single parameter on distance to patch that represents marginal cost of travel.

The factors that influence the decision to not fish are somewhat different in nature. In Section II.B, we assumed that utility from not fishing is simply 0 as a normalization. However, this utility varies across time due to market institutions, changing weather conditions, and changing economic opportunities outside of fishing. Following Smith [14] and Smith and Wilen [13], we use a dummy variable for Friday, Saturday, or Sunday (DWEEKEND) to capture the downturn in fishing activity as a result processor closures on Saturdays and Sundays as well as the Tokyo Central Wholesale Market closures on Sundays. Also, we use the same weather variables: 12-hour daily averages of wave height (WH), wind speed (WS), and wave period (WP) measured from a NOAA buoy off the coast of Point Arena. All of these factors increase the physical risk of sea urchin diving and, ceteris paribus, increases in them ought to decrease the probability of fishing. Unlike previous work with these data, introducing a dynamic framework requires an additional assumption about responsiveness to weather. Since we are testing whether urchin divers are forward-looking about their fishing decisions, we require that they have some forward-looking assessment about the weather. As a first step, we assume that they have perfect foresight about the 12-hour averages on wave height, wind speed, and wave period.

With the choice of not fishing, there are a total of five possible choices on each open season day (j=1, ... 5). We denote choice-specific constants as $\alpha_j$. Given the independent variables that we identify above and the model of revenue updating, following (1.10) the empirical specification for expected utility is as follows:

$$U_{ijt} = \begin{cases} \alpha_j + \beta_1 * DISTANCE_j + \bar{r}_{ijt} + \varepsilon_{ijt}, & for\ j = 1,...,4 \\ \alpha_j + \beta_2 * DWEEK + \beta_3 * WH_t + \beta_4 * WS_t + \beta_5 * WP_t + \varepsilon_{ijt}, & for\ j = 5, \end{cases} \tag{1.18}$$

where i indexes individuals. For identification, we restrict one of the $\alpha_j$'s to zero.

## IV. Estimation Strategy

Estimation of the likelihood function (1.17) involves a nested inner algorithm in which a stochastic dynamic programming (SDP) algorithm approximates the value function $v_t$ for each day of the season, and an outer gradient algorithm that searches for the values of $\{\boldsymbol{\beta}, \boldsymbol{\sigma}^2, \mu, \delta\}$ that maximizes the likelihood function. With $v_t$ known for each day of the season, the gradient algorithm is routine. Nonetheless, because an SDP must be solved at *each iteration* of the gradient algorithm, a premium is placed on parsimony in both the specification of the model –in particular, the number of parameters in the utility function –and the approximation of the value function.

Four factors effectively determine the size of the estimation problem: the number of parameters to be estimated, the number of fishers in the sample, the number of time periods in the SDP problem, and the number of state variables in the SDP problem. The first three of these have an essentially linear effect on the size of the problem. The fourth has an exponential effect, and thus deserves especially close scrutiny in the development of the estimation algorithm.

The value function depends on 26 state variables: the five exogenous state variables directly entering the utility function; the binary state concerning whether the fishery is closed on day *t*; the fleet-wide standard deviation of revenues associated with patch *j* on day t (four state variables, one for each patch, approximated from the complete data set for the fishery); the fleet-wide average revenues associated with patch *j* on day *t* (four state variables, one for each patch, approximated from the complete data set for the fishery); the random component of utility (four state variables, one for each patch); the number of trips taken by the boat to patch *j* (four state variables, one for each patch); and total revenues generated by the boat in patch *j* (four state variables, one for each patch). Fortunately, the first fourteen state variables are not affected by a fisher's trip decision, and therefore can be treated as conditioning variables that do not contribute to the dimensionality problem associated with SDP problems (though see below for a discussion of a negative implication of this treatment of these variables), and the state variables concerning the random component of utility, by virtue of their time independence and assumptions on their distributional form, are addressed in the estimation with relatively little computational effort. This still leaves eight state variables (which are necessary to calculate average revenues for each patch) contributing to a formidable computational challenge. Four of these state variables –those concerning the number of trips to a patch –are integer-valued, and so, insofar as the maximum number of trips to any patch in the sample is relatively small, the value function in the dimension of these state variables can be calculated exactly. The other four state variables are continuous, and in the dimension of these state variables the value function must be approximated.

