On the Frontier of Generating Revealed Preference Choice Sets:
An Efficient Approach

David Scrogin
Department of Economics
University of Central Florida
Orlando, FL 32816-1400
dscrogin@bus.ucf.edu

Richard Hofler
Department of Economics
University of Central Florida
Orlando, FL 32816-1400
rhofler@bus.ucf.edu

Kevin Boyle
Department of Resource Economics and Policy
University of Maine
Orono, ME 04469-5782
kboyle@maine.edu

J. Walter Milon
Department of Economics
University of Central Florida
Orlando, FL 32816-1400
wmilon@bus.ucf.edu

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Denver, Colorado, August 1-4, 2004

On the Frontier of Generating Revealed Preference Choice Sets: An Efficient Approach

Abstract

Deterministic rules for generating choice sets are often employed by analysts confronting universal sets with large numbers of alternatives. For destination choice analysis, site exclusion rules defined by travel time, distance, or quality have a behavioral appeal, yet are fundamentally limited by their one-dimension scope. To remedy this shortcoming while maintaining the concept that trips require costly inputs to yield utility generating outputs, we develop and test an exclusion rule for generating choice sets defined by efficiency measures derived from stochastic frontier econometric models. Choice set composition, site choice efficiency and probability of selection, and consumer surplus may be compared with results obtained under alternative exclusion rules.

Keywords: Choice sets; Destination choice; Discrete choice models; Exclusion rules; Stochastic frontier models

On the Frontier of Generating Revealed Preference Choice Sets:
An Efficient Approach

1. Introduction

Proper selection of individual choice (or consideration) sets is an ongoing concern of choice analysts. While factorial design algorithms and efficiency statistics have been developed to assist in the design of stated choice experiments (e.g., conjoint analysis and multi-attribute contingent valuation), such tools elude the analyst of revealed preference (RP) data. Instead, the individual's true, underlying choice set is not known with certainty; as a result, defining RP choice sets appropriately is often problematic. The task is confounded when the universal set of alternatives—such as recreational sites, houses, and shopping locations—is unusually large.

In confronting the choice set issue in the RP context, recent studies of destination choice [7,13,23] have evaluated criteria defined by trip inputs (distance and time) and outputs (site quality) for excluding sites from universal sets comprised of hundreds of alternatives. Similar to McFadden's [11] random draws approach, these deterministic rules avoid the computational burden and memory requirements required to estimate random utility models (RUMs) with the full set of alternatives. Furthermore, the rules possess a behavioral appeal: sites that are sufficiently distant from the point of trip origin or deficient in particular aspects may, one might argue, reasonably be discarded. The finding shared by the studies is that RUMs and consumer surplus tend to be robust to the exclusion rule because the estimation results were similar to those obtained with the universal set of alternatives. The conclusion has followed that sites failing to satisfy the exclusion rule may be excluded without cause for concern.

Yet closer examination reveals the shortcomings of deterministic exclusion rules. In particular, because sites that fail to satisfy a rule defined by trip inputs (or output) are excluded regardless of the size of the respective output (or inputs), the rule fails to accommodate tradeoffs between trip inputs and attainable site quality. Thus, high-quality sites may be discarded if a conservative input rule is adopted, while sites within close spatial or temporal proximity to the point of trip origin may be discarded if a conservative output rule is adopted. It follows that deterministic rules are not consistent with utility maximization if they discard sites from the universal set that dominate sites that are retained in the reduced (or censored) choice sets.

In light of these issues, we test an exclusion rule that accommodates tradeoffs between trip inputs and outputs and which is defined in terms of productive efficiency. Stochastic frontier models are estimated to identify each individual's frontier of efficient sites; choice sets are then generated by excluding sites from the universal set that deviate sufficiently from the frontier. Typically, frontier modeling is used for supply-side analyses, yet the methods readily extend to the study of consumer behavior as demonstrated here. Our approach permits direct comparisons to deterministic input and output rules and contributes to the literature on econometric approaches for choice set generation (e.g., [5,20,22]). Coupled with its ease of implementation in the discrete choice framework, stochastic frontier modeling may serve as a useful tool in future choice analyses.

Additional background is provided in the next section. Using a short-run production framework relating trip inputs to attainable site quality, the frontier model is developed for choice set generation in section three. Choice set composition, site choice efficiency, and estimated RUMs obtained with the efficient exclusion rule are contrasted with results produced with deterministic choice sets in the case of angler site choice. The implications of the analyst's choice of exclusion rule for measuring non-market use values are explored in the fourth section. Section five concludes.

2. Trip inputs, destination quality, and the choice set definition

We begin in the conventional way by assuming individuals derive utility from destination quality and a numeraire good but are constrained by income (see e.g., [4,6]). Quality is obtained by traveling a known distance or length of time to any of *n* heterogenous destinations in the universal choice set *J*. Assuming utility maximization as the objective, an individual *i*'s problem is written:

$$\max_{j \in J} V(q_{j,i}; y_i - p_{j,i}) \qquad i \in I \tag{1}$$

where $V(.)$ is the indirect utility function, $q_{j,i}$ denotes the quality of site *j* attainable by the individual, y_i is income, and $p_{j,i}$ is the implicit price of site access.[1]

Deterministic exclusion rules defined by trip inputs and outputs can be motivated directly from (1). With the former, since $p_{j,i}$ is a function of round-trip distance and travel time, defining a rule solely by trip inputs is equivalent to defining the rule in terms of travel costs. If site *j* is sufficiently distant from the point of trip origin—and, hence, costly to access—then, all else constant, the utility derived from the site may be sufficiently low for the site to be excluded from the universal set if it fails to satisfy the input criterion. Similarly, with exclusion rules defined in terms of the trip outputs, if the quality of site *j* is sufficiently poor then, all else constant, the utility derived from visiting the site will be sufficiently low, and again, the site may be excluded from *J*.

Recent studies have examined the sensitivity of destination choice models to deterministic exclusion rules. [13] evaluated a rule defined by travel time in modeling angler destination choice in Maine. Their findings indicated that estimated RUMs and non-market benefits were robust as sites were excluded from the universal set in thirty minute increments from the point of trip origin.

[23] evaluated both input and output exclusion rules in studying angler site choice in the southeast United States. The trip input was defined by distance, and two definitions of expected site quality were evaluated: the historic (five-year) daily catch rate of fish and an individual-specific daily catch rate generated from an estimated Poisson regression model (see e.g., [10,12]). With sites excluded in sixty mile increments, the findings of [13] were supported. Similarly, for both definitions of site quality the estimated RUM parameters quickly converged upon those obtained with the universal set of sites. [7] evaluated the travel time and distance rules relative to 'familiar' choice sets constructed with the aid of individual responses to survey questions for studying Maryland beach access. Again, the estimates quickly converged upon those produced with the universal set of sites, which may be attributed to a high degree of correlation between travel time and distance.

