

Using Auxiliary Population Samples for Sample-Selection Correction in Models Based on Crowd-Sourced Volunteered Geographic Information

Trudy Ann Cameron *Professor Emeritus, University of Oregon, Eugene; cameron@uoregon.edu*

Sonja H. Kolstoe *Research Economist, USDA Forest Service, Portland, Oregon; sonja.kolstoe@usda.gov*

ABSTRACT *Data from citizen science (CS) projects (and some social media) can offer selected samples with extensive information about human interactions with the natural world. Independently, we elicit levels of engagement with the eBird project from members of the eBird CS project and from a general-population sample. The general-population sample allows an ordered-probit model to explain propensities to engage with eBird at different levels, which we transfer to predict selection-correction terms for our independent sample of eBird members. We illustrate our method with a question posed only to our eBird-member survey sample about the radii of their individual spatial consideration sets for typical one-day birding excursions. (JEL Q26, Q57)*

1. Introduction

Observations on human interactions with nature are becoming increasingly plentiful with the growth in volunteered geographic information (VGI) that people contribute to citizen science (CS) projects (and via some types of social media).¹ VGI data provide a vast

amount of granular individual-level information about people's interactions with environmental goods and services—a potential gold mine of data for environmental and resource economists. However, the amount of VGI data a single person provides depends on their level of engagement with the data source, be it a CS project or a social media platform. In addition, these data pertain only to the contributing set of users—citizen scientists, social media users, and so on. These are samples of convenience rather than random samples from the overall population. Sample-selection bias is an obvious concern. Users' intensity of engagement or participation with the project or platform affects the likelihood that they appear in any data set used for statistical analysis. With careful attention to selection corrections, to control for intensity of engagement or participation, CS and other sources of VGI data may be a valuable research resource for environmental economists seeking to provide scalable and policy-relevant inferences.

Nonrepresentative voluntary surveys are often used by environmental economists to collect data. As a consequence, a variety of methods have been developed to correct for respondents' differing propensities to respond to the survey and be part of the estimating sample. These methods certainly include the traditional method of Heckman (1979).² Alternative ad hoc approaches have been proposed, as in Cameron and DeShazo (2013), Johnston and Abdulrahman (2017), and Kolstoe and Cameron (2017). Proper attention to systematic selection, and corrections (if indicated) can be important. Failure to account for sample selection can misrepresent the influence of low- and high-intensity participants

¹“Citizen science” or “community science” projects recruit volunteers from the general population to help scientists gather data about the natural world. CS projects have proliferated because of the growing ability of participants to contribute real-time field observations using convenient smartphone apps. As of February 2022, there are more than 2,000 active CS projects, according to the Citizen Science Association (see <https://citizenscience.org>), 493 are registered in the federal crowd-sourcing and CS registry (see the catalogue offered at www.citizenscience.gov).

²Heckman (1979) is the foundational paper for the least-squares context, now cited more than 11,500 times in Web of Science.

and may bias estimates of the population-level behavioral parameters of interest to economists. For data collected from CS participants, self-selection bias may arise from the potential correlation between the unobserved components of (a) their propensities to engage with the CS project to different degrees, and (b) their outcome variable of interest in statistical models concerning the environmental good being studied.

In this article, we develop a new approach to sample-selection correction for CS/VGI data. Our goal is to make any potential inferences based on such data more useful for policy makers. We illustrate our selection-correction strategies for a sample of birdwatchers who participate in the eBird CS project. The eBird project has proven to be a valuable CS/VGI data source for natural scientists and social scientists alike (e.g., Kolstoe and Cameron 2017; Kolstoe, Cameron, and Wilsey 2018; Roberts et al. 2017; Rosenberg et al. 2019). Furthermore, birdwatching is a very popular pastime. About 45.1 million people observed birds in the United States, both around home and away from home, according to the U.S. Fish and Wildlife Service's 2016 survey on Fishing, Hunting, and Wildlife-Associated Recreation (FHWAR) report (U.S. Fish and Wildlife Service 2018).