For the case at hand Chebychev projection methods are the preferred approach to approximating the value function in the dimensions of the four continuous state variables (the state variables concerning the fisher's total revenue in each of the four patches). The advantages

of using Chebyshev polynomials to approximate functions are well-known. Not only do Chebyshev polynomials satisfy certain minimax theorems of approximation (theorems concerned with whether a polynomial minimizes the maximum approximation error), but coefficients of the polynomials are obtained by exceptionally rapid algorithms.

The SDP problem is solved via backward recursion, and with regard to the recursion, two observations with implications for the estimation strategy deserve comment. The first is that at each stage of the recursion, and for each patch $j$, the expectation of the value function in the previous stage (that is, the expectation of the value function at time $t+1$) must be taken (see equation (1.15)). This requires numerical quadrature over the distribution of revenues, which increases the size of the estimation problem by the number of quadrature points. Second convergence on an infinite-horizon solution is not possible because the value function for any day $t$ is implicitly conditioned by exogenous variables such as wave height ($WH$) and fleet-wide mean revenues for patch $j$. This means that the SDP recursion proceeds completely from the last day of the sample period to the first day, though on the many days when the season is closed the decision space is empty and so the algorithm is considerably smaller than indicated by the total number of days in the sample. Moreover, it is clear that the value function does converge to a fixed point as the number of trips to each patch increases; intuition suggests, for instance, that when the number of trips to all patches is, say, 1000, the value function is essentially a set of constants equal in number to the number of patches and independent of the trip decision at time t. This strongly suggests that the SDP problem at hand has convergent properties that can be exploited in estimation.

## References

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

2. 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), 1191-1197.

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

4. Erdem, T. and M.P. Keane (1996), "Decision-making Under Uncertainty: Capturing Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science* 15, 1-20.

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

6. Larson, D.M., W.R. Sutton, and J.M. Terry (1999), "Toward Behavioral Modeling of Alaska Groundfish Fisheries: A Discrete Choice Approach to Bering Sea/Aleutian Islands Trawl Fisheries," *Contemporary Economic Policy* 17:2, 267-277.

7. Mistiaen, Johan A. and Ivar E. Strand (2000), "Supply Response under Uncertainty with Heterogeneous Risk Preferences: Location Choice in Longline Fishing," *American Journal of Agricultural Economics* 82(5), 1184-1190.

8. Provencher, Bill and Richard C. Bishop (1997), "An Estimable Dynamic Model of Recreation Behavior with an Application to Great Lakes Angling*," Journal of Environmental Economics and Management* 33, 107-127.

9. Sanchirico, James N. and James E. Wilen (1999), "Bioeconomics of Spatial Exploitation in a Patchy Environment," *Journal of Environmental Economics and Management* 37, 129-150.

10. Smith, Martin D. (2000), "Spatial Search and Fishing Location Choice: Methodological Challenges of Empirical Modeling," *American Journal of Agricultural Economics* 82(5), 1198-1206 .

11. Smith, M.D. (2003), "State Dependence and Heterogeneity in Fishing Location Choice," Selected Paper, North American Association of Fisheries Economists, Williamsburg, VA.

12. Smith, M.D. and J.E. Wilen (2002), "State Dependence in Modeling the Spatial Behavior of Renewable Resource Users," forthcoming, *Proceedings of the11th Biennial Conference of the International Institute for Fisheries Economics and Trade*.

13. Smith, M.D. and J.E. Wilen (2003), "Economic Impacts of Marine Reserves: The Importance of Spatial Behavior," forthcoming, *Journal of Environmental Economics and Management*.

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