In deciding upon a deterministic rule for choice set generation, it is natural to question whether a tradeoff exists between trip inputs and attainable site quality. In many instances, quality can reasonably be expected to increase with the time or distance traveled, at least over some range.[2] If empirical findings support this hypothesis, then concern arises about the merits of deterministic rules. Specifically, if the rules are defined solely by the spatial or temporal proximity of sites to the trip departure points, then high-quality sites may be discarded. Alternatively, if the rules are defined solely by site quality, then convenient and readily accessible sites may be discarded. As a result, deterministic rules may be inconsistent with utility maximization if sites are retained from the universal set that are dominated by sites that are excluded. The implications of this censoring for empirical analysis are evaluated below.

3. Frontier modeling and individual choice

To remedy the limitations of deterministic exclusion rules within a utility maximization framework, we test an approach for choice set generation that accommodates tradeoffs between trip inputs and outputs and which is defined in terms of productive efficiency. Stochastic frontier models (see e.g., [1,9]) are employed to quantify relationships between trip inputs (travel time or distance) and attainable quality.[3] By measuring a site's divergence from the individual's frontier of efficient sites—referred to here as *site efficiency*—choice sets are generated by excluding sites that fail to satisfy a specific efficiency criterion. And similar to [7,13,23], the criterion may be varied to evaluate the effects on individual choice sets, probability of site choice, and non-market benefits.

Trip inputs and site quality are assumed to be continuous, non-negative and positively related.[4] In the absence of site *in*efficiency, the production function relating the quality of site j attainable by individual i ($q_{j,i}$) to the quantity of the trip input required to access the site ($x_{j,i}$) is defined as:

$$q_{j,i} = f(x_{j,i};\beta) \tag{2}$$

where β is a positive parameter. Attaching a per-unit input price to $x_{j,i}$ (e.g., cents-per-mile or opportunity costs of time) yields the implicit price $p_{j,i}$ appearing in (1). Empirically, (2) will likely fail to hold perfectly in most, if not all, situations. In addition to stochastic variation in the input-output relationship, quality will likely differ between sites that are equidistant from the trip departure point (or approximately so). The site that provides the maximum quality for a given amount of the trip input is considered to be the most efficient (or dominant) site.[5] To accommodate divergence from the efficient frontier, an efficiency term $\varepsilon_{j,i}$ defined over the interval (0,1] is included in (2):

$$q_{j,i} = f(x_{j,i};\beta)\cdot\varepsilon_{j,i} \tag{3}$$

Site j lies on the frontier if $\varepsilon_{j,i}$ = 1, while the site is considered increasingly inefficient as $\varepsilon_{j,i} \rightarrow 0$.

To accommodate stochastic variation, a two-sided random error term denoted by $exp(\phi_{j,i})$ is included in (3). The resulting stochastic frontier function is given by:

$$q_{j,i} = f(x_{j,i};\beta)\cdot\varepsilon_{j,i}\cdot exp(\phi_{j,i}) \tag{4}$$

Taking natural logarithms, (4) is rewritten:

$$ln(q_{j,i}) = ln\Big(f(x_{j,i};\beta)\Big) - u_{j,i} + \phi_{j,i} \tag{5}$$

where $u_{j,i} = -ln(\varepsilon_{j,i})$. The stochastic frontier model in (5) may be estimated by maximum likelihood using specific distributional assumptions about the transformed one-sided efficiency term u and the two-sided random term ϕ. Typically, ϕ is assumed to be distributed as $N(0,\sigma_\phi)$, while u may be assigned one of several distributions (e.g., truncated normal, half normal, and exponential).

Using ε, exclusion criteria may be defined in terms of site divergence from an individual's efficient frontier. The approach taken here constructs a lower bound from the frontier that is defined by a linear combination of the mean and standard deviation of the efficiency distribution (μ_ε and σ_ε , respectively). The n^* sites that lie between the frontier and the lower bound are retained from the n sites contained in the universal set J to form the individual choice set J^*. The sites that lie below the lower bound are excluded. Formally, the efficient site exclusion rule is defined as:

$$\text{For individual } i \in I, \text{ site } j \notin J^* \text{ if: } \quad q_{j,i} < f(x_{j,i}; \beta) \cdot \varepsilon_{j,i}^{-1} \cdot exp(v_{j,i}) \cdot (\mu_\varepsilon + \alpha\sigma_\varepsilon) \qquad (6)$$

The calibrating parameter α controls the location of the lower bound, and hence the size and composition of an individual's choice set. If $\alpha = 0$, the frontier is scaled downward by the mean of the efficiency distribution to form the lower bound. In general, $J^* \rightarrow J$ as $\alpha \rightarrow -\infty$, while the lower bound collapses upon the frontier from below as $\alpha \rightarrow \infty$.

To assess the size and composition of choice sets generated from a given exclusion rule (efficient or deterministic) and the relative efficiency of individual choice conditional upon the choice set, efficiency rankings [8] may be constructed by placing the n^* sites contained in J^* in descending order by their efficiency values and then assigning consecutive integers (1, 2, 3, ..., n^*) to the sites. The ratio of the rank of the chosen site to n^* provides a metric for evaluating site choice efficiency. A ratio of one indicates that the individual chose the least efficient site; a ratio approximating zero indicates that a relatively efficient site was chosen. A second metric is constructed with t-scores calculated from the efficiency values.[6] The ratio of a choice set's mean efficiency t-score to the efficiency t-score of the chosen site measures the relative efficiency of site choice. A ratio less than one indicates the chosen site was relatively efficient; a ratio greater than one indicates the site was relatively inefficient.

3.1. Model specifications

Following the usual convention (see e.g., [1,9]), the deterministic portion of the stochastic frontier model is specified as double-log: $ln(q_{j,i}) = \beta_0 + \beta_1 ln(x_{j,i})$. The random error term ϕ and the log-transformed efficiency term *u* are assumed to be distributed as normal and truncated normal, respectively.[7] The resulting log-likelihood function is:

$$ln\, L = \sum_{j=1}^{J} \left\{ -\frac{1}{2} ln(2\pi) - ln\,\sigma_S - ln\Phi\left(\frac{\mu}{\sigma_S\sqrt{\gamma}}\right) + ln\Phi\left(\frac{(1-\gamma)\mu - \gamma\kappa_{j,i}}{\left\{\sigma_S^2\gamma(1-\gamma)\right\}^{1/2}}\right) - \frac{1}{2}\left(\frac{\kappa_{j,i}+\mu}{\sigma_X}\right)^2 \right\} \quad (7)$$

where $\sigma_S = \left(\sigma_u^2 + \sigma_\phi^2\right)^{1/2}$, $\gamma = \sigma_u^2 / \sigma_S^2$, $\kappa_{j,i} = ln(q_{j,i}) - \beta_1 ln(x_{ji}) - \beta_0$, and $\Phi(.)$ is the standard normal cumulative distribution function. The individual-specific efficiency measure for site *j* is:

$$\varepsilon_{j,i} = \left(\frac{1-\Phi(\sigma^* - \mu_{j,i}^* / \sigma^*)}{1-\Phi(-\mu_{j,i}^* / \sigma^*)}\right) exp\left(-\mu_{ji}^* + \frac{1}{2}\sigma^{*2}\right) \quad (8)$$

where the terms μ_{ji}^* and σ^* are defined:

$$\mu_{ji}^* = \frac{-\kappa_j\sigma_u^2 + \mu\sigma_\phi^2}{\sigma_S^2} \quad \text{and} \quad \sigma^* = \frac{\sigma_u\sigma_\phi}{\sigma_S} \quad (9)$$