We use a survey of eBird CS members in the Pacific Northwest and a completely independent nationally representative sample from a survey of the general population of the United States. Both samples include specific information from respondents about the degree to which they participate in the eBird project, so that we can distinguish the extensive margin (whether an individual does or does not participate in eBird), and the intensive margin (the degree to which they engage with this CS project).

We propose three strategies for sample-selection correction. At the most basic level (with its details included the [Appendix](#)), we use our two samples to construct estimated heterogeneous sampling weights (for different levels of engagement intensity, controlling for the mix of individual characteristics in each sample). These weights serve to adjust the relative frequencies at different engagement

levels in our survey of eBird members so they more closely match the analogous relative frequencies at each level in the general-population sample.

Our second strategy, which is a more structural approach, adapts the standard two-stage Heckman correction method. We replace the Heckman first-stage binary-probit selection equation with an ordered-probit selection equation to explain six levels of engagement intensity. This selection equation still permits the calculation of an inverse Mills ratio term like the one that is key to the Heckman two-stage method. However, we estimate the selection equation using our general-population sample and then transfer it to our eBird member survey sample. As with standard selection-correction methods, this approach relies on strong assumptions about the joint error distribution and allows only the expected value of the outcome variable (i.e., the intercept of the outcome model) to be distorted by sample-selection bias.

Our third approach is more ad hoc. We transfer an engagement propensity function, estimated using our general-population sample, to the eBird member survey sample. Demeaned individual predicted engagement propensities in the eBird member survey sample, normalized on the mean engagement propensity in the general-population sample, are allowed to shift the intercept and slope parameters in the outcome equation of interest. We can then simulate the desired outcome equation if everyone in the eBird sample shared the same engagement propensities, identical to the mean selection propensity in the general-population sample.

To demonstrate our selection-correction strategies for CS data employing an auxiliary general-population sample, we model one particular outcome variable from our eBird survey: the radius of the respondent's so-called consideration set for one-day birding excursions. This variable is complementary to the idea of the relevant spatial market extent (or economic jurisdiction) for a specific recreational destination, as discussed by Loomis (1996), Walsh, Milon, and Scrogin (2011), and Glenk et al. (2020). Our study presents a unique opportunity to address consideration

sets because we included in the eBird member survey a specific question about how far each respondent would be willing to travel on a typical one-day birding excursion.³

Most previous research concerning recreational destination choices (e.g., Dundas and von Haefen 2020) has tended to use a common consideration-set radius for all individuals, often choosing a distance that has been used in other studies concerning similar environmental goods. Sometimes an assumption about a single common consideration-set radius is loosely informed by the upper percentiles of the observed marginal distribution of distances traveled across all trips in the data, as in Kolstoe and Cameron (2017). Other recent analyses have grid-searched across possible consideration-set radii and employed for all individuals the single radius that maximizes the model's likelihood (Holland and Johnston 2017). Here, we seek to identify individual systematic variations across our sample of eBird members in their directly elicited consideration-set radii. Our fitted radius function may then be transferable to other samples of birders from the general population, but only if the estimates are corrected for self-selection bias in our sample of eBird citizen scientists.

The consideration-set radius for an individual is related to other concepts in the revealed and stated preference literature. For example, Sen et al. (2014) adapt the terminology of "trip generation functions" (TGF) from the transportation economics literature on destination choices. A TGF, however, models the number of trips by an individual as a function of the observed travel time for a visit (controlling for origin and destination attributes). This approach does not focus on the maximum distance willingly traveled by a single person. Nor does it emphasize heterogeneity in this maximum distance across people with different characteristics.⁴

³The question explicitly excludes trips to destinations with reported rare bird sightings. Distances willingly traveled in those special situations can be much larger. This is consistent with the distinction between iconic and noniconic destinations in the related literature. See n. 5 in Glenk et al. (2020).

