Full information selection bias correction for discrete choice models with observation-conditional regressors

Y. Allen Chen[a], Alan C. Haynie[b], and Christopher M. Anderson[c]

[a] Fishery Resource Analysis and Monitoring Division, Northwest Fisheries Science Center, NOAA Fisheries, 2725 Montlake Blvd. E., Seattle, WA 98112, allen.chen@noaa.gov

[b] Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Bldg. 4, Seattle, WA 98115, alan.haynie@noaa.gov

[c] School of Aquatic and Fishery Sciences, University of Washington, Box 355020, Seattle, WA 98195, cmand@uw.edu

Abstract

We examine self-selection in polychotomous choice models that construct attribute values for each alternative conditioned on observed choices. Using observations made only when the alternative was chosen ignores private information which was a basis for the decision, biasing resulting estimates. We suggest a full information maximum likelihood procedure that performs well at the extremes of the choice set in our sample, and use an "identification at infinity" weighting to identify levels. We apply the model to understanding fishing location choice in the economically

significant Bering Sea pollock fishery, where expected catches at each location are constructed from harvests observed when that location is chosen.

1. INTRODUCTION

In econometric models of discrete choice, agents choose between options based on the expected attributes of the alternatives available to them. We investigate a class of models where certain attributes are only observed for the alternative actually selected by the agent, and show how private information impacts the agent's selection criteria and the data a researcher observes. An example is the eponymous "Roy Model" of migration (Roy 1951), where a researcher may hypothesize workers choose their eventual state of residence depending on the expected wages they will receive across locations. Because they only observe the realized wages in the state chosen, the researcher creates proxies from observed data for the other locations (Dahl 2002, Bertoli et al. 2013), in order to compare different geographic states. Similar intuition is applied in research explaining how households trade off climate amenities and expected wages (Sinha et al. 2018), how expected wages explain human migration (Parey et al. 2017), how recreators choose between recreational sites when some site amenity data are missing (Kinnell et al. 2006), how child care costs impact female labor supply (Kornstad & Thoresen 2007), or how teacher quality and expected test scores affect school choice and teacher choice (Jacob & Lefgren 2007), among others. In fisheries models of location choice, fishers choose where to fish based in part on their expectations of catch across polychotomous locations.

To model the spatial decisions of fishers, existing methods use observation-conditional catch data to predict expected catch at various locations. A researcher only observes catches at the locations chosen by the fishers. To create proxies, researchers frequently regress researcher-observed catches on chosen covariates (such as fisher characteristics or lagged catches), and then use the parameter estimates from the catch equation model to predict unobserved catches for locations.

Examples of such models evaluate how fishers trade off catch and cost expectations (Eales & Wilen 1986), vessel willingness to avoid common-pool bycatch (Abbott & Wilen 2011), the effect of spatial closures and marine reserves (Haynie & Layton 2010, Smith 2005), or the extent of information-sharing across fishermen (Smith 2000).

Such catch data are non-randomly sampled. A fisher may possess a diversity of private information not known to the researcher when they make a decision where to fish. Fishers may share information with each other in ways researchers cannot observe. In addition, fishers may follow an aggregation of fish across areas, such that they know catches will be large at their next location, even in the absence of previous visits (and therefore researcher-observed data) at that location. However, even if the distribution of the error with which researchers estimate expected catch is mean zero, the expected value of that error conditional on observing the catch is not. When fishers are more likely to choose locations with larger catches, researchers are also more likely to observe large, positive shocks.

A number of solutions exist to correct for selection bias in the sample of data a researcher uses to create predicted values, although they may either require strong distributional assumptions about the error terms or may not be generalized to models with polychotomous choices. In a Roy (1952) model estimating how migration is affected by expected earnings across locations, Dahl (2002) suggests a semiparametric correction function, noting that the mean of the conditional error term can be written as an invertible polynomial function of the probability that the location was chosen

(Ahn & Powell 1993),[1] which allows the researcher to forgo assumptions about the joint distribution of the error terms (e.g. Lee (1983) examines a similar problem where the distribution is assumed jointly normal). We contribute to the broader literature of modeling and correcting for selection bias, the seminal example of dichotomous choice found in Heckman (1979), by applying a full information correction that simultaneously estimates model parameters with correction functions in a polychotomous choice setting.

First, we propose an extension to previous models by simultaneously estimating attribute expectations (i.e., expected catches) within the discrete choice model. Instead of estimating the probabilities of choosing a location in a first stage, which are needed as covariates in Dahl's (2002) correction function, we simultaneously estimate the catch equation with a correction function and the discrete choice problem using full information maximum likelihood. To our knowledge, the first stage with correction has not been modeled jointly with the second-stage problem, as the second-stage equation of interest is not always a discrete choice problem, but may be a linear function instead (e.g. examining the magnitude of migration flows in Dahl 2002). Our Monte Carlo experiments suggest that the full information approach performs well at identifying coefficients at the extremes of the choice set. Second, we apply an "identification at infinity" weighting approach (Andrews & Schafgans 1998, Chamberlain 1986) that allows us to identify levels in the attribute equation; an intercept in the first-stage equation typically cannot be identified due to estimation of the correction function (Dahl 2002), however, we do so with an extension of the weighting

[1] The continuous nature of catch and revenue data makes the fisheries context a particularly suitable application of the correction.

approach to a polychotomous context. This is broadly useful for any application where the level of predicted values (e.g. wages, costs, test scores, fishery catches), is important.[2]

In the remainder of this paper, we first explain how the fisher uses private information about catches when they choose locations, and how expected catch is proxied by the researcher with error due to selection. Monte Carlo experiments illustrate how this biases parameter estimates, and how a correction function approach can test and correct for the bias. The experiments also suggest that a full information maximum likelihood procedure performs well at the extremes of the choice set, which is important in estimation of the discrete choice parameters. Finally as an example, we demonstrate the importance of selection in the U.S. Bering Sea catcher vessel pollock fishery. We can test the statistical significance of the correction function in order to ascertain whether self-selection exists in a model relying on non-standardized catch data recorded by onboard observers, and find the use of uncorrected fishery-dependent data results in underestimated welfare effects from a hypothetical spatial closure

2. Location choice and expected catch with error

Consider a stylized model where a fishing fleet harvests fish from the spatial distribution of a fish population that is on average time-invariant, such that some locations have larger catches on average than others. However, specific catches also vary from averages across time in some

[2] For example, Dahl (2002) does not require wage levels in his analysis of migration flows, however, in an application where wages enter a second-stage discrete choice problem (Bertoli et al. 2013), the wage intercept is not separately identified from the polynomial intercept.

unobservable, non-systematic way (e.g., as fish move to different locations). We can write the true, realized weight of fish caught (Y_{itk}) by fisher i at location k for observation t as a function of covariates (X_i, potentially vessel-specific), a location-specific parameter β_k that scales vessel characteristics to catch, and a stochastic catch deviation term u_{itk}, such that:

$$Y_{itk} = X_i{'}\beta_k + u_{itk}. \tag{1}$$

In (1), the attribute catch varies by location, and depends on covariates such as the size of the vessel, and we assume u_{itk} is a stochastic term representing the myriad of influences that can impact the fisher's catch that cannot be captured by the researcher's model. Therefore, $X_i{'}\beta_k$ represents the time-invariant average catch at location k for fisher i, but then catch can deviate from this average at any given observation.

We assume the stochastic catch deviation can be written as two parts, one part the fisher observes (u_{itk}^f), and one part the fisher does not observe (u_{itk}^s), such that

$$u_{itk} = u_{itk}^f + u_{itk}^s. \tag{2}$$

Furthermore, we make the following assumption such that both are independently and identically distributed mean zero random variables.

Assumption 1. $\begin{bmatrix} u_{itk}^f \\ u_{itk}^s \end{bmatrix} \sim \boldsymbol{N}\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\sigma_f} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma_s} \end{bmatrix}\right)$

Then it follows that u_{itk} is also a mean zero normally distributed random variable, a common assumption used in empirical studies.

Denote the fisher's information set I_f, which can contain private information that catches will be good at their next chosen location despite having not fished there yet. Specifically, we can define

$I_f = \{\beta_k,\ X_i,\ u_{itk}^f;\ \forall\ k\}$. The term u_{itk}^f allows fishers to share information amongst themselves through complex social networks in a way not observable to the researcher, for example.[3] Or, more skillful vessel skippers would know when catches are larger than average at a location and act accordingly. Although the fisher does not observe part of the stochastic deviation u_{itk}^s, fisher-specific knowledge would allow fishers to choose locations when they know the deviations of u_{itk}^f are positive and catches are larger.

Conversely, the researcher does not observe β_k or attribute levels Y_{itk} at locations not chosen. Rather, they only observe realized catches at locations fishers choose, denoted $\tilde{Y}_{itk}$, as well as fisher characteristics X_i, such that the information set of the researcher $I_r = \{X_i, \tilde{Y}_{itk}\}$. Then, the researcher must construct a proxy of attribute levels in order to compare locations, without observing the variation from the stochastic error, or knowing the true expectation function.

To create attribute expectation proxies researchers can regress observed catches on known covariates, and use the estimated $\widehat{\beta_k}$ to construct counterfactual estimates of expected catch:

$$E[\widehat{Y_{itk}}|I_r] = \hat{Y}_{itk} = X_i{}'\widehat{\beta_k}. \quad (3)$$

Note that (3) is generalizable to match contemporary methods of constructing catch expectations in fisheries economics. X_i could include covariates such as average catches over a more recent

[3] Studies such as Abbott & Wilen (2010) and Evans & Weninger (2014) have investigated if fishers choose to share information about catches amongst each other, although the existing research does not always find benefits to fishers.

period of time relative to the fisher's choice occasion (Eales & Wilen 1986), or weighted moving averages of different lag lengths to include both fine-grained and historical information (Abbott & Wilen 2010). Here we focus on a common approach that can be thought of as a vessel-specific average catch over the entire sample of data available to the researcher.[4] Our specification also corresponds better to more general economic models: for example, we could imagine expected wages on the left-hand side as a function of education, in models of human migration. Importantly, in all specifications the researcher does not observe the variation in catch expectations at each location (u_{itk}^{f}) that is observed by the fisher.

Because the researcher does not observe attribute levels at all locations, but only at locations chosen by the fisher, $\widehat{\beta_k}$ is a biased estimator. The fisher's choice problem in a standard random utility model assumes fishers choose to fish in location k if its expected utility U_k is greater than the utility in all other locations, or

$$U_k > U_m \; \forall \, m \neq k. \tag{4}$$

We assume the fisher's utility from alternative k depends on the marginal utility they derive from catch α, their starting location j, vessel- and location-specific variables that are costly to the fisher (Z_{ijk}; e.g. travel), a parameter γ that scales cost conditional on vessel characteristics, and a portion of utility unknown to the researcher ε_{itjk}:

[4] We would expect that as the catch expectation function becomes more fully specified, and fewer variables are omitted, the amount of private information available only to the fisher could decrease. However, note that the problem we describe in this paper pertains to a scenario where any information about catches remains available to the fisher but not to the researcher.

$$U_{itjk} = V_{itjk} + \varepsilon_{itk} = \alpha * (X_i{}'\beta_k + u^f_{itk}) - \gamma(Z_{ijk}) + \varepsilon_{itjk}. \quad (5)$$

The fisher's expected catch can be written as $E[Y_{itk}|I_f] = X_i{}'\beta_k + u^f_{itk}$, as they observe the part of the stochastic catch deviation that corresponds to their private information, and their expectation of u^s_{itk} equals zero ($E[u^s_{itk}|I_f] = 0$). We assume the unknown portion ε_{itjk} is assumed to be independently and identically distributed extreme value (Gumbel), and that the marginal utility of catch is positive.

Assumption 2. $\varepsilon_{itjk} \sim \boldsymbol{GEV}(\boldsymbol{\mu} \in \mathbb{R},\ \boldsymbol{\beta} > \mathbf{0},\ \boldsymbol{\xi} = \mathbf{0})$

Assumption 3. $\alpha > 0$

Then, the true probability fisher *i* chooses location *k* can be written as:[5]

$$Prob\left(U_{itjk} > U_{itjm}, \forall\, m \neq k; \alpha, \gamma, \beta_k, Z_{ijk}, X_i\right) = \quad (6)$$

$$\frac{\exp\left({}^{\alpha}/_{\sigma_{scale}} * \left(X_i{}'\beta_k + u^f_{itk}\right) - {}^{\gamma}/_{\sigma_{scale}}(Z_{ijk})\right)}{\sum_{m=1}^{m=M} \exp({}^{\alpha}/_{\sigma_{scale}} * (X_i{}'\beta_m + u^f_{itm}) - {}^{\gamma}/_{\sigma_{scale}}(Z_{ijm}))}.$$

Notice that in (6) the probability that the fisher chooses a location (and the researcher observes that catch) increases with larger, positive error realizations as long as $\alpha > 0$. The fisher's expected catch $E[Y_{itk}|I_f]$ depends on the private signal about catch deviations u^f_{itk}, and larger catches are associated with greater utility at a location. $E[u^f_{itk}|observe\ Y_{itk}] \neq 0$ is directly a result of the fisher's choice problem when specified as a random utility model (RUM), where fishers choose locations (and catches) that result in the greatest expected utility at that time, visiting locations when they have private information fishing is good at that location. Thus, the sample of observed

[5] Note that only 2 of the 3 parameters $(\alpha, \beta, \sigma_{scale})$ can be identified. In practical use these will typically be α and β divided by some unknown scale parameter.

catches is biased ($E[\tilde{Y}_{itk}] = X_i{}'\beta_k + E[u_{itk}|observe\ Y_{itk}]$), biasing estimates of $\widehat{\beta_k}$ as well.[6] Finally, any discrete choice model that empirically compares locations by inserting a prediction for the average catch $\hat{Y}_{itk}$ at each location, such as in equation (7), will also be biased.

$$Prob\left(U_{itjk} > U_{itjm}, \forall\ m \neq k; \alpha, \gamma, \beta_k, Z_{ijk}, X_i\right) = \frac{\exp\left({}^{\alpha}/_{\sigma_{scale}} * \hat{Y}_{itk} - {}^{\gamma}/_{\sigma_{scale}}(Z_{ijk})\right)}{\sum_{m=1}^{m=M} \exp\left({}^{\alpha}/_{\sigma_{scale}} * \hat{Y}_{itm} - {}^{\gamma}/_{\sigma_{scale}}(Z_{ijm})\right)}. \quad (7)$$

Because $\alpha > 0$ and catch enters utility positively in this example of fisher location choice, we expect that $E\left[u_{itk}^{f}\middle|observe\ Y_{itk}\right] > 0$, but we note in general the methods described in this paper are agnostic about the sign of the selection bias. As long as expected error in the conditional sample is non-zero, attribute level predictions are incorrect. This also implies we could assume the fisher has full information in the fishery ($u_{itk}^{s} = 0$) without loss of generality, as long as there is utility-maximizing behavior and $u_{itk}^{f} \neq 0$.[7] Specifically, additional noise from non-zero u_{itk}^{s} mitigates the impact from selection to the extent correlation between Y_{itk} and $E\left[Y_{itk}\middle|I_f\right]$ decreases.