And the empirical expression of the site exclusion rule (6) is:

For individual $i \in I$, site $j \notin J^*$ if: $q_{ji} < x_{j,i}^{\beta} \cdot \varepsilon_{ji}^{-1} \cdot exp(\phi_{j,i}) \cdot (\mu_\tau + \alpha\sigma_\tau)$ (10)

For modeling site choice, the utility derived by individual i from visiting site j is written:

$$V_{j,i} = v(q_{j,i};y_i - p_{j,i}) + \gamma_{j,i} \quad \forall j \in J^*; i \in I \tag{11}$$

where $v(.)$ is the deterministic utility component depicted initially in (1), and $\gamma_{j,i}$ is the random utility component. From (11), the probability of choosing site j from the set J^* is:

$$\begin{aligned} prob(choose\, j) = \; & prob[v(q_{j,i};y_i - p_{j,i}) + \gamma_{j,i} \; > \\ & v(q_{f,i};y_i - p_{f,i}) + \gamma_{f,i}] \quad \forall j,f \in J^*;\; i \in I \end{aligned} \tag{12}$$

Distributional assumptions about γ lead to specific RUMs. As with the frontier models, various RUMs were investigated.[8] McFadden's conditional logit model, which results from assuming the γ are iid type-I extreme value, was selected. The conditional logit choice probability is:

$$prob(choose\, j) = \frac{e^{v_{j,i}}}{\sum_{j=1}^{J^*} e^{v_{j,i}}} \quad \forall j \in J^*; i \in I \tag{13}$$

And the conditional logit log-likelihood function is:

$$ln(L) = \sum_{i=1}^{I} \sum_{j=1}^{J^*} z_{j,i}(v_{j,i} - ln\sum_{j=1}^{J^*} exp(v_{j,i})) \tag{14}$$

where $z_{j,i}$ is equal to one if individual $i \in I$ selects site $j \in J^*$ and is zero otherwise. To evaluate the effects of the efficiency criterion (10) on the model, choice sets are generated by varying the calibrating parameter α through its range of values and then estimating the model at each value.

3.2. The data

The empirical application is individual choice of inland fishing destination in Maine during the 1999-2000 season. The data were obtained through mail surveys using the method of Salant and Dillman [17]. The surveys elicited the sites visited (lakes and ponds), the quantities of ten species of fish caught and harvested, and respondent socio-economic characteristics. The sample employed here consists of 1,081 Maine residents who took 5,556 single-day trips to 483 sites.[9]

For estimation of the stochastic frontier models, the trip input is defined as the round-trip travel time between the centroid of the respondent's home zip-code and the site access points. One-way travel time averages approximately 1.25 hours and ranges from about zero to five hours (Table 1). Converted to natural logarithms, the input variable is referenced by *Ln(Time)*. The trip output assumed to be shared by the anglers is the daily catch rate of fish. We define the catch rate in terms of the four species of fish that comprised the bulk of reported catch: salmon, brook trout, lake trout, and brown trout. This measure of expected site quality is generated by estimating species-level count data models relating catch to characteristics of the sites and the respondents and then summing the individual-specific fitted values over the four species of fish (see [10,18]).[10] Converted to natural logarithms, the daily expected catch rate is referenced by *Ln(Quality)*.

For estimation of the site choice models, the indirect utility function *v(.)* is specified as linear-in-parameters and linear-in-variables. Included in *v(.)* are the species-specific expected catch rate variables (*Salmon*, *Brook Trout*, *Lake Trout* and *Brown Trout*) and a travel cost price proxy (*Travel Cost*) defined as the sum of explicit round-trip travel costs ($0.33/mile) and the opportunity cost of time valued at one-third of the estimated hourly wage rate.

3.3. *Estimation results*

3.3.1. *Stochastic frontier models*

The stochastic frontier models were estimated separately for each individual in the sample over the 483 sites defining the universal choice set. Results reported in Table 2 indicate that expected site quality increases significantly with travel time on average. As the *Ln(Time)* coefficients may be interpreted as elasticities, the results indicate that a one hundred percent increase in one-way travel time (about 1.25 hours) yields about a fifteen percent increase in expected daily catch on average. More than seventy-five percent of 5,556 *Ln(Time)* coefficients are positive and significant, while only two percent are negative and significant. We retain the former portion of the sample for modeling site choice. Table 2 also presents summary statistics on the estimated one-sided errors ($\hat{u}$'s) and the associated efficiency estimates ($\hat{\varepsilon}$'s). Overall, the average site is relatively efficient ($\bar{\hat{\varepsilon}}$ = 0.87), but across the 483 sites and 1,081 individuals the efficiency estimates vary sizably as gauged by the average standard deviation ($\bar{\hat{\sigma}}_{\varepsilon}$ = 0.25) and range (0.06 – 0.99) their values.

3.3.2. *Analysis of the choice sets*

Using the individual efficiency estimates for the 483 sites, choice sets are generated with the efficient exclusion rule (10) at different values of the calibrating parameter α. For comparison, travel-time dependent choice sets are also generated [7,13]. The left half of Table 3 reports summary results on the composition of the choice sets generated with the efficient rule and the right half reports results for choice sets generated with the deterministic rule. The universal choice set (*J*) is obtained with the efficient rule by relaxing the calibrating parameter α to a value of -3.0 standard deviations below the mean site efficiency. Increasing the parameter to α = 2.0 leads to the minimum feasible choice set, containing n^* = 2 sites. Evaluated at $\bar{\hat{\varepsilon}}$, the average choice set contains approximately $\bar{J}^*$ = 231 sites ($\hat{\sigma}_{J^*}$ = 17.1). With the deterministic rule, J is obtained by including all sites within five hours of one-way travel, and the minimum choice set is obtained by excluding all sites beyond eighteen minutes. Evaluated at the mean (1.25 hours), the average choice set contains about twenty percent more sites ($\bar{J}^*$ = 277) than that of the efficient rule; however, the spread in choice set size around the mean is more than four times as large ($\hat{\sigma}_{J^*}$ = 79.0).

The composition of the choice sets is evaluated by converting the site efficiency estimates into t-scores and averaging the resulting values at each level of the exclusion rules. The mean t-score with the efficient rule increases from approximately zero at the universal set to beyond two as α is increased (Table 3). In contrast, the choice set composition with deterministic rules remains approximately constant on average with the deterministic rule as gauged by the efficiency t-scores. Here, the average t-score reaches a maximum value of only 0.13 because both relatively efficient and inefficient sites are discarded as the permissible level of travel time decreases. This is noteworthy because it provides an explanation why RUM coefficients estimated with deterministic choice sets have exhibited little variation relative to those obtained with the underlying universal choice sets.