⁴An estimated TGF model could be solved for the average travel time (and therefore approximately the average

The spatial stated preference literature offers another related concept, referred to as "distance decay," reviewed by Glenk et al. (2020). Demand for visits to recreational sites is understood to decline with distance, holding everything else constant. However, distance decay can also reflect the fact that destinations at a greater distance face an increasingly large set of potential substitute destinations because the area of a circle around a given origin location increases much more quickly than the radius of that circle. Nevertheless, there appear to be very few examples in the literature where researchers have sought to identify individual-level heterogeneity in distance decay. A partial exception is Logar and Brouwer (2018), who find heterogeneity between urban and rural areas. In contrast, the individual consideration-set radius model we consider in this study acknowledges that there can be systematic differences across individuals in these radii, rather than just systematic differences over space that are shared by all individuals at a given location.

2. Ordered-Probit Strategies for Systemic Sample Selection

Our eBird member survey sample is self-selected, consisting only of eBird members who chose to respond to our survey. These birders are likely to participate in the eBird project with a different mix of engagement levels than might be expected for members of the general population. For this study, across several waves of the Qualtrics Omnibus (qBus) survey, we independently surveyed more than 4,000 respondents from that general-population panel. [Appendix A](#) offers some further discussion of our qBus sample, and [Appendix Table A2](#) contrasts a simple binary indicator for eBird CS participation, CS, with the greater level of detail in our six ordered categories of engagement intensity,

distance) at which expected trips fall to zero, conditional on origin and destination attributes. But these origin attributes would typically be population medians or proportions in the origin area, not individual characteristics.

CS6, elicited from the qBus data and our sample of eBird members.⁵

As noted in the introduction, standard selection-correction models use a binary selection model. We increase the level of detail by switching to an ordered-probit selection model, using six categories for our “selection into citizen science project participation intensity” model, where eBird is the specific CS project in question.⁶ Our general-population survey elicits six levels of eBird engagement intensity, and our eBird member survey questions elicit four corresponding levels of eBird engagement intensity, conditional (obviously) on at least some level of participation in eBird. Earlier binary selection models focus only on the extensive margin—the choice between participation versus nonparticipation. Our additional level of detail about engagement intensity provides unusual but valuable information about the intensive margin of participation in eBird for both samples.⁷

Selection in the General-Population (qBus) Sample

For the $i=1, \dots, N$ individuals in our general-population (qBus) sample, let CS participation intensity, $CS6_i$, take one of six levels, from “unfamiliar with the project” to “report virtually all of my observations.” For everyone, we have the same sociodemographic and income variables, Z_i , that we use to explain eBird participation intensity, where respondents $i=1, \dots, r$ participate in eBird at one of four different levels and respondents $i=s, \dots, N$ do not (but may either have heard about eBird, or not). If we sort these observations in decreasing order of participation intensity, the data for the selection model can be written as

$$[CS6]_{N \times 1} = \begin{bmatrix} 6 \\ \vdots \\ 5 \\ \vdots \\ 4 \\ \vdots \\ 3 \\ \vdots \\ 2 \\ \vdots \\ 1 \\ \vdots \\ \vdots \end{bmatrix}_{r \times 1}, \quad [Z]_{N \times k} = \begin{bmatrix} Z_{11} & \dots & Z_{k1} \\ \vdots & & \vdots \\ Z_{1r} & \dots & Z_{kr} \\ Z_{1s} & \dots & Z_{ks} \\ \vdots & & \vdots \\ Z_{1N} & \dots & Z_{kN} \end{bmatrix} [1]$$