3. Correcting selection bias

Because the researcher inserts incorrect proxies of catches in the discrete choice problem, they will misunderstand how fishers make trade offs between catches and costs. For example, if differences in expected catches between locations are underestimated, the researcher would observe fishers choosing to move to different locations, incurring travel costs, despite relatively small changes in

[6] Note $E[u_{itk}^{s}|observe\ Y_{itk}] = 0$, as neither the researcher nor fisher observes u_{itk}^{s}.

[7] To see this, notice that even if the fisher has perfect information and the researcher observes none of the stochastic portion of catch, but the fisher chooses locations randomly and not based on a selection criteria, parameter estimates in the catch equation would be unbiased.

proxied expected catch ($\hat{Y}_{itk}$). Then in order for the probability in (7) to match empirical choice patterns, the model would incorrectly infer fishers must derive large marginal utilities from small changes. A correction function approach allows us to both test for selection bias as well as estimate unbiased parameters for the catch distribution and choice components of the model.

We refer the reader to Dahl's (2002) paper for a complete explanation of the correction function, which approximates the conditional error as a polynomial function of the probability of visiting a location (p_{itjk}), where $\boldsymbol{\beta}_{prob}$ is a vector of coefficients to be estimated, with each coefficient corresponding to a polynomial term.[8] In addition, let $\widetilde{M}_{itjk}$ and M_{itjk} be indicator variables, the first denoting if the fisher moved or "stayed", and the second to which location they moved, which allows the conditional error to vary based on the moving decision. Note that moving or staying is not a nested decision, but rather "staying" denotes the fisher chose the same location (and incurred no moving cost).

To use the correction we assume that the probabilities used as covariates in the correction function are the only factors that influence the joint distribution (g_k) of the errors in the catch equation and a maximum order statistic summarizing the error terms in the selection equation. If we follow

[8] For example, a 3rd order polynomial correction function for a fisher that stayed at location *k* could be written as $c + \beta_{prob1} * p_{itjk} + \beta_{prob1} * p_{itjk}^2 + \beta_{prob3} * p_{itjk}^3$, where $\boldsymbol{\beta}_{prob}$ and constant c are estimated, and the probabilities p_{itjk} of fisher *i* staying at location *k* are included as covariates.

Dahl's notation such that $\vec{q}$ represents the chosen subset of the full migration probabilities $\{p_{itj1}, \ldots, p_{itjN}\}$ of moving to *{1…N}*, this can be written as:

Assumption 4. $g_k(u_{itk}, \max_m(V_m - V_k + \varepsilon_{itjm} - \varepsilon_{itjk}) \mid V_1 - V_k, \ldots, V_N - V_k)$

$$= g_k(u_{itk}, \max_m(V_m - V_k + \varepsilon_{itjm} - \varepsilon_{itjk}) \mid \vec{q}\,)$$

Then, if catches follow the process in (1), ($Y_{itk} = X_i{}'\beta_k + u_{itk}$), estimates of $\widehat{\beta_k}$ can be obtained by including an approximation of the conditional expectation $E[u_{itk} | observe\, Y_{itk}] \approx \eta(\widetilde{M}_{itjk}, M_{itjk}, p_{itjk}, \boldsymbol{\beta}_{prob})$ in ordinary least squares estimation of the regression:

$$\tilde{Y}_{itjk} = X_i{}'\beta_k + \eta(\widetilde{M}_{itjk}, M_{itjk}, p_{itjk}, \boldsymbol{\beta}_{prob}) + v_{itk}. \quad (8)$$

Following Dahl, we include a separate correction function for each location when a fisher moves, and for each location when a fisher "stays", thus allowing the conditional error to be different depending on the move/stay decision. With *K* locations there are therefore a total of *K*2* correction functions. Note that v_{itk} is an error term with mean zero in the *conditional* sample and u_{itk} is estimated as a function of the probability of moving to or staying at location *k*:

$$\eta\left(\widetilde{M}_{itjk}, M_{itjk}, p_{itjk}, \boldsymbol{\beta}_{prob}\right) = \quad (9)$$

$$\widetilde{M}_{itjk} \sum_{\mathrm{k}=1}^{K} [M_{itjk} * \eta_{itjk}(p_{itjk})] + (1 - \widetilde{M}_{itjk}) \sum_{\mathrm{k}=1}^{K} [M_{itjk} * \eta_{itjk}(p_{itjk})] =$$

$$\widetilde{M}_{itjk} \sum_{k=1}^{K} [M_{itjk} * (\sum_{q=1}^{q=Q} \beta_{prob,k,q} * p_{itjk}{}^{q} + \sum_{\tilde{q}=1}^{\tilde{q}=\tilde{Q}} \beta_{prob,k,\tilde{q}} * \left(p_{itjk}\tilde{p}_{itjj}\right)^{\tilde{q}}\,)\,)] +$$

$$(1 - \widetilde{M}_{itjk}) \sum_{\mathrm{k}=1}^{K} [M_{itjk} * (\sum_{\mathrm{q}=1}^{q=Q} \beta_{prob,k,q} * \tilde{p}_{itjj}{}^{q})].$$

The selection bias for each location is approximated in equation (9) with a polynomial function of *q* degrees. The probability that fisher *i* chooses location *k* is denoted p_{itjk}, while the probability that they stay is denoted $\tilde{p}_{itjj}$, where q is the power of the polynomial. Also note that in the correction function for movers, the polynomial of the moving probability and the polynomial of the interaction term need not be the same degree ($q \neq \tilde{q}$). The number of total parameters in the correction function then depends on the number of alternatives and the degree of the polynomial.[9]

By approximating the conditional error term with a polynomial function, and including it in the catch regression, we can purge the bias in $\widehat{\beta_k}$ and therefore obtain unbiased predictions of expected catch, which leads to accurate estimation of the discrete choice parameters. In addition, an advantage to using the correction function approach is that we can estimate the statistical significance of the correction functions. When the correction terms jointly are statistically significant, they indicate whether the conditional error is significantly different from zero, and whether self-selection occurs in the sample of data available to the researcher.

4. Full information maximum likelihood estimation

In order to empirically estimate probabilities (to then insert into the correction function), Dahl suggests partitioning data into "cells", where individual fishers within a cell have similar characteristics. The probabilities can be recovered as the proportion of individuals who move to each location, which allows individuals with different characteristics to be more or less likely to move to a given location, on average. Alternatively, the probabilities can be estimated from a first-stage discrete choice model (e.g., with conditional logit). Dahl notes the danger in using these probabilities in a two-stage approach if two locations are perceived to be similar (rather than independent) by individuals, potentially violating the independence of irrelevant alternatives assumption.

[9] Specifically, $(2(Q+1)+\tilde{Q})K$ parameters in the correction functions with K alternatives.

We evaluate two model-based methods of estimating the probabilities: a two-stage model using nonparametric cell probabilities, as well as a full-information model simultaneously estimating the probabilities as a function of catches. Our Monte Carlo experiments suggest the full-information model performs well at the extremes of the choice set in our example.[10] When evaluating our model using nonparametric cell probabilities, we calculate probabilities as the proportion of observations in which each vessel visits a given location (essentially treating each individual vessel as a "cell"), because we can exploit repeated observations from each fisher in our model, a unique feature of our fisheries data.

Conversely, with full information, the probabilities p_{itjk} in the correction function of the catch equation are no longer fixed, but rather updated as a function of the parameters in the fisher's utility. Specifically, we take advantage of the fact that the probability of choosing a location (or staying in the original location) can be calculated as part of the full likelihood:

$$Prob\left(U_{itjk} > U_{itjm}, \forall\, m \neq k; \alpha, \gamma, \beta_j, Z_{ijk}, X_i\right) \tag{10}$$

$$= \frac{\exp\left(\alpha/\sigma_{scale} * \hat{Y}_{itk} - \gamma/\sigma_{scale}\left(Z_{ijk}\right)\right)}{\sum_{m=1}^{m=M} \exp\left(\alpha/\sigma_{scale} * \hat{Y}_{itm} - \gamma/\sigma_{scale}\left(Z_{ijm}\right)\right)}$$

$$s.t.\ p_{itjk}{}^{n} = \left(\frac{\exp(\alpha/\sigma_{scale} * \hat{Y}_{itk} - \gamma/\sigma_{scale}(Z_{ijk}))}{\sum_{m=1}^{m=M} \exp(\alpha/\sigma_{scale} * \hat{Y}_{itm} - \gamma/\sigma_{scale}(Z_{ijm}))}\right)^{n}.$$

The full likelihood that fisher *i* chooses location *k* is then:

---

[10] An example of joint estimation of catch and location choice is the expected profit model of Haynie and Layton (2010), although we explicitly correct for selection in our problem.

$$l_{itjk} = \left(\frac{2\pi^{-\frac{n}{2}}}{\sigma_{catch}^{n}} \exp\left[\frac{-\Sigma\left(\tilde{Y}_{itk} - X_i{}'\beta_k - \eta\left(\widetilde{M}_{itjk}, M_{itjk}, p_{itjk}, \boldsymbol{\beta}_{prob}\right)\right)^2}{2\sigma_{catch}^2} \right] \right) \quad (11)$$

$$* \left(\frac{\exp\left({}^{\alpha}/_{\sigma_{scale}} * X_i{}'\beta_k - {}^{\gamma}/_{\sigma_{scale}} \left(Z_{ijk}\right)\right)}{\sum_{m=1}^{m=M} \exp\left({}^{\alpha}/_{\sigma_{scale}} * X_i{}'\beta_m - {}^{\gamma}/_{\sigma_{scale}} \left(Z_{ijm}\right)\right)} \right)$$

$$s.t.\ \eta\left(\widetilde{M}_{itjk}, M_{itjk}, p_{itjk}, \boldsymbol{\beta}_{prob}\right) =$$

$$(\widetilde{M}_{itjk}) \sum_{k=1}^{K} [M_{itjk} * \eta_{itjk} \left(\frac{\exp\left({}^{\alpha}/_{\sigma_{scale}} * \hat{Y}_{itk} - {}^{\gamma}/_{\sigma_{scale}} (Z_{ijk})\right)}{\sum_{m=1}^{m=M} \exp\left({}^{\alpha}/_{\sigma_{scale}} * \hat{Y}_{itm} - {}^{\gamma}/_{\sigma_{scale}} (Z_{ijm})\right)} \right)] +$$

$$(1 - \widetilde{M}_{itjk}) \sum_{k=1}^{K} [M_{itjk} * \eta_{itjk} \left(\frac{\exp({}^{\alpha}/_{\sigma_{scale}} * \hat{Y}_{itk} - {}^{\gamma}/_{\sigma_{scale}} (Z_{ijk}))}{\sum_{m=1}^{m=M} \exp({}^{\alpha}/_{\sigma_{scale}} * \hat{Y}_{itm} - {}^{\gamma}/_{\sigma_{scale}} (Z_{ijm}))} \right)].$$

Note that if the correction is successful and the parameters β_k are estimated without bias, the researcher is comparing unbiased estimates of average catch across locations in the discrete component of the likelihood.[11] The estimated correction $\eta(\cdot)$ varies across individual fishers, across chosen locations, and depending on whether the fisher moved or stayed, as it is a function of the indicator variables $\widetilde{M}_{itjk}$ and M_{itjk}, as well as the probabilities p_{itjk} that are updated as a function of the parameters in the fisher's utility, which depend on fisher characteristics. [12] The

[11] Note that the correction polynomial is not included in the discrete component of the likelihood because inclusion of the correction implies the researcher would be comparing $E[Y_{itk}|observe\ Y_{itk}]$ with $[Y_{itm}|observe\ Y_{itm}]\ \forall\ m \neq k$. Instead, we include the correction in the catch portion of the likelihood to obtain unbiased estimates of average catch, and then compare unconditional expectations of catch across locations.

[12] As noted above there are *(2*(Q+1)+$\tilde{Q}$)*K* parameters in the correction functions with *K* alternatives. Then, the total number of parameters here would equal *(2*(Q+1)+$\tilde{Q}$)*K* + $K*X_N$ +

correction provides $\sqrt{n}$-consistent and asymptotically normal estimates in the catch equation with continuous covariates and as the number of basis functions increase with the sample size (Andrews 1991, Newey 1997).