3.3.3. Analysis of site choice

Conditional upon the choice sets, the efficiency of site choice is measured and the probability of site choice modeled. Beginning with the 483 site choice set (J), Table 4 shows that for both exclusion rules the absolute and relative site choice rankings averaged 163 and 0.34, respectively, and the efficiency t-score ratio averaged approximately 0.0. Considering the range of the rankings, in all cases some individuals selected the most efficient sites in their choice sets (Min = 1), while others selected highly inefficient sites (Max $\cong n^*$). Comparing the efficiency and deterministic rules, the relative efficiency of the chosen sites differs considerably as sites are excluded from J. With the efficient rule, the relative site ranking increases to approximately 1.0 on average at α = 2.0. Similarly, the ratio of the mean efficiency t-score relative to the t-score of the chosen site rises from 0.0 to beyond 3.0 as α increases to 2.0. These results indicate that as the lower bound approaches the efficient frontier, the chosen site becomes relatively inefficient within J^*. In contrast, with the deterministic rule the relative site rankings are nearly constant and the mean efficiency t-score ratio rises only slightly from that observed with J on average. Hence, both the choice set composition and the relative location of the chosen site within the choice set are largely constant when evaluated in terms of the site efficiency estimates.

The estimation results from the site choice models are examined graphically in Figures 1 and 2 (see Appendices D and E for numerical results).[11] With the efficient exclusion rule, the estimated *Travel Cost* coefficients are negative as anticipated but increase consistently as α increases above -1.5 (Figure 1). This finding can be explained with the results of the efficiency analysis (Tables 3-4). Specifically, because the relative efficiency of a chosen site decreases as the lower bound approaches the frontier (i.e., as α is increased), it follows that the site lies below the frontier by an amount that is relatively greater than that of the collection of remaining sites. Hence, the estimated marginal effect of travel cost increases with the efficiency of the choice set, all else held constant. In contrast, the coefficient estimates obtained with the deterministic choice sets are virtually constant across values of travel time, which may be explained by the findings that the choice set composition and the location of the chosen sites within the choice sets tended to be invariant.

As shown in Figure 2, the estimated catch rate coefficients diverge downward in all cases as α increases. In fact, at relatively conservative levels of the rule (i.e., $\alpha > 0$), the coefficient estimates are negative for at least two of the four species (i.e., *Brook Trout* and *Brown Trout*). This seemingly counter-intuitive result has a straight-forward explanation in terms of site efficiency. In particular, as α increases, the chosen sites become increasingly less and less efficient within the choice sets. Thus, for a given amount of travel cost the quality of the chosen sites is relatively low. As a result, the expected catch coefficients become negative as α rises sufficiently. On the other hand, the expected catch coefficients are robust with the deterministic rule, which can be explained in terms of site efficiency in a manner similar to that of the *Travel Cost* coefficients.

To summarize, the behavior of the RUMs estimated with deterministic choice sets is largely consistent with [7,13,23], and the robustness of the estimation results may be explained with the site efficiency estimates obtained from the stochastic frontier models. In addition to the choice set composition remaining largely constant as sites are excluded from J (Table 3), the relative position of the chosen sites within the deterministic choice sets is approximately constant. In contrast, convergence of the parameter estimates upon those obtained with J occurs at a much slower rate with choice sets generated from the efficiency rule. A natural question is how stringently the efficiency rule should be applied. As behavioral models underlying destination choice analyses have generally assumed a positive and monotonic relationship between utility and site quality, the results of this study indicate that the calibrating parameter (α) should be set at a value no greater than zero in order for the estimation results to be consistent with utility maximization.

4. Implications for benefits measurement

A common application of RUMs is to estimate the non-market benefits associated with exogenous changes in some or all of the alternatives comprising individual choice sets. With public resources, such as the water bodies and adjoining lands considered here, changes in policy may affect site quality or accessibility. For example, to mitigate environmental degradation or stock depletion, state and federal agencies often impose regulations upon particular activities in which the public may engage—directly or indirectly affecting site quality—or instead restrict public access through site closures (see e.g., [18]). As the estimated welfare effects of a given policy change are a function of the choice set definition and econometric specification of the site choice models, the exclusion rule serves a crucial role in benefits measurement.

To evaluate these effects, we focus upon a controversial set of five large lakes contained in the "China Lakes" region of the Maine [14]. While the lakes are among the state's most coveted coldwater-fisheries, agricultural and forestry production coupled with residential developments have led to increased levels of eutrophication and toxic loading in their waters. In exploring the implications of the site exclusion rules for welfare analysis, the focus is upon closure of the five lakes for recreational fishing. Thirteen of the 483 site access points are affected by the closures.

Individual i's compensating variation (CV_i) for the loss of one or more sites is calculated as:

$$CV_i = \frac{1}{MU_y}\left[ln\sum_{j=1}^{n^*} e^{V_{ji}} - ln\sum_{j=1}^{n^{**}} e^{V_{ji}}\right] \tag{15}$$

where MU_y represents the marginal utility of income, and n^* and n^{**} denote, respectively, the number of sites in the choice set J^* before and after the site closures. Estimated values of CV_i are obtained by replacing MU_y by an estimated *Travel Cost* coefficient and *v(.)* by its fitted values (see e.g., [4,6]).

As with the choice set analysis, we begin with the universal set of sites (Table 5). Here, the thirteen sites are included in one hundred percent of the individual choice sets. On average, the mean welfare loss is calculated to be thirty-two cents per trip under both exclusion rules. With the efficient rule, the number of affected sites in the choice set declines to less than one on average as α is increased, but the mean welfare loss per trip rises sharply. Given that the percentage of individuals affected by the site closures falls as α increases, the results indicate that those who are affected experience large welfare losses.

Compared to the efficient rule, the results indicate that the number of affected sites contained in the choice sets was considerably larger on average with the deterministic rule, especially at relatively conservative levels. Further, there is a tendency for the deterministic rule to overstate the portion of individuals affected by the site closures relative to the efficient rule, with the results again being amplified as J^* decreases. Given the robustness of the RUM estimates produced with deterministic choice sets, the mean welfare loss is found to be invariant across all but the most conservative values of travel time. Further, the percentage of anglers affected by the site closures is notably larger. When considered at the individual level, the welfare losses produced under the efficient rule are larger than with the deterministic rule—and in some cases strikingly so. However, when the standard deviation of the estimates is taken into consideration, the results indicate that there is much more heterogeneity within the population when evaluated with choice sets generated from the efficient rule.

5. Conclusions

Efficient exclusion rules for generating choice sets with RP data have several advantages relative to deterministic rules. First, by overcoming the ‘dominance’ problem associated with deterministic rules, efficient rules are behaviorally more appealing and economically intuitive. Second, implementation of the stochastic frontiers approach in the discrete choice framework requires no additional data beyond that required for a defensible model of destination choice (i.e., travel costs and site attributes). While individual characteristics may additionally be incorporated into the choice set generation process, such data are not required as they typically are with probabilistic approaches for modeling choice sets (see e.g., [5,22]).

On a final note, because the production frontier model employed in this study has a dual—the stochastic cost frontier—the efficient exclusion rule may be extended for cost or price analyses. As an example, residential housing choice studies to date have used deterministic rules to construct individual choice sets (e.g., [2]). However, the number of homes for sale in residential housing markets during a given period may be of sufficient size for concern to arise about the use of deterministic rules in these situations. Given the shortcomings of deterministic rules and the popularity of discrete choice models for policy analysis, additional research efforts in both the stochastic production and cost contexts are warranted.