For the qBus sample, we can model an underlying continuous latent propensity to be a member of eBird (denoted with an asterisk) as $CS6_i^* = Z_i \gamma + \eta_i$. We have observations at all six levels of participation intensity for the qBus sample. If the qBus sample also included information on our outcome variable of interest and a vector of regressors, called just y_i and X_i for now, there would be enough information in the qBus sample alone to estimate a selectivity-corrected outcome model, $y_i = X_i \beta + \varepsilon_i$. We could implement either a standard binary-probit selection model or the six-level ordered-probit selection model we develop in this study. But in this case, there are no data in the qBus sample for y or the X variables. That information is available only for our eBird sample.⁸

Outcome Variable for eBird-Member Survey Sample

For the $j=1, \dots, J$ observations from our eBird member survey sample, we have Z_j sociodemographic and income variables that conform to the Z_i variables in the qBus sample, but we have no information about anyone for whom $CS6_j = 1$ or $CS6_j = 2$ (i.e., everyone in this sample is a member of eBird). In this case, the process of selection into eBird

⁵The Qualtrics Omnibus surveys have been discontinued, but there remain numerous other Omnibus options. See Appendix Table A3.

⁶We assume, in this proof-of-concept example, that respondents to the qBus questions are essentially a representative sample of the general population and respondents to the analogous questions posed to our eBird member survey are essentially a representative sample of eBird members.

⁷Practitioners may be aware that an adjustment could be necessary, to the intercept of the fitted propensity variable, depending on how the ordered probit algorithm has been parameterized.

⁸One could attempt to collect all the variables provided by our eBird member survey and its linked eBird CS observations from a large sample of respondents drawn from the general population. This would be impractical and duplicative, however, given the number of survey questions that would be required (and hence the cost of using a representative panel). The diary data in eBird also avoid the recall bias that would affect retrospective histories of birding activity.

membership cannot be modeled using the eBird data alone because there is no variation in the selection outcome for this group. However, we have data on an outcome variable of interest for this sample, y_j (in our illustration, the individual’s typical consideration-set radius—their maximum one-way distance for a regular one-day birding trip), along with a set of regressors, X_j , to explain this outcome, where none of this information is available for the qBus sample. Our eBird data for $CS6_j$, Z_j , y_j , and X_j can be summarized as

$$\begin{aligned}
 [CS6]_{J \times 1} &= \begin{bmatrix} 6 \\ \vdots \\ 5 \\ \vdots \\ 4 \\ \vdots \\ 3 \\ \vdots \end{bmatrix}, & [Z]_{J \times k} &= \begin{bmatrix} Z_{11} & \cdots & Z_{k1} \\ \vdots & & \vdots \\ Z_{1J} & \cdots & Z_{kJ} \end{bmatrix}, \\
 [y]_{J \times 1} &= \begin{bmatrix} y_1 \\ \vdots \\ y_J \end{bmatrix}, & [X]_{J \times m} &= \begin{bmatrix} X_{11} & \cdots & X_{m1} \\ \vdots & & \vdots \\ X_{1J} & \cdots & X_{mJ} \end{bmatrix} \quad [2]
 \end{aligned}$$

For the $j = 1, \dots, J$ observations in our eBird member survey sample, we assume the underlying population relationship between $CS6$ and the Z variables is identical to the analogous relationship in the qBus sample. If the complete six-level ordered probit-selection equation could be estimated for the $j = 1, \dots, J$ observations in the eBird member survey sample alone, the relevant pair of equations for our selection-correction model would be

$$\begin{aligned}
 CS_j^* &= Z_j \gamma + \eta_j \\
 y_j &= X_j \beta + \varepsilon_j \\
 (\eta_j, \varepsilon_j) &\sim BVN(0, 0, \sigma_\eta, \sigma_\varepsilon, \rho) \quad [3]
 \end{aligned}$$

Of course, this complete joint model cannot be estimated using our eBird member survey sample alone, because there are only eBird members in the $j = 1, \dots, J$ observations from that survey (i.e., there are no observations with $CS6_j = 1$ or $CS6_j = 2$).⁹

⁹For readers who want review the conventional Heckman two-step sample-selection correction procedure in more detail, we provide a summary in Appendix B.