There are advantages and disadvantages to this full information approach. For example, the correction assumes we know the true probabilities of moving, but Assumption 2 implies we are placing a parametric assumption on the estimation of the probabilities in our application: namely that the selection equation errors are distributed extreme value. Estimates of the probabilities could be mis-specified, compared to Dahl's nonparametric approach.[13] However, this also allows us to use multiple continuous covariates to calculate probabilities rather than discrete cells, relaxing Dahl's assumption that agents in a cell are affected by moving costs, catches, etc. in the same way on average. For example, in Dahl's approach, it would make little sense to include catch expectations in the estimation of probabilities, as we expect observations of catch to be biased, but by simultaneously estimating corrected estimates of catch we can provide a potentially richer distribution of probabilities. This is important for ensuring a large number of distinct probabilities, mimicking continuous covariates for the basis functions.

$N*Z_N + 2$ where X_N and Z_N are the number of covariates in the catch and cost portions of utility respectively and the last 2 parameters are σ^n_{catch} and ${}^{\alpha}/_{\sigma_{scale}}$.

[13] We thank an anonymous reviewer for highlighting this tradeoff. In robustness checks with normal errors in the selection equation we did not find significant differences in our Monte Carlo results (available from the authors upon request); however, further research is required.

In addition, as Dahl notes, it would be natural to include probabilities of choosing other locations besides the chosen location in the correction function as well, at the cost of increasing the dimensionality of the problem. For better comparison, we follow Dahl's suggestion of adding only the probability of "staying" in the correction function. While it is feasible to include only the probability of the chosen location, as long as this probability conveys all information about catches in a chosen location (a condition Dahl refers to as the index sufficiency assumption), we note that an additional advantage of the full information estimation is that we can use the probabilities corresponding to an individual's 2nd-, 3rd-, 4th-, etc., best choices as well, as these are estimated in the full information method but not observed in cell probabilities.

Because previous literature typically estimates a first-stage regression with correction, and inserts predicted values using the first-stage estimates in a second-stage equation of interest, we compare non-corrected, two-stage (using cell probabilities), and full-information correction approaches in the next section. There are potential benefits from simultaneously estimating the corrected first-stage with the second-stage equation of interest, and we illustrate the asymptotic behavior of the full information estimation method with Monte Carlo simulations to demonstrate that selection is of empirical concern.

5. Monte Carlo experiment illustrates how Catch and Discrete Choice estimates are biased

We use a stylized model in a Monte Carlo experiment to demonstrate that fishers choose locations based on private information not known to the researcher, and this biases estimates of the marginal utility from catch in random utility models of location choice. For the data-generating process in our experiment there are *K=4* locations, where catch and utility vary across locations.[14] A given fisher *i* that is currently in location *j* chooses between *K* potential utilities:

$$U_{itjk} = \alpha * E\left[Y_{itk}|I_f\right] - \gamma(distance_{jk} * hp_i) + \varepsilon_{itjk}. \quad (12)$$

Here costs depend on the distance from their current location *j* to potential location k. In addition, distance is interacted with a fisher characteristic (e.g., vessel "horsepower" hp_i); vessels with more horsepower may have higher or lower costs of travel. We randomly generate uniformly distributed variables for horsepower such that $hp_i \sim U[1,10]$; note that the scale of the distribution is chosen for convenience and is generalizable as long as costs are scaled appropriately to the other variables in fisher utility (e.g., by scaling the γ coefficient on distance instead).

Fishers choose locations on a square grid, where the Euclidean distance to the adjacent grid square is parameterized to be 1.5 units.[15] In addition, ε_{itjk} is distributed Extreme Value Type I ($G(0,1)$) with mean equal to the Euler-Mascheroni constant (0.5772) and variance equal to $\pi^2/6$.

We assume catches follow:

$$(Y_{itk} = \beta_k * (grtons_i) + u_{itk}), \quad (13)$$

[14] We choose a relatively smaller number of locations which allows a relatively pronounced bias and computational simplicity that makes it easier to study correction.

[15] And the distance to the diagonal location is 2.12 units.

where fishers in vessels with greater gross tonnage catch more fish on average. The fisher characteristic $grtons_i$ is a scalar distributed $U[1,5]$,[16] while the error on the researcher's catch regression $u_{itk} = u_{itk}^f$ is normally distributed ($N(0,3)$), and $u_{itk}^s = 0$. For simplicity we will initially investigate a corner condition where the fisher observes Y_{itk} for all k, and chooses a location based on its observation, while the researcher constructs $\hat{Y}_{itk} = \hat{\beta}_k * (grtons_i)$ as described in Section 2. As the ratio of u_{itk}^s to u_{itk}^f increases, we may expect the effect from selection bias to decrease. The true catch coefficients β_k are described in the first column of Table 1.

Accordingly, the estimated probability that the fisher chooses location k is:

$$Prob\left(U_{itjk} > U_{itjm}, \forall\, m \neq k; \alpha, \gamma, \beta_k, distance_{jk}, hp_i, grtons_i\right) = \tag{14}$$

$$\frac{\exp(\alpha/\sigma_{scale} * \hat{Y}_{itk} - \gamma/\sigma_{scale}(distance_{jk} * hp_i))}{\sum_{m=1}^{m=M} \exp(\alpha/\sigma_{scale} * \hat{Y}_{itm} - \gamma/\sigma_{scale}(distance_{jm} * hp_i))}.$$

Because the scale parameter σ_{scale} cannot be identified, for the purposes of comparison in the Monte Carlo experiments we do not estimate and fix the cost parameter γ to (-1), and estimate the catch parameter α and σ_{scale}. This allows us to compare the marginal utility of catch parameter to its true value, and focus on determining the magnitude of the bias (as well as how sensitive fishers are to catch). Because catches are not observed by the researcher at every location for a given observation, we first estimate β_k in a first-stage regression, then use $\hat{\beta}_k$ to create proxies of

[16] Again note that we have chosen the scale of the distribution without loss of generality, as other variables and coefficients (such as α or β_k) in the fisher utility can be appropriately scaled if the fisher characteristic were to be changed.

catch $\hat{Y}_{itk}$, which are inserted in the fisher's utility for the discrete choice second-stage, and estimated using a conditional logit.

We generate 1000 choice occasions at each initial location k, where fishers on each choice occasion choose between K utilities and catches, given randomly drawn fisher characteristics, and the fisher chooses the location according to the selection criteria in (4), $(U_k > U_m \; \forall \, m \neq k)$. Note this model does not include or account for state dependence or dynamic choice: researchers observe a number of haul occasions, the locations chosen, the catches at the chosen locations, and the location of the previous haul.

5.1 Uncorrected results

As a baseline, first note in the left column of Figure 1 that when the private signal is absent in fisher catch expectations (i.e., $u_{itk} = u_{itk}^{f} = 0$), the conditional logit[17] produces unbiased estimates of α, where the true parameters are given in the last column. Next, when we introduce error in the catch equation, estimates for the marginal utility from catch are almost twice as large as the true means (in the second column of Table 1 and Figure 1 respectively). The reported values in Table 1 are the median values from 100 Monte Carlo iterations.

Fishers will choose locations with a smaller true mean whenever the private signal is positive and large for that time period. Researchers observe the catch plus the private signal, predicted catches are overestimated, and tradeoffs such as from a spatial closure can be underestimated. We examine possible welfare effects in our empirical example in Section 6.

[17] All discrete choice models use modified routines from the FishSET R package.

Note in the second column of Table 1 that when we estimate catch according to (3), without correction, the bias is not proportional across locations (for each estimated beta). For example β_4 is estimated less accurately compared to locations 1-3. The selection bias for locations with smaller average catches tends larger because a particularly large shock is necessary for a fisher to visit that location. This is important for the estimation of a second-stage conditional logit model because the selection bias does not fall out of the probability in the second stage when we use the estimated betas to create predicted catch values as proxies for expectations. Differences in catch across locations are underestimated, which causes overestimation of the marginal utility from catch in the conditional logit step (third column of Table 1).[18] The researcher incorrectly believes the fisher is willing to move to locations for small increases in catch, while the true fisher expectation at those locations is actually larger than what the researcher estimates.

Finally the similarity in researcher-estimated expected catches matches empirical data patterns in many fisheries where catches are “hyperstable”, and do not change much across locations (e.g. Rose and Kulka 2011). Hyperstable catches can suggest fish stocks are healthy, however, our model shows that hyperstability can be a result of selection, and not the true underlying heterogeneity in site quality, as our true means across locations do vary.

[18] Here the number of observations in the data-generating process is important, potentially impacting the direction of the bias in the marginal utility of catch, which can be positive or negative. The direction depends on the similarity of catches across locations, and we examine some causes in Appendix A.

Table 1: Monte Carlo comparison of catch parameter and marginal utility from catch estimates, between no correction, two-stage, and full information correction models.

	I. True parameters	**II. Catch with error**			**III. Two-stage catch equation with correction function**			**IV. Full information maximum likelihood with corrected catch**		
Parameter		*Estimated parameters*	*Standard error*	*Percent bias (from true)*	*Estimated parameters*	*Standard error*	*Percent bias (from true)*	*Estimated parameters*	*Standard error*	*Percent bias (from true)*
$\boldsymbol{\beta_1}$	1.50	2.01	0.02*	34.0	1.40	0.04*	-6.7	1.48	0.04*	-1.3
$\boldsymbol{\beta_2}$	1.25	1.88	0.02*	50.4	1.22	0.05*	-2.4	1.24	0.03*	-0.8
$\boldsymbol{\beta_3}$	1.00	1.75	0.02*	75.0	1.04	0.06*	4.0	0.98	0.04*	-2.0
$\boldsymbol{\beta_4}$	0.75	1.61	0.03*	114.7	0.86	0.07*	14.7	0.72	0.06*	-4.0
$\boldsymbol{\alpha}$	3.00	5.88	0.27*	96.0	4.29	0.20*	43.0	3.10	0.37*	3.3
			*α=0.05		F-test[40, 3956] = 40.02*					

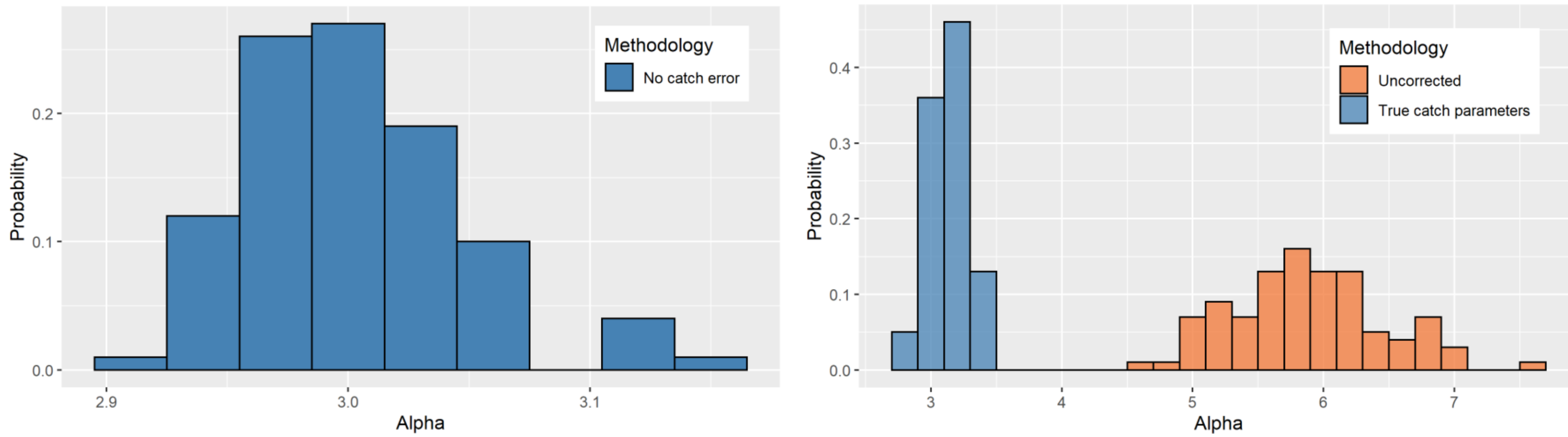

Figure 1: Discrete choice estimates from a baseline model without private signal, and uncorrected estimates when there is error in the catch equation.

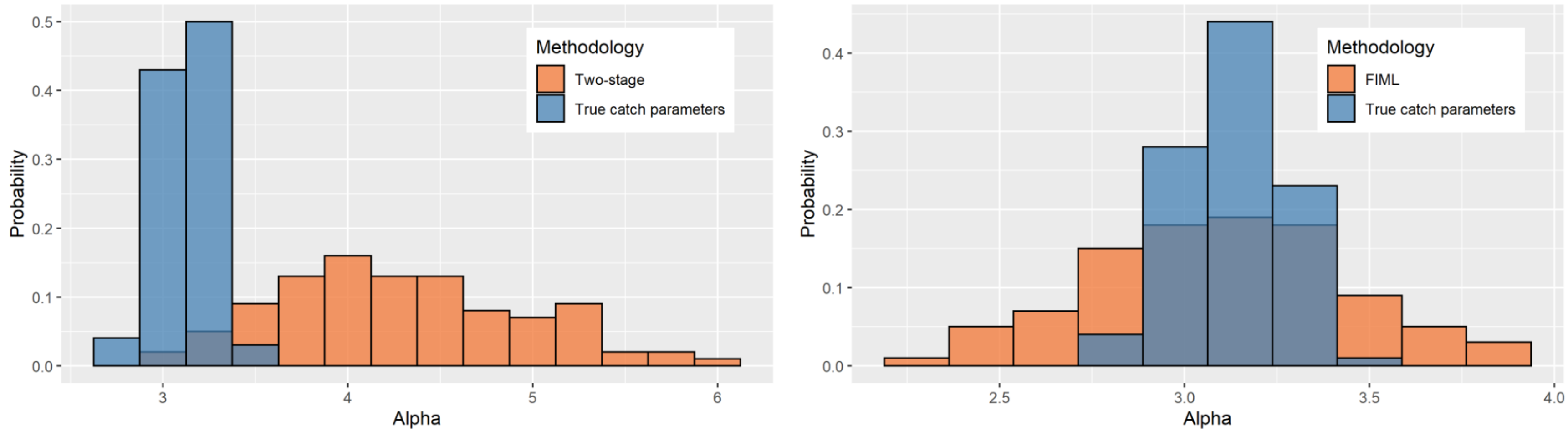

Figure 2: Corrected discrete choice estimates, and full information maximum likelihood discrete choice estimates.