References

[1] D.J. Aigner, C.A.K. Lovell, and P. Schmidt, Formulation and estimation of stochastic frontier production function models, Journal of Econometrics 6 (1977) 21-37.

[2] H.S. Banzhaf, V.K. Smith, Meta analysis in model implementation: choice sets and the valuation of air quality improvements, Discussion Paper 03-61, Resources for the Future, 2003.

[3] A.C. Cameron, P.K. Trivedi, Regression Analysis of Count Data, Cambridge, Cambridge University Press, 1998.

[4] A.M. Freeman III, The measurement of environmental and resource values: theory and methods, Washington D.C., Resources for the Future, 2003.

[5] T.C. Haab, R. Hicks, Accounting for choice set endogeneity in random utility models for recreational demand, Journal of Environmental Economics and Management 34 (1997) 127-147.

[6] T.C. Haab, K.E. McConnell, Valuing environmental and natural resources, Cheltenham, UK, Edward Elgar Publishing Limited, 2002.

[7] R.L. Hicks, I.E. Strand, The extent of information: its relevance for random utility models, Land Economics 76 (2000) 374-385.

[8] U. Jensen, Is it efficient to analyse efficiency rankings? Empirical Economics 25 (2000) 89-208.

[9] S.C. Kumbhakar, C.A.K. Lovell, Stochastic Frontier Analysis, Cambridge, Cambridge University Press (2000).

[10] K.E. McConnell, I.E. Strand, L. Blake-Hedges, Random utility models of recreational fishing: catching fish using a Poisson process, Marine Resource Economics 10 (1995) 247-261.

[11] D. McFadden, Modelling the choice of residential location, in: A Karlqvist, L. Lundquist, F.

Snickars, J Weibull (Eds.), Spatial Interaction Theory and Planning Models, Amsterdam, North-Holland, 1978, pp. 75-96.

[12] E.R. Morey, D.M. Waldman, Measurement error in recreation demand models: the joint estimation of participation, site choice, and site characteristics, Journal of Environmental Economics and Management 35 (1998) 262-276.

[13] G. Parsons, B. Hauber, Spatial boundaries and choice set definition in a random utility model of recreation demand, Land Economics 74 (1998) 32-48.

[14] G.R. Parsons, A.J. Plantinga, K. J. Boyle, Narrow choice sets in a random utility model of recreation demand, Land Economics 76 (2000) 86-99.

[15] Paterson, R., D. Scrogin, K. Boyle, D. McNeish, Maine open water fishing survey: summer 1999, Maine Agricultural and Forest Experiment Station Paper No. 493, University of Maine, 2001.

[16] S. Rabe-Hasketh, A. Skrondal, A. Pickles, Estimation of generalized linear models, The Stata Journal 2 (2002) 1-21.

[17] P. Salant, D. Dillman, How to conduct your own survey, New York, John Wiley and Sons, Inc., 1994.

[18] D. Scrogin, K. Boyle, G. Parsons, A.J. Plantinga, Effects of regulations on expected catch, expected harvest, and site choice of recreational anglers, Forthcoming in American Journal of Agricultural Economics (2003).

[19] V.K. Smith, R.B. Palmquist, P. Jakus, Combining Farrell frontier and hedonic travel cost models for valuing estuarine quality, Review of Economics and Statistics 73 (1991) 694-699.

[20] J. Swait, M. Ben-Akiva, Empirical test of a constrained choice discrete model: mode choice in

Sao Paulo, Brazil, Transportation Research-B 21B (1987) 103-115.

[21] K. Train, Mixed logit models for recreation demand, in: C. Kling, J. Herriges (Eds.), Valuing the Environment using Recreation Demand Models, New York, Elgar, 1997, pp. 140-163.

[22] R.H. von Haefen, Latent consideration sets and continuous demand system models, working paper, Department of Agricultural and Resource Economics, University of Arizona (2003).

[23] J.C. Whitehead, T.C. Haab, Southeast marine recreational fishing statistical survey: distance and catch based choice sets, Marine Resource Economics 14 (1999):283-298.

Footnotes

1. The access price $p_{j,i}$ is usually assumed to consist of explicit and, perhaps, implicit travel costs (e.g., fuel expenditures and the opportunity cost of time, respectively). The former are a function of the round-trip distance between the trip departure point and the destination; the latter are usually expressed as a function of travel time and income.

2. For example, to the extent that public lands and waterways within city boundaries can be expected to be of lesser quality than those in the surrounding rural countryside, it follows that residents must travel a sufficient time or distance to experience the higher quality destinations. The exception is individuals who reside adjacent to the highest quality sites.

3. To our knowledge, frontier models have only been implemented in the leisure demand context by [19]. While their application also entails angling, they used data envelopment analysis (DEA) methods to estimate a deterministic (rather than stochastic) hedonic travel-cost frontier model.

4. While the hypothesis is maintained that trip inputs and attainable quality are positively related, this relationship is assumed to be correlational rather than causal.

5. Although quality is represented here by a single, continuous variable, it may be considered a multi-dimensional vector of site attributes.

6. An individual's n site efficiency t-scores are calculated by subtracting the mean of the

individual's n efficiency estimates from the site-specific estimates and then dividing the differences by the standard deviation of the efficiency estimates.

7. The double-log specification has the benchmark in the frontiers literature and permits convenient calculation of ε and its interpretation as an efficiency measure. Models were also estimated with u distributed as half-normal and exponential and the deterministic relation specified as semi-log, linear, and quadratic. Estimation results are reported in Appendix C.

8. Random parameters (or mixed) logit models (see e.g., [21]) were estimated by Gaussian-Hermite and adaptive quadrature methods using the gllamm routine in Stata version 8.2 [16]. While the adaptive method is computationally more efficient, run time with the smallest data set used in this study (N = 55,111) and four random parameters exceeded ninety hours. Nested logit models were also considered, yet these require the analyst to place a common and potentially arbitrary structure upon the individual decision process. Although the mixed and nested models relax the IIA assumption of conditional logit, their empirical implementation in this study is problematic. However, the general conclusions of the analysis are expected to be upheld by the alternative models.

9. By focusing on single-day trips, differences in trip outputs that may occur with multiple-day trips are avoided. Details on the survey instruments, sample, and responses are contained in [15].

10. Because an angler's catch of individual species–and, hence, aggregate catch–is unknown *a priori*, the catch variables appear as expected values. The expected catch models were specified as

zero-inflated negative binomial as it outperformed Poisson, negative binomial, and zero-inflated Poisson models in all cases (see [3]). Variable definitions, summary statistics, and estimation results are reported in Appendices A and B.

11. All but two of the fifty-five RUM coefficients are significant (0.01 level) with the efficient rule, while all fifty-five of the coefficients are significant with the deterministic rule.