Transferring a Fitted Selection Equation

Again, the challenge for selection-correction for our eBird member survey sample is that we do not have data for the y_i outcome variable and the X_i explanatory variables for people in the qBus sample who happen to be eBird members. We have these variables only for our completely independent sample of eBird members, where this sample allows linkages to extensive profile and birding-related data collected by eBird. If we can assume that participation in eBird among the general-population qBus sample follows the same data-generating process as the one that determines participation in eBird among people in our eBird member survey sample, perhaps we can assume that the underlying statistical relationship $(CS6^*, y) \sim BVN(Z\gamma, X\beta, 1, \sigma_\varepsilon, \rho)$ applies for both the $i = 1, \dots, N$ members of our qBus sample and for the $j = 1, \dots, J$ members of our eBird sample.¹⁰

The crux of this approach is that we use the $\hat{\gamma}^q$ estimates (from an ordered-probit model fitted to the qBus data on $CS6_i$ and Z_i) to construct a predicted index for each member of the eBird member survey sample, $Z_j \hat{\gamma}^q$, that can reflect a different mix of Z_j characteristics in our eBird member survey sample. We can use this predicted index in the selection-correction process for the eBird sample, even though we have no data from non-eBird members in the eBird member survey sample. With the bivariate normality assumption, the conditional expected value and variance for y_j will be calculated as follows, noting the j subscripts for the eBird data:¹¹

¹⁰Mechanically, it would be possible to pool our two samples and use the combined data set to estimate one common selection equation. The advantage of using the qBus sample alone for the selection equation is that the qBus data represent a random sample from the general population. Pooling it with the eBird sample, however, produces a data set that no longer represents the general population.

¹¹It is not uncommon for samples to have error distributions with different scales. Probit (and ordered-probit) models normalize their parameters on the error standard deviation for the model, so the estimated coefficients in the selection models we estimate using the qBus data are known only up to a scale factor. Each γ coefficient is implicitly γ^* / σ_η , where the σ_η applies to the qBus data. If the value of σ_η is larger or smaller for the eBird sample, employing the coefficients estimated on the qBus sample would lead to predicted engagement intensities in the eBird sample that are

$$\begin{aligned}
 E[y_j | y_j \text{ observed}] &= E[y_j | CS6_j^* > -Z_j \hat{\gamma}^q] \\
 &= X_j \beta + \rho \sigma_\varepsilon \lambda(-Z_j \hat{\gamma}^q) \\
 &= X_j \beta + \beta_\lambda \lambda(-Z_j \hat{\gamma}^q) \\
 \text{Var}[y_j | y_j \text{ observed}] &= \text{Var}[y_j | CS6_j^* > -Z_j \hat{\gamma}^q] \\
 &= \sigma_\varepsilon^2 [1 - \rho^2 \delta(-Z_j \hat{\gamma}^q)] \quad [4]
 \end{aligned}$$

The inverse Mills ratio (IMR), denoted as $\lambda(-Z_j \hat{\gamma}^q)$, is equal to $\phi(-Z_j \hat{\gamma}^q) / [1 - \Phi(-Z_j \hat{\gamma}^q)] = \phi(Z_j \hat{\gamma}^q) / \Phi(Z_j \hat{\gamma}^q)$, where $\phi(\cdot)$ is the standard normal probability density function (pdf) and $\Phi(\cdot)$ is the corresponding cumulative density function (cdf). The term $\delta(-Z_j \hat{\gamma}^q) = \lambda(-Z_j \hat{\gamma}^q) [\lambda(-Z_j \hat{\gamma}^q) + Z_j \hat{\gamma}^q]$. The desired unconditional (i.e., not systematically selected) expectation for y_j can be simulated, counterfactually, by setting $\rho = 0$, so that $E[y_j | y_j \text{ observed}] = X_j \beta$ and $\text{Var}[y_j | y_j \text{ observed}] = \sigma_\varepsilon^2$.