5.2 Corrected results: two-stage and full-information

To eliminate the bias, we introduce and compare two correction functions, a two-stage model mimicking Dahl's method and a full-information model. We follow the convention described in Section 3, with both "stayer" and "mover" correction functions of degree 3, for a total of 8 correction functions, with a 2^{nd}-order polynomial in the interaction between the probability of moving and the probability that they stayed. We generally find the choice of polynomial is robust to smaller-order polynomials in the simulation example (as low as 2^{nd}-order polynomials), and the use of larger-order polynomials is at the cost of computational efficiency.[19]

Note that the two-stage application effectively uses Dahl's cell probabilities. Because we can exploit repeated observations from each fisher in our model, we calculate probabilities as the proportion of observations in which each vessel visits a given location (essentially treating each individual vessel as a "cell"). Then, individuals with the same characteristics are affected by differences in catch and moving costs in the same way on average. Using these cell probabilities we estimate equation (13) with correction functions appended, in a first stage with ordinary least squares and recover $\hat{\beta}_k$. Then we create proxies based on those estimates, which are inserted into the discrete choice problem in a second stage (i.e. equation (14)).

[19] In the simulation as well as the empirical example we tested multiple models with different polynomial degrees, as well as models with and without a "stayer" correction function. While this was important to ensure robustness, additional work on best practices to choose the polynomial degree q is required.

Column III of Table 1 reports an F-statistic that implies the data is inconsistent with the hypothesis that the correction function terms are equal to zero. Because the private signal in the uncorrected catch equation is proxied by the correction functions, we can conclude that selection bias occurs in this simulated fishery, given the terms of the correction are jointly significant (they are different from zero at any level of statistical significance). However, although the corrected estimates improve the conditional logit estimates of the cost and catch coefficients, they cannot completely correct the second-stage bias. The traditional two-stage correction appears to work best for median values of β: the bias is larger for β_1 and β_4 in the third column of Table 1. Even if the bias in the corrected attribute-level equation is much smaller than the uncorrected estimates, we still consistently underestimate locations with larger true catches and overestimate locations with smaller true catches. This structure at the extremes of the choice set turns out to have implications for estimation of the discrete choice parameters, and the marginal utility of catch (α) remains overestimated in column III, because we observe fishers moving to locations for small perceived increases in catch.

Jointly estimating the discrete choice portion of the likelihood and the corrected catch function in a full-information model (equation (11)), we find that the selection bias in the catch equation is close to zero, while estimates of the marginal utility from catch appear both unbiased and consistent in the right frame of Figure 2. Because the second-stage equation of interest is often not the discrete choice problem itself in Roy models of migration, joint estimation to our knowledge has not been investigated in this literature. However, by examining the fourth column of Table 1 we also see that estimates of α are improved because small differences in the catch parameters can result in relatively large biases in the discrete choice estimates.

Although the two-stage method corrects much of the selection bias, our Monte Carlo experiments suggest the remaining structure of the bias that remains can have a large effect on the discrete choice parameters and any welfare implications drawn from the models. Alternatives with larger catches are underestimated, while those with smaller catches are overestimated. Meanwhile, the full-information model performs relatively well at the extremes of the choice set, which allows it to recover the discrete choice parameters more reliably. However, we note these results are specific to the data-generating process we've investigated here, and additional work is required.

5.3 Bounding the efficacy of full information maximum likelihood

Finally, one factor that explains differences among uncorrected, two-stage, and full-information models is the quantity of private information available to the fisher, and specific fisheries may have more or less private information that the researcher cannot observer. Therefore, we investigate the robustness of the proposed methods, by repeating the Monte Carlo experiments above, re-estimating the model (and correction functions) as we increase the private information available to the fisher relative to average catches. These experiments follow the data-generating process outlined in Section 5, but vary the standard deviation of $u_{itk}^{f} \sim N(0, \sigma)$. Performance is measured by estimation of the marginal utility of catch parameter (α) in the second-stage discrete choice model, whose true value is still equal to 3.

Figure 3 shows that full information maximum likelihood performs well even as we increase the variance of the error term in the catch equation (on the x-axis), and that joint estimation maintains its advantage over the two-stage model as selection bias increases. These data point is the median value from 100 Monte Carlo iterations, for each unique standard deviation of u_{itk}^{f}. Unsurprisingly, when there is no private information all methods perform well at recovering α. However, both the two-stage and uncorrected methods perform worse as the private signal becomes larger.

While the two-stage correction estimator corrects most of the bias in the first-stage and improves the estimation of the discrete choice parameters, the two-stage model still overestimates the impact of catch on fisher utility in this example, because differences in expected catch across locations are still underestimated. Again, our simulation suggests that when the second-stage equation of interest is a discrete choice problem, small errors in the catch equation can have large effects in the second stage, in particular when there is structure in the bias across alternatives (here from underestimating differences). In contrast, full-information maximum likelihood estimation behaves well even as the variance of the error term increases. However, additional work is required to investigate the robustness of these results to other data-generating processes beyond this sample of data.

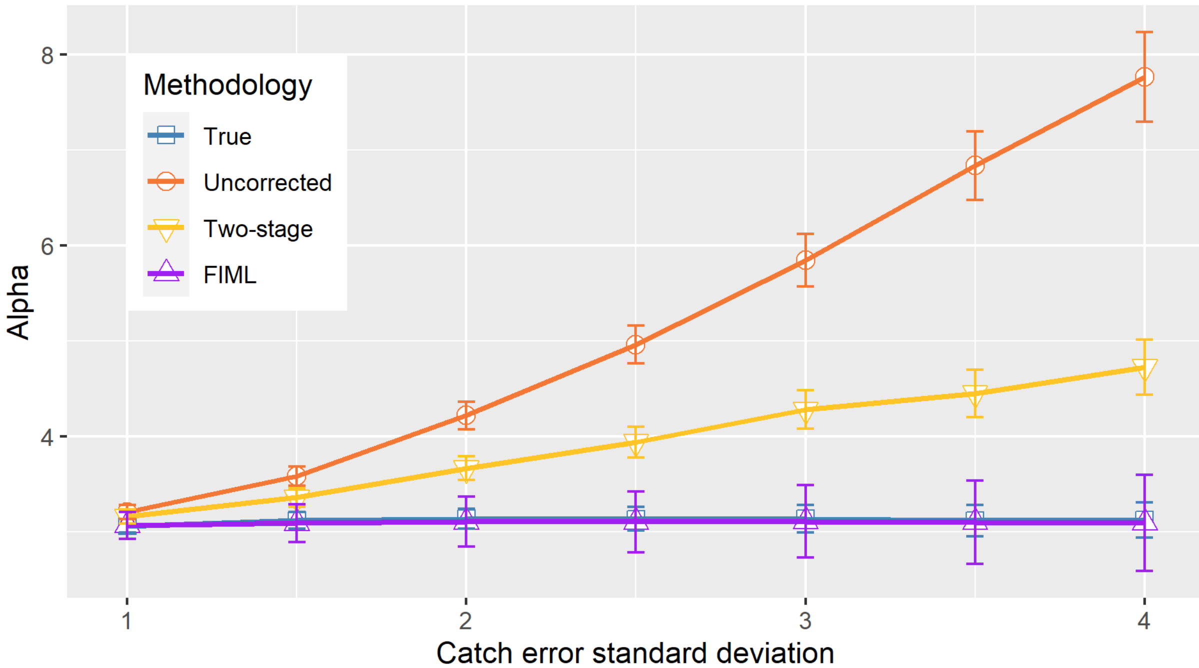

Figure 3: Bias in Monte Carlo discrete choice estimates increases with catch error.

6. Empirical example in the Bering Sea pollock catcher vessel fishery

We demonstrate the importance of correcting for selection bias with a hypothetical closure applied to an empirical example in the Bering Sea pollock fishery for the 2015 summer "B-season". In this fishery and year-season, 72 catcher vessels delivered to the inshore processing sector, comprising approximately 45 percent of the total catch in that year-season (the total catch includes catcher vessels that deliver to shore-based processors and fish that are caught and processed at sea). Table 2 suggests these catcher vessels exhibit considerable variance in the size (by gross tons), age, and horsepower across the fleet. For the purposes of estimation, we normalize vessel characteristic data such that the mean is one for each characteristic, and catch and distance are rescaled ensuring they are of similar magnitude (divided by one hundred).

The choice set for the individual fisher is discretized into areas that are 1 decimal degree east-west by 0.5 degrees north-south, known as "Stat6" areas designated by the Alaska Department of Fish and Game (ADFG). There is a tradeoff between a finer spatial resolution and maintaining enough observations in each grid cell to identify the coefficients in the correction functions; recall that the estimated probabilities need to exhibit considerable variation over different vessels and tastes. As opposed to states or cities in a Roy model which are well-defined alternatives, analysts must make judicious choices as to how to discretize their study area at sea.[20] While standard best practices have been elusive, with potential options varying across fisheries depending on the natural

[20] We thank an anonymous reviewer for helping draw the distinction between alternatives available in fishery applications and that of Roy models in labor settings, and for emphasizing the importance of robustness checks in different choices of grids.

variability of catches and the definition of time intervals (see e.g., Depalle et al. 2021), we note this is an area for future study. Analysts would be well-served to check the robustness of their results to different grid discretization choices.

Figure 4 maps the areas visited by fishers in the B-season of 2015, as well as sample sizes and average catches at each location with a minimum of 20 observations. Catcher vessels that operate in the fishery tend to choose locations closer to Dutch Harbor and Akutan (the two offloading ports). These vessels also participate in a number of inshore cooperatives as a result of the American Fisheries Act, and we can investigate and test whether member vessels may share information (which implies the amount of private information available to fishers but not observed by researchers could be large).[21] Stat6 areas to the northwest generally have fewer observations and smaller average catches, but the researcher cannot ascertain if differences in catch are understated due to selection, or if observed catches actually describe the underlying state of the stock.

We choose to examine the B-season because the tradeoffs across locations (i.e., between catch and distance) are substantially different in the winter A-season when high-valued roe enters the choice calculus of the fisher and at which point vessels are also restricted by ice cover at different times. For each haul, the researcher observes the vessel's starting location (the end point of the last

[21] For example, in-season management is dictated by a cooperative manager who is responsible for communication within the fleet.

haul),[22] the vessel's characteristics, the location the vessel chooses, and the weight of the catch at the chosen location. We note again that we abstain from dynamic planning, potentially ignoring non-independence of repeated samples. The pollock catcher vessel fishery tends to have fewer hauls within each trip, before returning to port (compared to pollock catcher-processors). We speculate that the separate corrections for movers versus stayers may allow vessels that choose a location and stay there to be treated differently from vessels that are actively searching and following fish aggregations; however additional study is warranted.[23]

[22] Here we abstain from using the first haul of a trip as the previous location is the nautical port.

[23] In addition, we also note again that repeated observations may actually assist in calculating cell probabilities in a two-stage application, potentially providing as many cells as the number of vessels.

Table 2: Vessel characteristics in 2015 B-season.

	Age (years)	**Horsepower**	**Gross tons**	**Catch per haul (metric tons)**
1st quantile	35.0	1200.0	193.0	52.1
Mean	37.5	1901.0	372.8	95.9
3rd quantile	40.0	2000.0	394.0	129.2

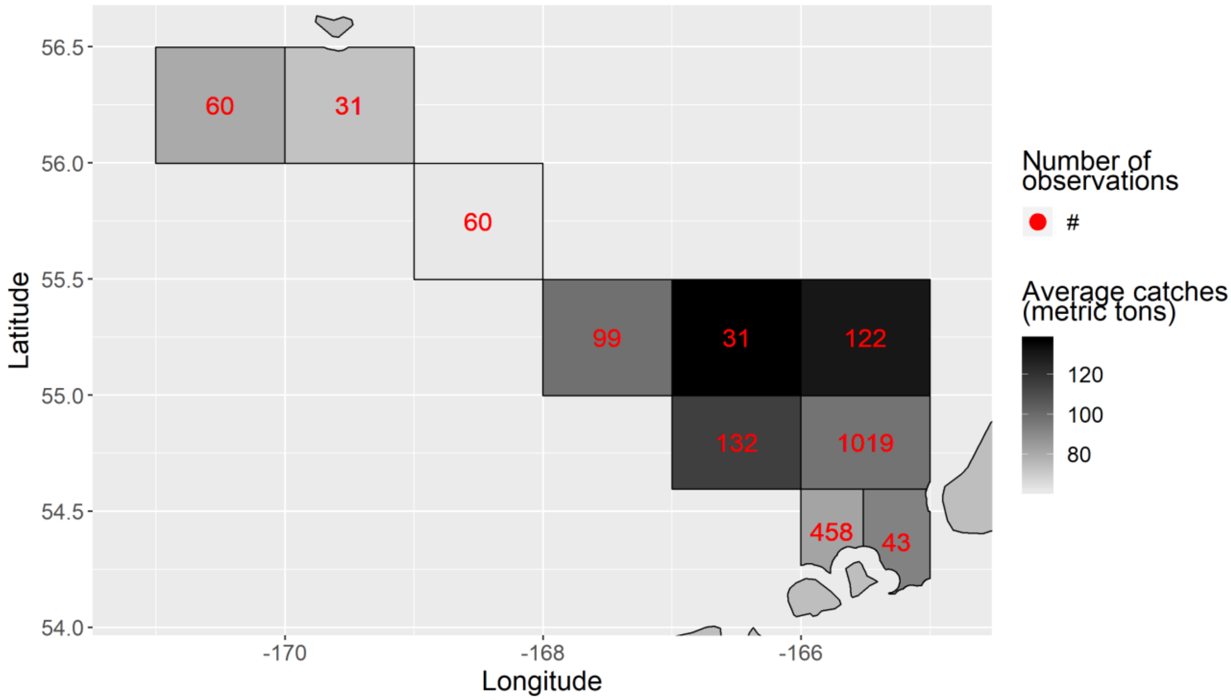

Figure 4: Number of hauls and observed average catch (metric tons) per location.