Table 1. Variable definitions and summary statistics

	Variable Definition	Mean	Standard Deviation
Stochastic Frontier Models		*N* = 2,683,548	
Ln(Quality)	Natural log of expected catch per day	-1.02	1.04
Ln(Time)	Natural log of round-trip travel time	0.76	0.70
Random Utility Models		*N* = 2,056,131	
Salmon	Expected catch of salmon per day	0.15	0.25
Brook Trout	Expected catch of brook trout per day	0.30	0.49
Lake Trout	Expected catch of lake trout per day	0.06	0.12
Brown Trout	Expected catch of brown trout per day	0.07	0.12
Travel Cost	0.325 x (Round trip miles) + time cost	51.99	33.40

Notes: The aggregate and species-specific expected catch rates are obtained from estimated zero-inflated negative binomial models relating reported catch to site and individual characteristics. Variable definitions, summary statistics, and estimation results are reported in Appendices A and B. The time cost component of *Travel Cost* is defined as one-third of the hourly wage; hourly values of the wage are derived from respondent annual income and a 2,000 hour work year (40 hours per-week x 50 weeks per-year).

Table 2. Estimated stochastic frontier models averaged over trips

Number of Trips: 5,556 Number of Sites: 483	Truncated Normal Stochastic Frontier Models			
Average Coefficient Estimates				
Ln(Time)	0.146* (0.040)			
Intercept	-0.904* (0.208)			
Summary of *Ln(Time)* Coefficients				
% Positive and Significant	76.6			
% Negative and Significant	2.1			
% Insignificant	21.3			
Summary of Efficiency Estimates	Mean	Standard Deviation	Min	Max
Estimated One-Sided Term: $\hat{u}$	0.22	0.47	1.02e-07	2.84
Estimated Site Efficiency: $\hat{\varepsilon}$	0.87	0.25	0.06	0.99

Notes: The dependent variable is the natural logarithm of expected daily catch. The number of estimated models is equal to the number of trips (5,556), and the number of observations per model is equal to the number of sites (483). Reported estimates are obtained by averaging the individual coefficient estimates. Standard errors are reported in parentheses. * denotes significance at the 0.01 level. The reported percentages are based upon two-tailed *t*-tests of the null H_o: $\beta_1 = 0$ conducted at the 0.05 level. Approximately seventy percent of the intercepts were significant.

Table 3. Choice set characteristics: efficient vs. deterministic exclusion rules

	Efficient Site Exclusion Rule						Deterministic Site Exclusion Rule						
α	Mean Sites	Standard Deviation	Min	Max	Mean Efficiency t-score	N	Travel Time	Mean Sites	Standard Deviation	Min	Max	Mean Efficiency t-score	N
-3.0	483.0	0	483	483	0.0001	2,056,131	5.0	483	0	483	483	0.0001	2,056,131
-2.5	482.8	1.1	476	483	0.001	2,055,233	4.5	482.8	0.4	480	483	-0.00002	2,055,488
-2.0	482.0	2.1	475	483	0.005	2,051,944	4.0	481.0	3.1	470	483	-0.001	2,047,542
-1.5	469.6	12.7	436	483	0.05	1,999,106	3.5	476.7	5.7	468	483	-0.004	2,029,142
-1.0	385.0	13.8	336	435	0.33	1,639,319	3.0	467.4	17.3	388	483	-0.01	1,989,703
-0.5	300.8	18.6	232	337	0.63	1,280,306	2.5	447.5	42.0	260	483	-0.01	1,905,148
0.0	230.9	17.1	175	274	0.90	982,877	2.0	409.5	67.2	130	470	-0.01	1,743,404
0.5	165.6	13.4	127	209	1.15	704,921	1.5	328.7	77.5	59	445	-0.02	1,399,332
1.0	93.2	10.7	72	134	1.46	396,593	1.0	214.7	70.5	12	317	-0.03	913,855
1.5	38.0	9.0	17	73	1.80	161,727	0.5	73.4	33.9	4	146	0.06	312,449
2.0	12.9	6.1	2	33	2.05	55,111	0.3	38.1	20.6	3	85	0.13	161,982

Table 4. Relative efficiency of individual choices: efficient vs. deterministic exclusion rules

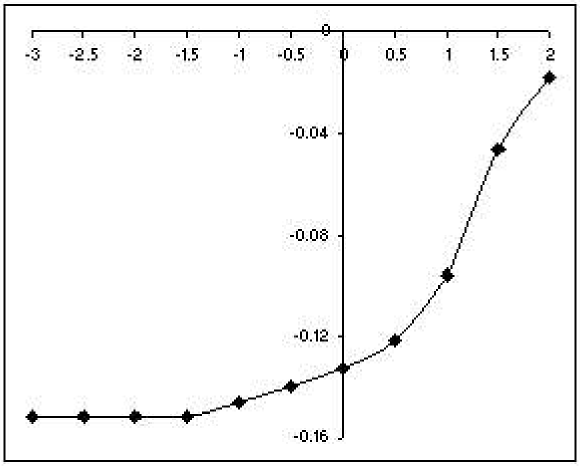
	Efficient Site Exclusion Rule						Deterministic Site Exclusion Rule				
	Efficiency Rankings			Relative	Relative		Efficiency Rankings			Relative	Relative
α	Mean	Min	Max	Rank of Chosen Site	Efficiency of Chosen Site	Travel Time	Mean	Min	Max	Rank of Chosen Site	Efficiency of Chosen Site
-3.0	163.1	1	480	0.34	0.0001	5.0	163.1	1	480	0.34	0.0001
-2.5	163.1	1	480	0.34	0.002	4.5	163.1	1	480	0.34	-0.00003
-2.0	163.1	1	479	0.34	0.01	4.0	162.4	1	477	0.34	-0.002
-1.5	163.1	1	478	0.35	0.08	3.5	160.7	1	475	0.34	-0.01
-1.0	161.2	1	421	0.42	0.56	3.0	157.0	1	471	0.34	-0.01
-0.5	151.1	1	331	0.50	1.08	2.5	150.1	1	463	0.34	-0.02
0.0	133.6	1	260	0.58	1.54	2.0	137.2	1	449	0.33	-0.02
0.5	107.0	1	205	0.65	1.97	1.5	109.5	1	402	0.33	-0.04
1.0	69.3	1	128	0.74	2.50	1.0	71.3	1	298	0.33	-0.04
1.5	33.3	1	73	0.86	3.09	0.5	25.9	1	123	0.35	0.10
2.0	12.5	1	33	0.97	3.51	0.3	14.3	1	62	0.38	0.21

Notes: The relative ranks are defined by ordering the sites by the efficiency estimates and then assigning consecutive integers (1, 2, 3, ..., n^*) to the sites. The relative efficiencies are constructed by converting the site efficiency estimates into t-scores and then calculating the mean t-score of the choice set to that of the chosen site.