The calculated IMR selectivity-correction term, based on the $\hat{\gamma}^q$ ordered-probit estimates from the qBus sample and the Z_j variables from the eBird sample, can be appended to the list of regressors, X_j , in the outcome equation of interest for the eBird member survey sample. This is completely analogous to the standard binary-probit two-step selectivity correction.¹²

Ad Hoc Alternative: Interactions with Demeaned Propensities

In lieu of a formally derived Heckman-type selection-correction model, an alternative ad hoc approach can be used. We use the

biased proportionately downward or upward, respectively. Joint estimation using the two samples is feasible in principle but prohibitively difficult in the current case because of the strategy we must use to deal with missing variable values, discussed below. Here we assume the qBus and eBird selection-equation error distributions are identical.

¹²It will be appropriate in future research to graduate to full-information maximum likelihood joint estimation of the selection equation and the outcome equation. The ordered-probit form for the selection model is atypical, so no packaged algorithms exist to permit FIML estimation of a selection-on-ordered-probit model. We note that there is a packaged algorithm for ordered probit models with selection, but this is not what we need. That model has a conventional binary-probit selection model and an outcome equation that may be estimated as an ordered probit. In Appendix C, we provide a detailed discussion of the types of outcome models where it may be appropriate to contemplate adding an IMR term to correct for sample selection.

estimated engagement propensity model from our first-stage selection model to calculate fitted propensities to engage with eBird at six levels (in the qBus sample) or predicted propensities to engage with eBird at four levels (in the eBird member survey sample). For any individual in the eBird member survey sample with a given set of X_j variables, their “predicted engagement propensity” can be used just like any other variable that controls for individual-specific heterogeneity, such as indicators for gender, age, employment status, or educational attainment.

In a true random sample from the general population, every individual in the population is equally likely to show up on the sample. If we treat our qBus sample as representative of the general population, the predicted engagement intensities for our eBird member survey sample can be demeaned relative to the average fitted engagement propensity for the general-population qBus sample. This fitted demeaned engagement propensity variable can then be allowed to shift all the β parameters in the outcome model. After estimation, this demeaned response propensity can be counterfactually set to zero, effectively dropping all the interaction terms in which it is involved. The resulting outcome equation, without these interaction terms, then applies (in principle) to the case where everyone in the estimating sample shares an engagement propensity equal to the average engagement propensity in the general-population qBus data—namely, for a “representative” sample.¹³

3. Selection Model: Engagement Intensity

Available Variables for Selection Model

Our selection equation requires conformable measures of the Z_i and Z_j variables (i.e., these variables must be measured in the same way

¹³For identification, one or more exogenous explanatory variables need to be included in the predicted $Z_j \hat{\gamma}^q$ index that yields the fitted engagement propensities but excluded from the $X_j \beta$ index that represents the conditional expected value of the outcome variable of interest.

for the qBus and eBird member survey data sets). For the qBus data, unless one wishes to pay for additional questions, it is necessary to make do with the default set of sociodemographic and geographic characteristics that are included for all qBus panelists. Thus, we aggregate the Z variables for both the qBus and the eBird member survey to the same level. This yields conformable sets of indicator variables for the different levels of each of seven individual characteristics that can be allowed to influence the different intensities of engagement in our ordered-probit selection equation.

The available variables for our selection model, conformably aggregated across our two samples, are as follows (see Table 1 for additional details):

- Annual birding excursions of more than one mile from home (12 bins).
- Whether the individual has participated in the Audubon Christmas Bird Count (0/1).
- Whether the individual also hunts birds (0/1).
- Sociodemographics: Gender = female (0/1); Age (6 brackets); Race (4 groups); Ethnicity (2 groups); Income (5 brackets); Geography (4 regions); Employment status (5 categories); Educational attainment (5 levels).