To ascertain if vessels tend to travel farther distances only when catches will be good in those locations, we use the correction function in a joint estimation methodology. The catch equation we estimate is similar to (1), except with a scalar represented by vessel age (*age*) interacted with vessel horsepower (*hp*) as the single vessel-specific covariate (15), and a constant c_k multiplied by unity. Meanwhile, we assume our cost equation (16) is a function of vessel characteristics (including gross tonnage, *grtons*) interacted with distance, as well as a linear component on mileage.

$$Y_{itk} = c_k + \beta_k * (age_i * hp_i) + u_{itk}. \tag{15}$$

$$C_{ijk} = \gamma_1 * (distance_{jk}) + \gamma_2 * (distance_{jk} * grtons_i) + \gamma_3 * (distance_{jk} * hp_i) + \gamma_4 * (distance_{jk} * age_i). \tag{16}$$

A potential issue arises if an intercept exists in the catch equation. As Dahl (2002) notes, an intercept in the equation of interest is not separately identified from the constant in the correction polynomial. Even if we seek to impose a restriction such that the constant in the catch equation equals zero, for example to ensure that a physically non-existent vessel with zero horsepower or age must have zero catches and that catches remain non-negative, the constant that remains in the polynomial still absorbs any explanatory power that would be attributed to the catch equation constant.

We use an extension of a weighting method for dichotomous problems from Andrews & Schafgans (1998) that works reasonably well for polychotomous situations in Monte Carlo simulations (Appendix B), where we only estimate the catch equation constant for a location as the probability of choosing that location goes to unity. The intuition from Heckman (1990) is that as the probability of choosing a location goes to unity, the selection bias term should go to zero. Equation

(17) illustrates how the weighting function $K(p_{itjk})$ weights both the polynomial and the catch constant in the full likelihood:

$$l_{itjk} = \left(\frac{2\pi^{-\frac{n}{2}}}{\sigma_{catch}^{n}} \exp\left[\frac{-\Sigma\left(\tilde{Y}_{itk} - K(p_{itjk})c_k - \beta_k * (age_i * hp_i) - (1 - K(p_{itjk}))\eta(\tilde{M}_{itjk}, M_{itjk}, p_{itjk}, \boldsymbol{\beta}_{prob})\right)^2}{2\sigma_{catch}^2}\right]\right)$$
$$* \left(\frac{\exp\left({}^{\alpha}/_{\sigma_{scale}} * (c_k + \beta_k * (age_i * hp_i)) - {}^{\gamma}/_{\sigma_{scale}}(Z_{jk})\right)}{\sum_{m=1}^{m=M} \exp\left({}^{\alpha}/_{\sigma_{scale}} * (c_m + \beta_m * (age_i * hp_i)) - {}^{\gamma}/_{\sigma_{scale}}(Z_{jm})\right)}\right) \quad (17)$$
$$s.t.\ K\left(p_{itjk}\right) = 1 - \exp\left(-\frac{p_{itjk}}{bw - p_{itjk}}\right).$$

The weighting function we use is suggested in Andrews and Schafgans (1998), where we choose a bandwidth of unity. Note a restriction of $c_k = 0$ still requires the weighting if the restriction is to hold. While previous methods would allow the recovery of the returns from vessel gross tons parameter, here we can recover levels and estimates of corrected catches at different locations as well, and we are unaware of previous applications of this method to specifically polychotomous models.

Recall that we have normalized vessel characteristics to unity, and therefore the marginal disutility of distance evaluated at the mean can be written as the sum of the cost function parameters. In Table 3 we find full-information estimation with correction infers a smaller marginal utility of catch, as well as a smaller marginal disutility of distance, and the ratio of catch to distance is both smaller and significantly different compared to uncorrected estimates[24]. While we cannot directly

[24] The standard errors for the disutility of distance are calculated using the delta method. Taking the ratio of utility from catch to disutility from distance accounts for the unknown scale parameter. A full suite of parameter estimates of the FIML model can be found in Appendix C.

compare likelihoods and model criterion as the underlying data is not the same (the full information model also includes the likelihood for the catch equation), we can compare likelihoods associated with the choice probabilities $\left(\frac{\exp\left(\alpha/\sigma_{scale} * X_i{'}\beta_k - \gamma/\sigma_{scale}(z_{jk})\right)}{\sum_{m=1}^{m=M}\exp\left(\alpha/\sigma_{scale} * X_i{'}\beta_m - \gamma/\sigma_{scale}(z_{jm})\right)}\right)$, where the full information model maximum log-likelihood is larger (-1816.83 versus -1820.30).

When we do not include a correction for selection, we infer larger predicted catches at locations that require larger travel costs, such that tradeoffs between locations will be underestimated. Figure 5 illustrates that uncorrected predicted catches are very similar across all locations, including those not visited often in the northwest (which are larger compared to the minimum predicted catch). Vessels are only willing to go to locations further away when catches are especially good, or when catches are poor elsewhere, biasing predicted catches in those locations upwards. To show this, we can test whether our approximation of the conditional error is significantly different from zero. We also can directly compare likelihoods to a joint model with no correction function, which is a nested model.

Table 3 shows a likelihood ratio test rejects the null (no correction) model. Correction functions for each location, as well as statistical significance of individual correction functions can be found in Appendix D. Seven out of 10 of the correction functions enter significantly at the median probability. These results suggest that selection bias is of empirical concern in this fishery. Interestingly, Table 3 also shows the pseudo R^2 of both models are very similar, which implies a model with reasonable fit[25] can still estimate inaccurate welfare impacts and incorrect predicted catches.

[25] Defined as a percentage as the starting log-likelihood less the fitted, divided by the starting. McFadden (1977) notes that values of 0.2-0.4 are reasonably well fit for the pseudo R^2.

Table 3: Discrete choice parameter estimates and model statistics.

	α	**γ_1**	**γ_2**	**γ_3**	**γ_4**	**$\Sigma\gamma$**	**Choice log-likelihood**
FIML	3.49	1.09	-0.25	-0.44	-5.12	-4.71	-1816.83
SE	*0.78*	*0.80*	*0.19*	*0.32*	*0.75*	*0.14*	
Uncorrected	8.31	-3.96	0.10	-0.96	-0.86	-5.68	-1820.30
SE	*0.63*	*1.14*	*0.30*	*0.45*	*1.04*	*0.15*	

Model statistics	**FIML**	**Joint with no correction**
Joint log-likelihood	-3123.40	-3262.64
AIC	6338.80	6557.28
AICc	6340.95	6557.55
BIC	6597.69	6647.33
Pseudo R^2	0.26	0.23
LR test (H0: joint estimation with no correction; dof = 30)	278.47	

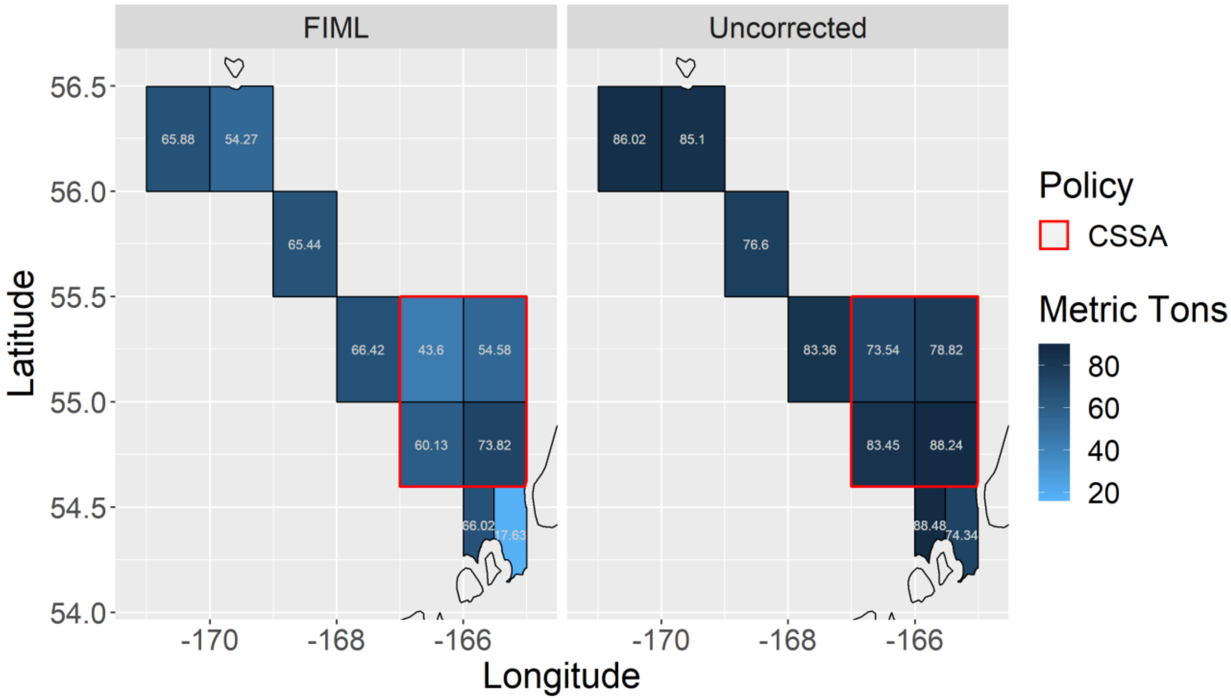

Figure 5: Predicted catches.

In addition, *absolute* catches are predicted to be larger under the uncorrected model as well. Average catches in the FIML model are 57 metric tons, with a standard deviation of 16, while average catches in the uncorrected model are both larger and exhibit less variance (82, standard deviation 6). A full table of predicted catches can be found in Appendix C: Table of predicted catches. These results overestimate the quantity of fish in the sea, along with misestimating welfare effects. Fishers and regulators often arrive at different conclusions as to the health of fishery stocks, and selection by the fisher can be one possible reason, as fishers tend to visit locations where fishing is good and catches are bountiful.

Finally, we can use the log-sum formula (Train 2009) to calculate percentage welfare changes from a hypothetical spatial closure. The enclosure in Figure 5 delineates the areas in the choice set that overlap with the Chinook Salmon Savings Area (CSSA) as defined by Amendment 58 (2000)[26]. The CSSA was closed in the B-season after September 15th if a fixed limit of Chinook salmon bycatch was attained. This CSSA became a back-up regulation after rolling hotspot closures became regulator measures in the fishery in 2006 and the closure was subsequently removed in 2011 when Chinook catch limits and other bycatch reduction measures were implemented through Amendment 91 to the BSAI FMP. We show the welfare loss to the fleet if a

[26] The Chinook Salmon Savings Area actually extends an additional 0.10 decimal degrees south into Stat6 areas 655409 and 655401; however, for the purposes of this hypothetical illustration we only examine closing intact Stat6 areas. Also note that the CSSA is larger than the shown enclosure, which only represents the areas in the choice set that overlap with the CSSA.

hypothetical 2015 summer season-long closure had been implemented, in Figure 6, faceted by vessel horsepower.

Welfare losses are much larger under full-information estimation than the uncorrected model, while the difference increases with horsepower. In addition, we note that absolute welfare losses tend to decrease as vessel horsepower increases, consistent with previous findings (Haynie & Layton 2010). These vessels have more fishing power and size, and spend more time fishing on trips where others may be limited by keeping fish fresh enough to deliver (Watson and Haynie 2018).