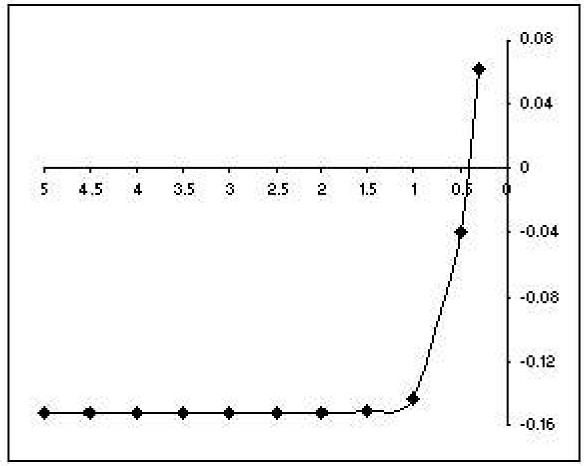
Table 5. Welfare losses for site closures: efficient vs. deterministic exclusion rules

Efficient Site Exclusion Rule					Deterministic Site Exclusion Rule				
			Welfare Losses					Welfare Losses	
α	Mean sites closed	Percent of anglers affected	Mean per trip	Standard deviation	Travel time	Mean sites closed	Percent of anglers affected	Mean per trip	Standard deviation
-3.0	13	100	$0.32	0.50	5.0	13	100	$0.32	0.50
-2.5	13	100	0.32	0.50	4.5	13	100	0.32	0.50
-2.0	13	100	0.32	0.50	4.0	13	100	0.32	0.50
-1.5	13	100	0.32	0.50	3.5	13	100	0.32	0.50
-1.0	12.6	100	0.36	0.57	3.0	13	100	0.32	0.50
-0.5	8.3	96.7	0.41	0.66	2.5	13	98.2	0.32	0.50
0.0	3.0	62.0	0.38	0.70	2.0	12.5	98.2	0.32	0.50
0.5	0.7	26.7	0.30	0.78	1.5	12.3	97.4	0.32	0.50
1.0	0.2	11.5	0.86	4.26	1.0	10.2	89.6	0.33	0.50
1.5	0.1	7.8	12.47	71.19	0.5	3.8	37.8	1.05	1.60
2.0	0.1	6.8	84.17	429.70	0.3	1.4	22.5	-0.55	2.26

Figure 1. Sensitivity of travel cost coefficients:
efficient vs. deterministic exclusion rules

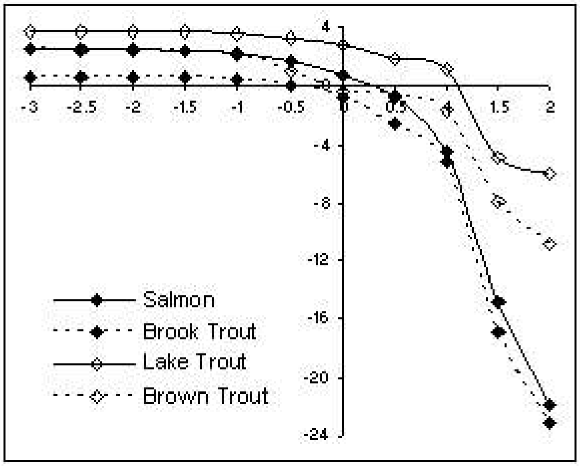


Efficient Choice Sets
Horizontal Axis: Calibrating Parameter (α)
Vertical Axis: Travel Cost Coefficient

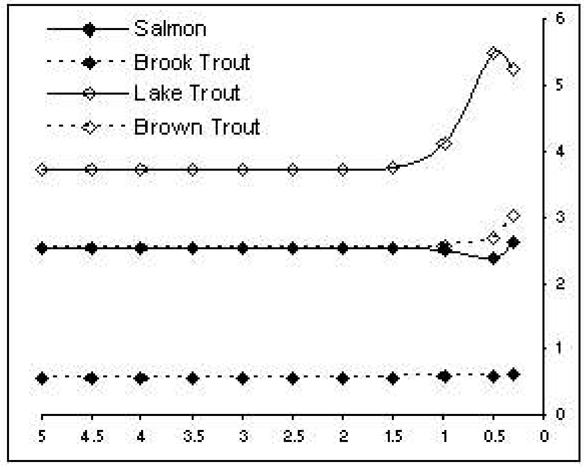


Deterministic Choice Sets
Horizontal Axis: One-Way Travel Time
Vertical Axis: Travel Cost Coefficient

**Figure 2. Sensitivity of expected catch coefficients:
efficient vs. deterministic exclusion rules**



Efficient Choice Sets
Horizontal Axis: Calibrating Parameter (α)
Vertical Axis: Expected Catch Coefficients



Deterministic Choice Sets
Horizontal Axis: One-Way Travel Time
Vertical Axis: Expected Catch Coefficients

Appendix A. Variable definitions and summary statistics

Variable Name	Definition		Mean	Standard Deviation
Angler Characteristics			I = 1,081	
Age	Angler age		40.58	12.91
Male	1 if angler is male; 0 if angler is female		0.88	0.32
Ln(Days)	Natural logarithm of single day trips		1.09	0.95
Targeted	1 if angler targeted the species; 0 otherwise:	Salmon	0.36	0.48
		Brook Trout	0.37	0.48
		Lake Trout	0.26	0.44
		Brown Trout	0.23	0.42
Fishery Characteristics			J = 483	
Species Not Present	1 if species not known to be present by managing agency; 0 otherwise:	Salmon	0.57	0.50
		Brook Trout	0.25	0.44
		Lake Trout	0.74	0.44
		Brown Trout	0.70	0.46
Species Not Abundant	1 if species not known to be abundant by managing agency; 0 otherwise:	Salmon	0.11	0.31
		Brook Trout	0.32	0.47
		Lake Trout	0.05	0.22
		Brown Trout	0.10	0.30

Appendix A continued

Stocked Species	1 if lake or pond stocked with species; 0 otherwise:	Salmon	0.26	0.44
		Brook Trout	0.32	0.47
		Lake Trout	0.08	0.28
		Brown Trout	0.20	0.40
Acres	Surface area of lake or pond in acres, scaled by 1,000		2.21	6.32
Depth	Depth of lake or pond in feet, scaled by 10		2.18	1.66
Elevation	Elevation of lake or pond above sea level, scaled by 100		5.45	4.52
Water Type 1	1 if oligotrophic water (low productivity; deep secchi disk readings); 0 otherwise		0.28	0.45
Water Type 2	1 if eutrophic water (high productivity; shallow secchi disk readings); 0 otherwise		0.34	0.47
No Live Bait	1 if no-live-bait regulation at lake or pond; 0 otherwise		0.13	0.34
Catch and Release	1 if catch-and-release regulation at lake or pond		0.01	0.12

Appendix B. Estimated zero-inflated negative binomial models of expected catch

Expected Catch	Salmon	Brook Trout	Lake Trout	Brown Trout
Modified Negative Binomial				
Log(Days)	0.80***	0.63***	0.79***	0.81***
	(0.07)	(0.07)	(0.09)	(0.09)
Targeted	1.92***	2.31***	2.31***	2.23***
	(0.16)	(0.16)	(0.21)	(0.20)
Age	0.02***	-0.02***	0.00	-0.01
	(0.00)	(0.01)	(0.01)	(0.01)
Male	-0.38	0.35	-0.54*	0.63*
	(0.23)	(0.23)	(0.32)	(0.33)
Acres	0.00	-0.01**	0.02**	0.00
	(0.00)	(0.00)	(0.01)	(0.01)
Depth	-0.04	0.00	-0.02	-0.14
	(0.07)	(0.03)	(0.08)	(0.11)
Elevation	0.04***	0.05***	0.00	0.05
	(0.02)	(0.02)	(0.03)	(0.03)
Water Type 1	0.34	-0.53***	0.59*	0.13
	(0.17)	(0.17)	(0.32)	(0.35)
Water Type 2	-0.21	-0.16	0.17	0.15
	(0.22)	(0.17)	(0.32)	(0.19)
No Live Bait	-0.39	-0.28	0.04	-1.12*
	(0.30)	(0.19)	(0.44)	(0.60)
Species Not Present	-0.20	-0.05	-0.47	-0.27
	(0.37)	(0.35)	(0.40)	(0.52)
Species Not Abundant	-0.57***	-0.99***	0.18	0.02
	(0.31)	(0.21)	(0.55)	(0.55)
Stocked Species	-0.36***	0.12	-0.19	0.40
	(0.20)	(0.15)	(0.24)	(0.46)
Catch and Release	0.04	-1.12**	-0.15	-0.37
	(0.46)	(0.48)	(0.52)	(0.76)
Constant	-2.37	-1.20***	-2.60***	-2.85***
	(0.44)	(0.38)	(0.59)	(0.63)