Across observations with no missing values, for the qBus data ($N = 4,161$) and for the eBird member survey data ($J = 1,081$), Table 1 summarizes the proportions of observations in each set of indicator variables. Note that respondents in the general-population qBus sample have two more response options than respondents in the eBird member survey sample. The qBus respondents can also choose the engagement categories “Unfamiliar with eBird CS project” or “Heard of eBird but not a member.” Consequently, it is not possible to compare directly the proportions in the other four eBird-member engagement-intensity categories across the qBus and eBird samples. However, if we calculate the simple qBus conditional distribution solely for engagement levels 3–6 (where a qBus respondent is at least a member of eBird), then the relative frequencies for

Table 1
Descriptive Statistics (Proportions) for Variables in First-Stage Engagement-Intensity Models

	qBus Sample Proportions	eBird Sample Proportions
Engagement data available	1.000	1.000
1=Unfamiliar with eBird CS project	0.802	0.000
2=Heard of eBird but not a member	0.083	0.000
3=eBird member, but report rarely	0.031	0.391
4=eBird member, report <1/2 of birds	0.029	0.280
5=eBird member, report >1/2 of birds	0.030	0.177
6=eBird member, report almost all birds	0.024	0.152
Travel 1+ mile data available	0.442	0.769
Trips 1+ miles = 0	0.348	0.277
Trips 1+ miles = (1,4)	0.063	0.113
Trips 1+ miles = (4,7)	0.065	0.065
Trips 1+ miles = (7,10)	0.048	0.025
Trips 1+ miles = (10,21)	0.076	0.093
Trips 1+ miles = (21,41)	0.065	0.065
Trips 1+ miles = (41,72)	0.063	0.078
Trips 1+ miles = (72,124)	0.065	0.067
Trips 1+ miles = (124,174)	0.063	0.052
Trips 1+ miles = (174,238)	0.063	0.032
Trips 1+ miles = (238,364)	0.062	0.054
Trips 1+ miles = 365	0.017	0.078
Audubon CBC data available	1.000	1.000
Has participated in CBC	0.092	0.528
Bird-hunting data available	1.000	1.000
Hunts birds	0.224	0.073
Gender data available	1.000	0.994
Gender: male	0.489	0.427
Gender: female	0.511	0.573
Age data available	1.000	0.993
Age: ≤24 years	0.125	0.018
Age: 25–34 years	0.224	0.065
Age: 35–44 years	0.196	0.089
Age: 45–54 years	0.135	0.146
Age: 55–64 years	0.175	0.311
Age: ≥65 years	0.145	0.370

(table continued on following page)

Table 1

Descriptive Statistics (Proportions) for Variables
in First-Stage Engagement-Intensity Models
(continued)

	qBus Sample Proportions	eBird Sample Proportions
Income data available	1.000	0.804
Income: <\$25K	0.179	0.072
Income: \$25K–\$50K	0.219	0.203
Income: \$50K–\$75K	0.189	0.231
Income: \$75K–\$100K	0.141	0.173
Income: ≥\$100K	0.272	0.321
Region data available	1.000	1.000
Region: West	0.225	1.000
Region: Northeast	0.186	0.000
Region: Midwest	0.217	0.000
Region: South	0.372	0.000
Employment status data available	1.000	0.849
Employment status: full time	0.473	0.359
Employment status: part time	0.132	0.080
Employment status: looking for work	0.057	0.008
Employment status: unemployed	0.145	0.066
Employment status: retired	0.193	0.487
Education data available	1.000	0.976
Education: high school	0.226	0.036
Education: some college	0.356	0.158
Education: college grad	0.263	0.288
Education: master's degree	0.118	0.396
Education: doctoral degree	0.038	0.121
Observations	4,161	1,081

Note: Availability indicators are proportions of the total sample; group shares are proportions of the available data.

engagement levels 3–6 (proportion in qBus, proportion in eBird) for these four engagement intensities are (0.273, 0.398), (0.252, 0.275), (0.265, 0.179), and (0.210, 0.146). Although these (marginal) relative proportions differ within each pair, it is also possible that the types of people who respond to the qBus survey may differ from the types of people who are enrolled in eBird and responded to our survey of a random sample drawn only from eBird members.