Because catches outside the hypothetical spatial closure are very similar and predicted to be larger under the uncorrected model, the welfare impact of the Chinook salmon savings area is underestimated. A spatial closure has very little effect on welfare in the uncorrected model as catches are predicted incorrectly to be similar everywhere. The correction in the full-information model infers that locations that are infrequently visited exhibit an upwards bias in predicted catch, and vessels only tend to visit those locations when fishing is good. The researcher would incorrectly believe the next-best options for fishers are relatively similar to catches within the hypothetical closure, and therefore inaccurately estimate smaller forgone benefits

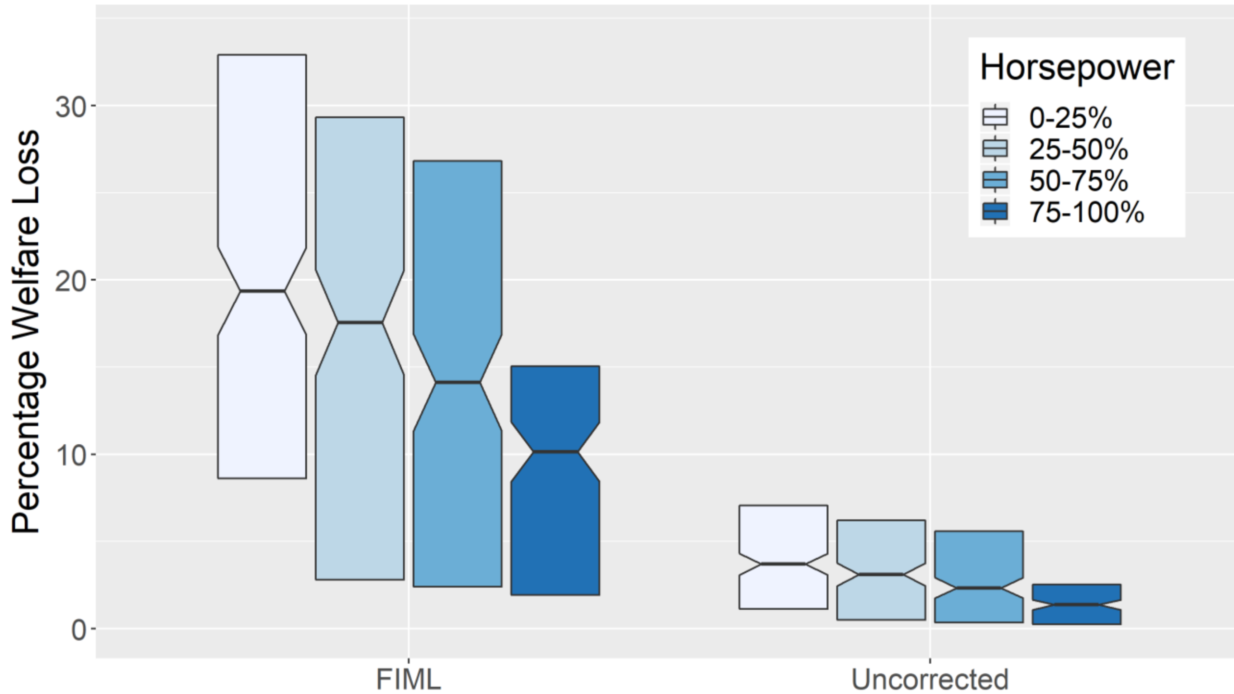

Figure 6: Welfare loss by vessel horsepower from hypothetical spatial closure of the Chinook Salmon Savings Area based on the 2015 summer Bering Sea pollock catcher vessel fishery.

7. Discussion

This paper illustrates how private information available to the fisher and unknown to the researcher is not accounted for in standard catch expectation proxies created by researchers in fisher discrete choice models. Because fishers are more likely to choose locations with larger catches, researchers are also more likely to observe large, positive catch deviations when a particular area is chosen. An empirical example in the Bering Sea pollock fishery shows that fishers only visit locations farther away when fishing in those areas is relatively good, which underestimates the welfare impacts from a hypothetical spatial closure.

We suggest an extension to the Dahl's (2002) correction function method by jointly estimating the corrected catch equation with the polychotomous discrete choice problem, in order to correct the selection bias that occurs in catch expectation proxies due to non-randomly sampled data. Using a Monte Carlo experiment, we show how full information maximum likelihood estimation can purge the bias from predictions of catch, which allows the researcher to correctly infer how fishers trade off expected revenues and costs. We find that while the two-stage method corrects much of the selection bias, the structure of the bias that remains can have a large effect on the discrete choice parameters. Applications where the second-stage equation is also the discrete choice problem lend themselves well to use a full-information model, and we show that simultaneous estimation performs well in correction at the extremes of the choice set. By applying a weighting method (Andrews & Schafgans 1998) to our polychotomous application, we are also able to recover the intercept in our first-stage catch equation. While levels in the first stage typically cannot be identified in polychotomous models correcting for selection, from a practical perspective it is broadly important in order to understand the health of fishery stocks.

Our methods explicitly acknowledge that the fisher has information not known to the researcher when the fisher makes a decision where to fish, and the sample of catches the researcher uses to construct catch expectation proxies is selected by the fisher with the intention of increasing their catch and maximizing their utility. This can occur when the availability of fish varies over time: for example, a skillful captain may be able to successfully follow an agglomeration of fish across space, or fishers may share information in a way a researcher cannot observe. Therefore, the researcher would tend to observe catches at certain locations when the fishing is good.

Incorrectly predicting the spatial opportunities for fishing implies researchers will underestimate the welfare effects from policies such as spatial closures. When relative differences across locations are underestimated, a researcher would inaccurately believe the next-best options for fishers are close substitutes. In reality, the researcher cannot observe catches at locations the fisher does not choose, and the fisher chooses infrequently visited locations only when they have private information the catches will be large there.

These methods may be extended to any polychotomous choice problem that requires constructing proxies for unobserved alternatives and are relevant to the broader literature examining self-selected data; for example, examining how migration flows are affected by expected wages across geographic regions. We note that the results we present are a function of the data and fishery we choose to investigate. The methods presented are agnostic to the nature and sign of the bias, and if no bias exists, the polynomial terms can be jointly tested under the null that the expected conditional error is equal to zero.

Finally, this paper uses a relatively stylized model that does not account for state dependence or dynamic decision-making, and treats the catch expectations associated with all hauls within one season of fishing as coming from the same choice set. An avenue for future work is to examine how the correction function works with more robust constructions of expected catch, such as weighted averages that use historical catches of different time series lengths and spatial sizes, and to investigate the magnitude of unobserved heterogeneity across various fisheries. Because the polynomial function used to approximate the conditional error is straightforward to add to any linear relationship, and can be used to test whether selection actually occurs in a given set of data, the methods outlined here are relevant to a large number of fisheries and econometric problems. Models that do not test and correct for selection bias risk incorrectly inferring how fishers make tradeoffs between catches and costs and underestimating the impacts from spatial policies that affect the fisher's choice set.

References

Abbott, Joshua K., & James E. Wilen (2011). Dissecting the tragedy: A spatial model of behavior in the commons. *Journal of Environmental Economics and Management*, *62*(3), 386–401. https://doi.org/10.1016/j.jeem.2011.07.001.

Abbott, Joshua K., & James E. Wilen (2010). Form and Structural Models Voluntary Cooperation in the Commons ? Evaluating the Sea State Program with Reduced Form and Structural Models. *Land Economics*, *86*(1), 131–154.

Ahn, H., & Powell, J. L. (1993). Semiparametric estimation of censored selection models witha a nonparametric selection mechanism. Journal of Econometrics, 58(1-2), 3-29.

Donald W. K. Andrews. (1991). Asymptotic Normality of Series Estimators for Nonparametric and Semiparametric Regression Models. Econometrica, 59(2), 307-345. doi:10.2307/2938259

Andrews, Donald W., & Marcia M. Schafgans (1998). Semiparametric estimation of the intercept of a sample selection model. *The Review of Economic Studies*, *65*(3), 497-517.

Bertoli, Simone, Jesus Fernández-Huertas Moraga, & Francesc Ortega (2013). Crossing the border: Self-selection, earnings and individual migration decisions. *Journal of Development Economics*, *101*(1), 75–91. https://doi.org/10.1016/j.jdeveco.2012.09.004.

Chamberlain, Gary (1986). Asymptotic efficiency in semi-parametric models with censoring, *Journal of Econometrics*, 32, issue 2, p. 189-218.

Dahl, Gordon B. (2002). Mobility and the Return to Education: Testing a Roy Model with Multiple Markets. *Econometrica*, *70*(6), 2367–2420.

Dépalle, M., Sanchirico, J. N., Thébaud, O., O'Farrell, S., Haynie, A. C., & Perruso, L. (2021). Scale-dependency in discrete choice models: a fishery application. Journal of Environmental Economics and Management, 105, 102388.

Department of Commerce, National Marine Fisheries Service, N. O. and A. A. (2010). 15 CFR Part 902; 50 CFR Part 679; Fisheries of the Exclusive Economic Zone Off Alaska; Chinook Salmon Bycatch Management in the Bering Sea Pollock Fishery; Amendment 91. *Federal Register*, *75*(167), 53026–53074.

Department of Commerce, National Marine Fisheries Service, N. O. and A. A. (2000). 50 CFR Part 679; Fisheries of the Exclusive Economic Zone Off Alaska; Bering Sea and Aleutian Islands Area; Amendment 58 To Revise the Chinook Salmon Savings Areas. *Federal Register*, *65*(198), 60587–60590. https://doi.org/10.1007/s10750-004-4538-3.

Eales, James, & Wilen, James E. (1986). An Examination of Fishing Location Choice in the Pink Shrimp Fishery. *Marine Resource Economics*, *2*(4), 331–351. https://doi.org/10.1086/mre.2.4.42628909.

Evans, Keith S., & Quinn Weninger (2014). Information Sharing and Cooperative Search in Fisheries. *Environmental and Resource Economics*, *58*(3), 353–372.

Haab, Tim C., & Robert L. Hicks (1997). Accounting for Choice Set Endogeneity in Random Utility Models of Recreation Demand. *Journal of Environmental Economics and Management*, 127–147.

Haynie, A. C., & David F. Layton (2010). An expected profit model for monetizing fishing location choices. *Journal of Environmental Economics and Management*, *59*(2), 165–176. https://doi.org/10.1016/j.jeem.2009.11.001.

Heckman, James J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, *47*(1), 153–161. https://doi.org/10.3982/ECTA9859.

Heckman, James J. (1990). Varieties of selection bias. *The American Economic Review*, *80*(2), 313-318.

Jackson, Chris (2007). *Multi-state modelling with R: the msm package*. Cambridge, UK.

Jacob, Brian A., & Lars Lefgren (2007). What Do Parents Value in Education? An Empirical Investigation of Parents' Revealed Preferences for Teachers. *The Quarterly Journal of Economics, 122*(4), 1603-1637.

Kinnell, Jason C., Matthew F. Bingham, Ateesha F. Mohamed, William H. Desvousges, Thomas B. Kiler, Elizabeth K. Hastings, and Karen T. Kuhns (2006). Estimating site choice decisions for urban recreators. *Land Economics*, *82*(2), 257-272.

Kornstad, Tom, & Thor Thoresen (2007). A Discrete Choice Model for Labor Supply and Childcare. *Journal of Population Economics, 20*(4), 781-803.

Lee, Lung-Fei (1978). Unionism and Wage Rates : A Simultaneous Equations Model with Qualitative and Limited Dependent Variables. *International Economic Review*, *19*(2), 415–433.

McFadden, Daniel (1977). *Quantitative methods for analyzing travel behavior of individuals: some recent developments*. Berkeley: Institute of Transportation Studies, University of California.

Morey, Edward R., & Donald M. Waldman (1998). Measurement error in recreational demand models: the joint estimation of participation, site choice and site characteristics. *Journal of Environmental Economics and Management*, *35*(3), 262–276.

Newey, W. K. (1997). Convergence rates and asymptotic normality for series estimators. Journal of econometrics, 79(1), 147-168.

Parey, Matthias, Jens Ruhose, Fabian Waldinger, & Nicolai Netz (2017). The selection of high-skilled emigrants. *Review of Economics and Statistics*, *99*(5), 776-792.

Rose, G. A., & Kulka, D. W. (2011). Hyperaggregation of fish and fisheries: how catch-per-unit-effort increased as the northern cod (*Gadus morhua*) declined. *Canadian Journal of Fisheries and Aquatic Sciences*, *56*(S1), 118–127. https://doi.org/10.1139/f99-207

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, *3*(2), 135–146.

Sinha, Paramita, Martha L. Caulkins, and Maureen L. Cropper (2018). Household location decisions and the value of climate amenities. *Journal of Environmental Economics and Management*, *92*, 608-637.

Smith, Martin D. (2000). Spatial Search and Fishing Location Choice. *American Journal of Agricultural Economics*, *82*(5).

Smith, Martin D. (2005). State dependence and heterogeneity in fishing location choice. *Journal of Environmental Economics and Management*, *50*(2), 319–340. https://doi.org/10.1016/j.jeem.2005.04.001.

Train, Kenneth E. (2009). Discrete choice methods with simulation. *Cambridge University Press, New York*, (2nd), 1148. https://doi.org/10.1016/S0898-1221(04)90100-9.

Wickham, Hadley (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Retrieved from http://ggplot2.org.

Willis, Robert J., & Sherwin Rosen (1979). Education and Self-Selection. *Journal of Political Economy*, *87*(5)

APPENDIX A: THE DIRECTION AND MAGNITUDE OF THE BIAS IN DISCRETE CHOICE ESTIMATES

Another factor that impacts potential bias in estimated expected catches, and then estimates of the discrete choice parameters, is performance in smaller samples. A commonality in the Monte Carlo experiments above are a large number of samples at each starting location (with 1000 observations at each location). The exact direction of the selection bias can vary upwards or downwards however, and we demonstrate here the dependence on sample size, and how the direction of the bias in the marginal utility of catch (α) can be explained by inaccuracy in the parameter estimates in the catch equation, in conjunction with how similar locations are.

We summarize two implications from the simulations in this appendix: the first is that even with an arbitrarily large number of observations, the researcher still underestimates differences across locations and overestimates catches in absolute terms. Due to selection these estimates are biased in any finite sample. The second implication, however, is that when the number of observations is small these estimates are *also* inaccurate, and it is more likely the researcher can incorrectly predict the ordinal ranking of locations, such that relatively unproductive locations have larger catches than productive locations, which is exacerbated when locations are relatively similar to each other.[27] The latter impacts the direction of the bias in the marginal utility of catch.