Appendix B Continued

Logit of Positive Catch				
Species Not Present	4.51***	3.12***	2.07***	1.73***
	(0.80)	(0.54)	(0.46)	(0.56)
Species Not Abundant	2.28***	1.64***	1.80**	2.06***
	(0.75)	(0.48)	(0.74)	(0.69)
Age	-0.03***	-0.05***	-0.02	-0.03**0
	(0.01)	(0.02)	(0.01)	(0.02)
Male	-1.97***	-0.35	-0.64	-0.17
	(0.75)	(0.46)	(0.49)	(0.61)
Dispersion Parameter				
Log(α)	0.36***	0.69***	0.47***	0.52*
	(0.12)	(0.13)	(0.22)	(0.27)
Prediction Statistics				
Mean	1.48	1.80	0.61	0.61
Standard Deviation	3.09	3.20	1.48	1.51
Minimum	0.00	0.01	0.01	0.00
Maximum	34.89	27.74	19.51	19.58
Sample Size	1,310	1,305	1,227	1,213
Log-Likelihood	-1,366.74	-2,022.56	-1,185.18	-1,141.60

Notes: The dependent variable is the reported catch per visited site during the open-water fishing season.

Appendix C. Comparison of estimated stochastic frontier models

Number of Trips: 5,556 Number of Sites: 483	Truncated Normal One-Sided Errors				Half-Normal One-Sided Errors				Exponential One-Sided Errors			
	Double-Log	Linear-Log	Linear	Quadratic	Double-Log	Linear-Log	Linear	Quadratic	Double-Log	Linear-Log	Linear	Quadratic
Coefficient Estimates												
Log(Travel Time)	0.147* (0.041)	0.086* (0.027)	---	---	0.146* (0.041)	0.086* (0.027)	---	---	0.147* (0.041)	0.086* (0.027)	---	---
Travel Time	---	---	0.043* (0.012)	---	---	---	0.043* (0.012)	---	---	---	0.043* (0.012)	---
Travel Time Squared	---	---	---	0.006* (0.002)	---	---	---	0.006* (0.002)	---	---	---	0.006* (0.002)
Intercept	-1.103* (0.217)	0.547* (0.059)	0.500* (0.063)	0.552* (0.113)	-1.062* (0.285)	0.540* (0.117)	0.493* (0.123)	0.552* (0.113)	-1.103* (0.217)	0.547* (0.059)	0.500* (0.063)	0.552* (0.113)
Coefficient Summary†												
% Positive and Significant	76.6	75.5	80.3	78.5	76.6	75.5	80.3	78.5	76.6	75.5	80.3	78.5
% Insignificant	21.3	21.5	18.3	20.9	21.3	21.5	18.3	20.9	21.3	21.5	18.3	20.9
% Negative and Significant	2.1	3.0	1.4	0.6	2.1	3.0	1.4	0.6	2.1	3.0	1.4	0.6

Notes: The dependent variable in the linear, quadratic, and semi-log models is defined as the sum of the salmon, brook trout, lake catch, and brown trout catch expectations generated from zero-inflated negative binomial (ZINB) models (see Appendix A). The dependent variable in the double-log models is the natural logarithm of the sum of the catch expectations. Coefficient estimates and standard errors (in parentheses) are obtained by averaging the coefficients over the 5,556 estimated stochastic frontier models. * denotes significance at the 0.01 level. † Significance based upon two-tailed *t*-tests (α = 0.05) of the null hypothesis H_o: β = 0. While not reported here, approximately 96 percent of the intercepts were significant (α = 0.05) with the levels and semi-log specifications; about 70 percent were significant with the double-log specifications.

Appendix D. Conditional logit estimation results: efficient choice sets

α	β_{Cost}	β_{Salmon}	$\beta_{Brook\,Trout}$	$\beta_{Lake\,Trout}$	$\beta_{Brown\,Trout}$	Pseudo R^2
-3.0	-0.152*	2.517*	0.555*	3.712*	2.527*	0.313
-2.5	-0.152*	2.517*	0.551*	3.711*	2.510*	0.313
-2.0	-0.152*	2.516*	0.546*	3.709*	2.482*	0.313
-1.5	-0.152*	2.512*	0.539*	3.704*	2.467*	0.310
-1.0	-0.146*	2.368*	0.474*	3.600*	2.197*	0.297
-0.5	-0.139*	1.784*	0.062	3.316*	0.955*	0.298
0.0	-0.132*	0.692*	-0.777*	2.861*	-0.410**	0.316
0.5	-0.121*	-0.782*	-2.633*	1.917*	-0.691*	0.359
1.0	-0.096*	-4.562*	-5.200*	1.116*	-1.613*	0.420
1.5	-0.046*	-14.763*	-17.018*	-5.212*	-7.958*	0.628
2.0	-0.016*	-22.355*	-23.952*	-5.640*	-10.658*	0.858

Notes: The choice sets were generated with the stochastic frontier model with the site efficiency component assumed to be distributed as truncated normal and the random error assumed to be normally distributed. * and ** denote significance at the 0.01 and 0.05 levels, respectively.

Appendix E. Conditional logit estimation results: deterministic choice sets

Travel Time	β_{Cost}	β_{Salmon}	$\beta_{Brook\,Trout}$	$\beta_{Lake\,Trout}$	$\beta_{Brown\,Trout}$	Pseudo R^2
5.0	-0.152*	2.517*	0.555*	3.712*	2.527*	0.313
4.5	-0.152*	2.517*	0.555*	3.712*	2.527*	0.313
4.0	-0.152*	2.517*	0.555*	3.712*	2.527*	0.312
3.5	-0.152*	2.517*	0.555*	3.712*	2.527*	0.311
3.0	-0.152*	2.517*	0.555*	3.712*	2.527*	0.309
2.5	-0.152*	2.517*	0.555*	3.712*	2.527*	0.304
2.0	-0.152*	2.517*	0.555*	3.714*	2.528*	0.291
1.5	-0.151*	2.516*	0.558*	3.750*	2.534*	0.261
1.0	-0.143*	2.511*	0.581*	4.127*	2.568*	0.197
0.5	-0.039*	2.362*	0.596*	5.493*	2.679*	0.081
0.3	0.062*	2.625*	0.638*	5.242*	3.040*	0.119

Notes: * denotes significance at the 0.01 level.