First, Figure 7 plots the discrete choice estimates when there are only 100 observations at each starting location. Unsurprisingly, the distribution of estimates is much more dispersed, but also notice that a proportion of uncorrected α are also smaller than the true value (which is still equal

[27] The intuition here more closely follows the results found by Morey and Waldman (1998), who investigated the impact of measurement error on discrete choice modeling. They suggest a correction based on the fact that the number of choices observed for a location provides information on expected catches at that location. Note however that selection still biases the catch and discrete choice estimates in any finite sample.

to 3), although the average uncorrected α remains similar to Figure 1 (when there were 1000 observations at each starting location).

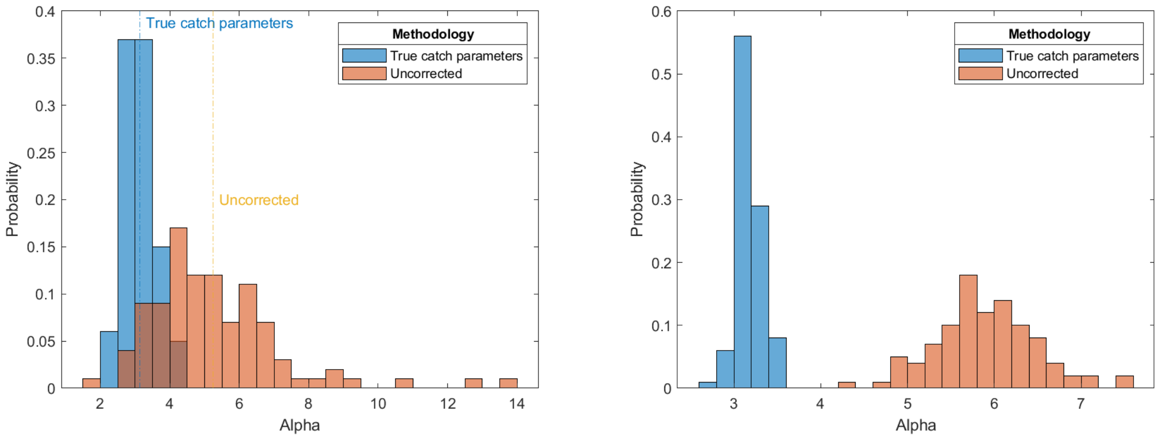

Figure 7: Uncorrected discrete choice estimates with 100 observations at each starting location (left) versus 1000 observations (right).

To understand the effect of a smaller sample, we illustrate a simplified example with only two locations, where we can simulate the effect from introducing bias in each catch parameter, while holding the other catch equation parameters (β_i) constant at their true values, and then estimating the discrete choice model using various values of β_i. In Figure 8 we re-simulate the discrete choice estimation to observe the effect on the marginal utility of catch (choices are not re-simulated, but rather we insert various values of β_i to observe the effect).

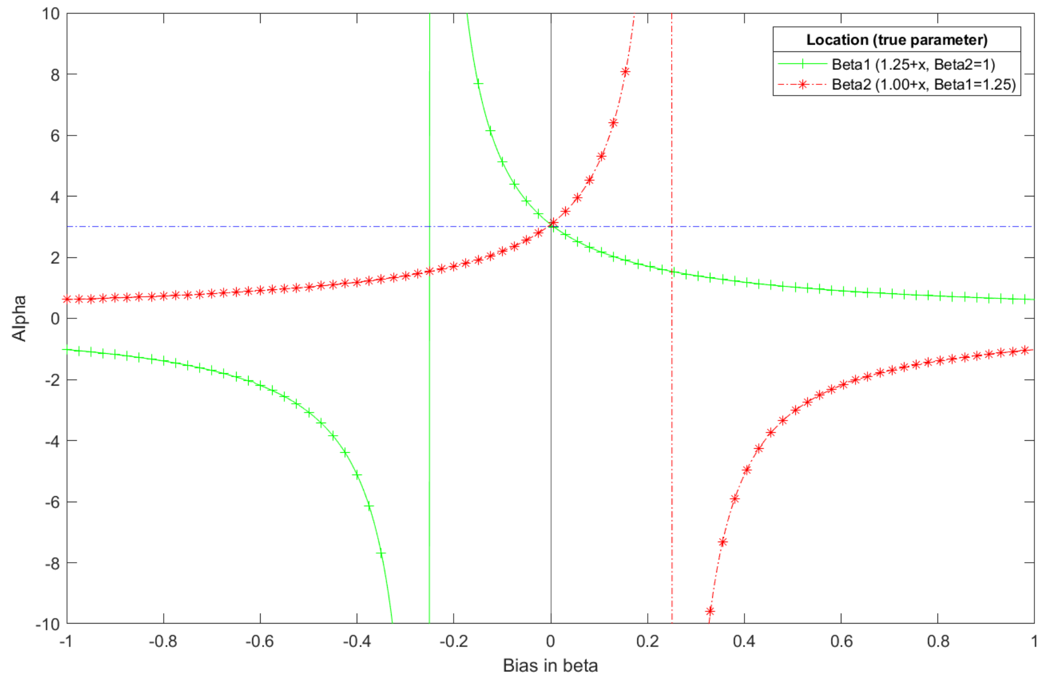

Figure 8: The impact of bias in the catch equation parameters on estimates of the marginal utility of catch.

First, when β_1 is underestimated, a location with larger catches on average, α is initially biased upwards. Conversely, overestimating β_2, a location with smaller catches, also biases α upwards. Interestingly, if instead of holding the other catch equation parameter constant, but bias in the catch equation parameters (β_i) was in the same direction and identical across β_i, there would be no bias in estimating α. These results are consistent with our previous Monte Carlo experiments, which emphasized the effect from underestimating differences across locations.

However, as bias in the catch equation parameters increases, the sign of the effect on α eventually changes. Notably, the effect on the marginal utility from catch (α) changes directions at asymptotes corresponding to 0.25 and -0.25, respectively.[28] With only two locations we can see the inflection point in the sign of the bias in α corresponds to when the researcher incorrectly changes the ordinal ranking of the locations by predicted catch. Specifically, the inflection point occurs when the researcher overestimates unproductive locations to the extent they believe the expected returns are larger than productive locations.

The ordinal ranking of locations is important because if the researcher observes vessels abstaining from visiting unproductive locations, but also incorrectly predicts large catches due to sampling variability, the model will infer vessels must suffer disutility from larger catches. This has the effect of changing the sign on estimates of the marginal utility from catch (α). Differences across locations are no longer underestimated, but rather the ordinal ranking of locations by expected catch has changed - locations with small catches are estimated to have large catches, and vice versa. These results are particularly stark with only two locations, but we can see similar patterns with four locations below.

The ordinal ranking of locations tends to be incorrect when the parameter estimates in the catch equation are inaccurate, such as when the researcher has few observations, or when locations are similar a smaller bias is sufficient to change the ordinal ranking of locations. Then, the researcher would observe unproductive locations with larger absolute catches than productive locations due to chance (i.e. sampling variability). Bias from selection therefore has two effects – not only are differences between expected catches across locations underestimated, but this also increases the likelihood that sampling variability might change the ordinal ranking of locations.

[28] The asymptotes occur because when catches are predicted to be the same across both locations, the model cannot identify the marginal utility from catch.

Notably, recall that the estimates in our Monte Carlo experiments exhibited an upwards bias in the marginal utility from catch. However, we are able to use a large number of observations and choose a data-generating process where the differences in average catches across locations are relatively large. An example such as Figure 8 shows that if observed catches across locations are similar, in a different fisheries context, and the researcher does not observe many samples, it would be possible for the marginal utility from catches to be biased downwards.

We also repeat the experiment with four locations. Again, we simulate the effect from introducing bias in each catch parameter, while holding the other catch equation parameters (β_i) constant at their true values. We re-simulate the discrete choice estimation to observe the effect on the marginal utility of catch (choices are not re-simulated, but rather we insert various values of β_i to observe the effect). We will refer to locations with larger average catches as "productive" locations, and locations with smaller average catches as "unproductive" locations.

First, when we underestimate β_i for productive locations, the marginal utility from catch α is initially biased upwards. For example, the first row of Figure 9 shows that underestimating β_1 results in estimates of α greater than the true value, when the bias ranges from 0 to approximately -0.2.[29] Conversely, overestimating unproductive locations (e.g. overestimating β_4) also biases α upwards.

[29] Recall that $\alpha_{true} = 3$.

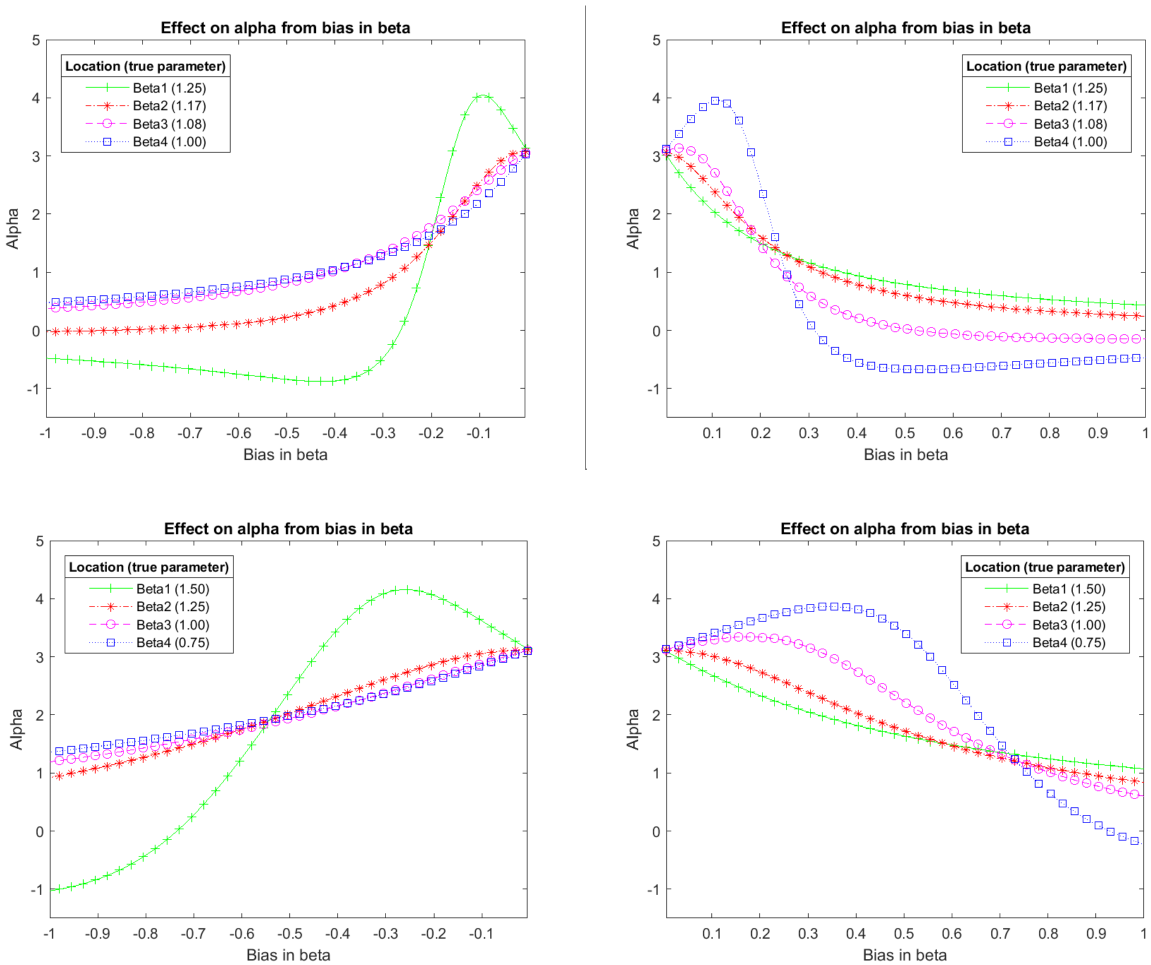

Figure 9: Bias in catch equation parameters.

However, as bias in the catch equation parameters increases, the sign of the effect on α eventually changes. For example, the first panel of Figure 9 shows that positive bias in the unproductive location corresponding with β_4 has a positive effect on α, but only while the bias in β_4 ranges from 0 to approximately 0.2. Subsequently, as the bias in β_4 continues to increase, the sign of the effect on the marginal utility from catch (α) flips, and estimates of α decrease below their true value: the bias in α, as positive bias in β_4 increases, is concave.

Again, there is an inflection point in the sign of the bias in α when the researcher incorrectly changes the ordinal ranking of the locations by predicted catch. We can see this in the second row of Figure 9, by investigating a data-generating process where the true differences across locations are more disparate. There, a larger bias (in absolute value) in β_i is required before the ordinal ranking of locations changes, and thus before the sign of the bias in α changes direction.

APPENDIX B: MONTE CARLOS WITH INTERCEPT

The experiments below follow the same as presented in the body of the paper, except with the inclusion of intercepts in the catch equations, whose true parameters are listed in the tables. The presented estimates are the median from 100 iterations. We also estimate the utility of catch $\boldsymbol{\alpha}$ and disutility of distance $\boldsymbol{\gamma}$ here, such that we should compare the ratios of $\boldsymbol{\alpha}$ to $\boldsymbol{\gamma}$ as both are proportional to some unknown scale parameter. In Table 4 we see that when we estimate the catch equation with error, and use those predicted catches in the choice model, the ratio of $\boldsymbol{\alpha}/\gamma$ is much larger than the true value of -3.

Table 4: Catch equation with error.

Location	**Estimated parameters**	**Standard error**	**True parameters**
$\boldsymbol{c_1}$	7.47	0.19	1.00
$\boldsymbol{c_2}$	7.44	0.19	3.00
$\boldsymbol{c_3}$	7.47	0.19	5.00
$\boldsymbol{c_4}$	7.73	0.19	7.00
$\boldsymbol{\beta_1}$	0.93	0.05	1.50
$\boldsymbol{\beta_2}$	0.95	0.05	1.25
$\boldsymbol{\beta_3}$	0.96	0.05	1.00
$\boldsymbol{\beta_4}$	0.91	0.05	0.75
$\boldsymbol{\alpha}$	2.75	0.14	3.00
$\boldsymbol{\gamma}$	-0.13	0.00	-1.00

Table 5: Full information maximum likelihood with corrected catch.

Location	Estimated parameters	Standard error	True parameters
$\boldsymbol{c_1}$	1.55	0.68	1.00
$\boldsymbol{c_2}$	3.60	0.45	3.00
$\boldsymbol{c_3}$	5.63	0.26	5.00
$\boldsymbol{c_4}$	7.49	0.18	7.00
$\boldsymbol{\beta_1}$	1.52	0.09	1.50
$\boldsymbol{\beta_2}$	1.26	0.06	1.25
$\boldsymbol{\beta_3}$	1.01	0.05	1.00
$\boldsymbol{\beta_4}$	0.78	0.04	0.75
$\boldsymbol{\alpha}$	0.46	0.05	3.00
$\boldsymbol{\gamma}$	-0.14	0.00	-1.00

If we use the weighting function in order to identify the intercepts in Table 5, we find that while the model performs much better, we are still unable to completely purge the bias from the catch constants. Better performance might be found in a different choice of weighting function or bandwidth, which we leave to further study. However, we do note that because the bias enters each location similarly (upwardly biased by roughly 0.50), it mostly falls out of the choice component, the returns from vessel gross tons remain accurately estimated, and the ratio of $\boldsymbol{\alpha}/\gamma$ is also similar to the true value (-3.28 vs. -3).

924 APPENDIX B: FIML MODEL FULL ESTIMATES

	Coef.	St. Err.	T-stat.
Marginal utility from catch	3.49	0.78	4.46
Catch beta 1	0.18	0.11	1.63
Catch beta 2	0.67	0.03	25.05
Catch beta 3	0.75	0.02	39.30
Catch beta 4	0.56	0.06	9.98
Catch beta 5	0.61	0.03	18.66
Catch beta 6	0.44	0.05	8.29
Catch beta 7	0.68	0.04	16.04
Catch beta 8	0.67	0.05	12.84
Catch beta 9	0.55	0.09	6.38
Catch beta 10	0.67	0.09	7.17
Polynomial constant 1	1.16	0.32	3.68
Polynomial constant 2	0.45	0.11	4.22
Polynomial constant 3	0.13	0.11	1.22
Polynomial constant 4	0.68	0.15	4.43
Polynomial constant 5	0.34	0.17	2.01
Polynomial constant 6	0.11	0.26	0.42
Polynomial constant 7	0.32	0.21	1.48
Polynomial constant 8	-0.26	0.25	-1.02
Polynomial constant 9	0.62	0.45	1.39
Polynomial constant 10	-0.14	0.47	-0.29
Polynomial 1st-order 1	-8.51	6.40	-1.33
Polynomial 1st-order 2	-1.37	0.87	-1.58
Polynomial 1st-order 3	-0.40	0.77	-0.52
Polynomial 1st-order 4	-2.11	1.48	-1.43
Polynomial 1st-order 5	1.03	1.56	0.66
Polynomial 1st-order 6	10.98	3.33	3.30
Polynomial 1st-order 7	-7.24	4.63	-1.56
Polynomial 1st-order 8	2.25	2.88	0.78
Polynomial 1st-order 9	-4.27	3.56	-1.20
Polynomial 1st-order 10	-1.68	3.63	-0.46
Polynomial 2nd-order 1	40.24	42.40	0.95
Polynomial 2nd-order 2	3.86	1.51	2.56
Polynomial 2nd-order 3	4.62	1.27	3.65
Polynomial 2nd-order 4	9.87	2.93	3.37
Polynomial 2nd-order 5	1.55	3.16	0.49
Polynomial 2nd-order 6	-24.97	10.06	-2.48
Polynomial 2nd-order 7	19.27	7.94	2.43
Polynomial 2nd-order 8	0.79	5.22	0.15
Polynomial 2nd-order 9	9.84	5.95	1.65

Polynomial 2nd-order 10	7.27	6.57	1.11
Disutility from distance linear miles	1.09	0.80	1.37
Disutility from distance miles and gross tons	-0.25	0.19	-1.38
Disutility from distance miles and horsepower	-0.44	0.32	-1.38
Disutility from distance miles and age	-5.12	0.75	-6.80
Catch function variance term	0.46	0.01	62.43

APPENDIX C: TABLE OF PREDICTED CATCHES

Table 6: Predicted catches between full information and uncorrected models.

ADFG Stat6 area	FIML (metric tons/100)	Uncorrected (metric tons/100)
655401	0.18	0.74
655409	0.66	0.88
655430	0.74	0.88
655500	0.55	0.79
665430	0.60	0.83
665500	0.44	0.74
675500	0.66	0.83
685530	0.65	0.77
695600	0.54	0.85
705600	0.66	0.86

APPENDIX D: CORRECTION FUNCTIONS

Figure 10 below illustrates the correction function for each location. Note that the shape of the corrections is explained by the weighting function that allows for identification of levels of catch at each location. Here, portions that are statistically significantly different from zero are highlighted in bold.

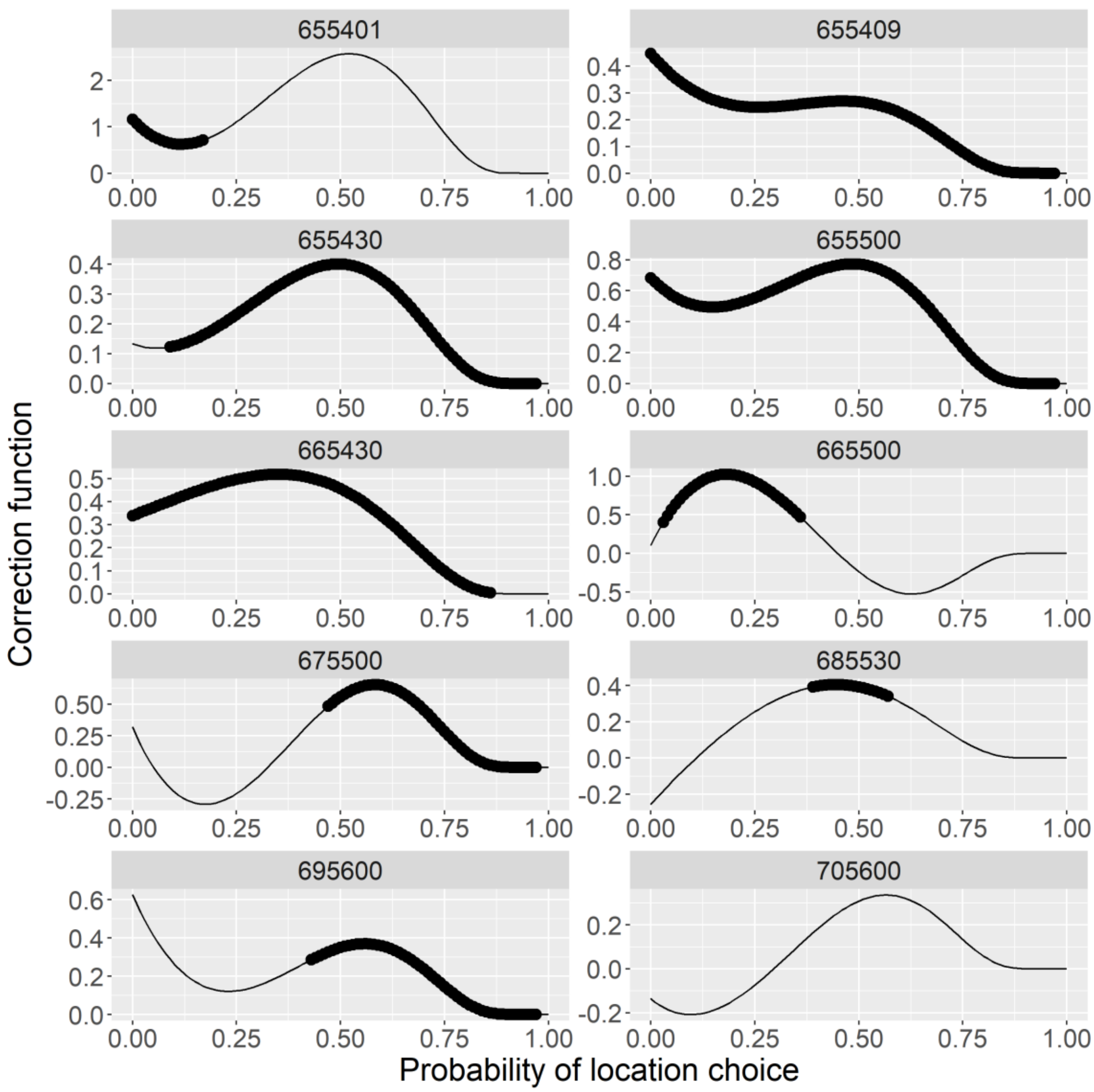

Figure 10: Correction functions at each location.

In addition, the statistical significance of each segment suffers when the support for the function is lacking. Figure 11 shows that the number of observations tends to match well with certainty around the correction function estimates. In addition, we generally have a good range of probabilities to estimate the correction function for each location, with the exception of ADFG areas 655401 and 665500.

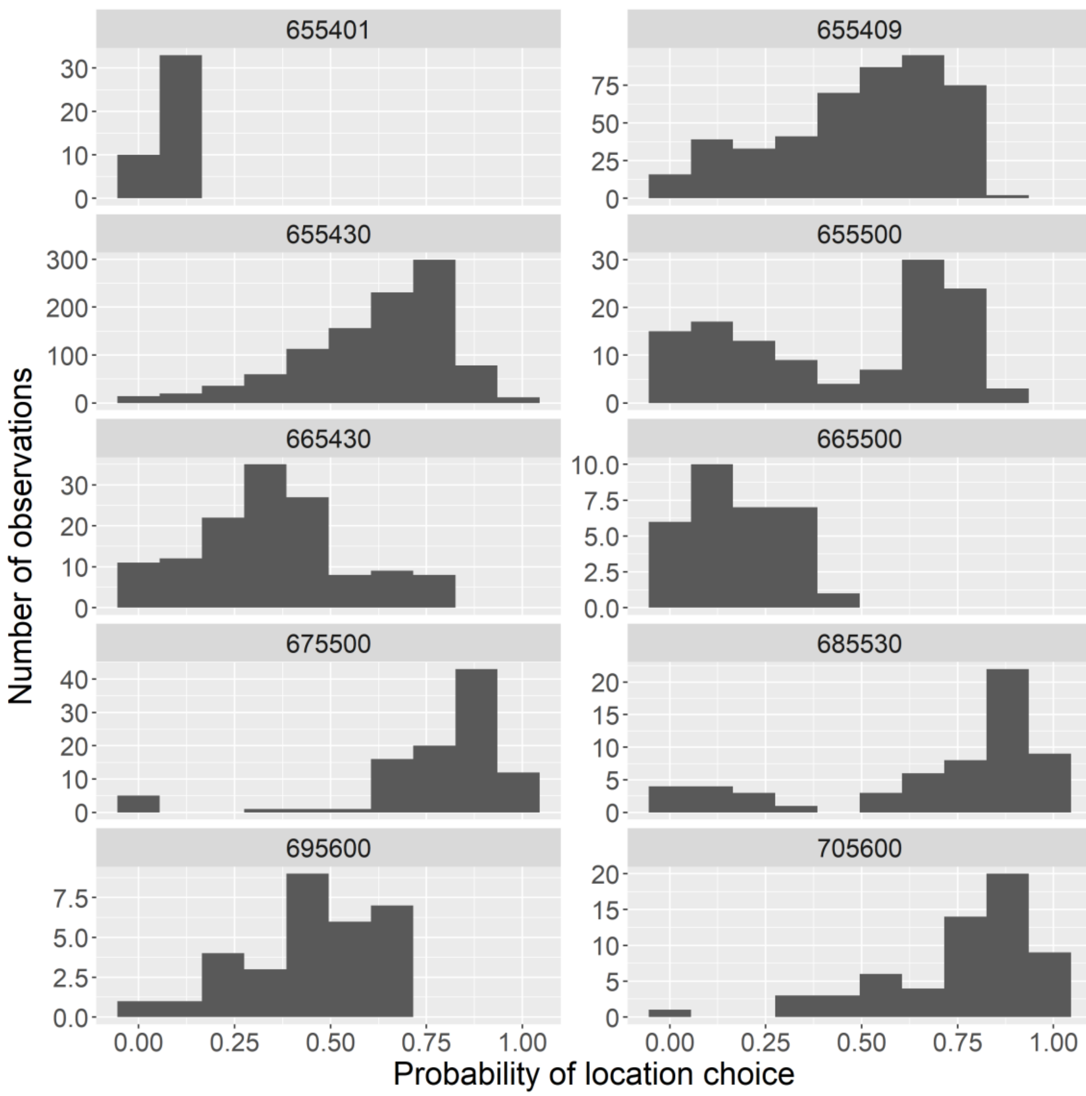

Figure 11: Number of observations given the probability of choosing a location.