Estimation Results for Selection Model

Ordered-Probit qBus Propensities to Engage with eBird

The qBus sample has virtually complete data for its Z_i variables (other than the annual number of days with birding trips more than one mile from home). This completeness stems from the fact that the standard demographic variables in our selection model are part of the “profile” data supplied for each qBus panelist, rather than being information we elicited via our questions. There are considerably more missing values for the Z_j variables from our eBird member survey, since all the socio-demographic information for that sample was collected during our survey, rather than being part of a standard profile. At least one relevant Z_j variable value is missing for 509 of the 1,081 respondents to the eBird member survey.

Our approach for dealing with these missing values is to transfer from the qBus sample, to each respondent in the eBird member survey sample, the richest possible specification of the ordered-probit selection model given the nonmissing data for that particular eBird respondent.¹⁴ To accommodate all of the patterns of missing values encountered in our eBird member survey data, we must estimate ordered-probit specifications with 30 different combinations of explanatory variables using the qBus data (as documented in [Appendixes E and F](#)). An analogous set of 30 ordered-probit models, but this time with just four engagement-intensity levels, can be estimated using the eBird member survey sample. These eBird ordered-probit models are required solely for the construction of our heterogeneous population weights, the discussion of which is included in the [Appendixes](#).¹⁵

¹⁴In an ideal world, all respondents would answer all questions in the survey and then only a single ordered-probit specification would be necessary. We could just estimate a simpler model based on the subset of data available for all respondents but this selection (based on item non-response) could introduce further non-representativeness.

¹⁵See Appendix G for the models using eBird data. Our heterogeneous population weights for the eBird member survey sample are described in Appendix H. Note that for the qBus models, “cut5” is the threshold between engagement levels 5 and 6, whereas for the eBird member survey models, “cut3” is the corresponding threshold between these

To illustrate just one of the 30 corresponding engagement-intensity models for the two samples, Table 2 presents the most complete specification that can be estimated using both the qBus and the eBird member survey samples. The complete set of Z_j variables is available for only 572 of the 1,081 eBird member survey respondents.¹⁶

To predict participation intensities for respondents to our eBird member survey, we use the $\hat{\gamma}^q$ coefficients estimated from the qBus specification, such as the first set of results in Table 2 (or the relevant version of the rest of the 30 models, to match the pattern of item nonresponse for each observation in the eBird member survey sample). We use these parameter estimates to calculate four predicted engagement-level probabilities (for construction of our weights) and predicted engagement intensities and predicted IMR terms to be used for sample-selection corrections in outcome equations that rely on only the eBird member survey data. The second set of results in Table 2 is estimated using the eBird member survey data alone. Again, we need these eBird ordered-probit models only to calculate fitted engagement-level probabilities in the eBird member survey sample, an ingredient in our heterogeneous sampling weights described in [Appendix H](#).¹⁷

One of the key innovations is the specification of this sample-selection model where the selection equation is an ordered-probit model. Of course, a binary-probit selection equation could be estimated and used in an analogous

manner, although it contains less information.¹⁸

In comparing the coefficient estimates for each sample in Table 2, we note numerous differences. These differences do not imply that it is inappropriate to transfer our qBus estimates to the eBird member survey sample for use in our selection-correction procedures. The eBird member survey sample is also a selected sample for these ordered-probit engagement-intensity models. The qBus model covers all six engagement-level propensities, including the roughly 88% of the qBus general-population sample who are not eBird members. In transferring the qBus propensity parameters to our eBird member survey sample, it is imperative to preserve the influence of the first two, non-eBird member engagement levels in our general-population qBus data.

Consider the signs and significance of the individual coefficient estimates in Table 2. For respondents who report having traveled at least one mile from home to see birds over the past year, the more days a year a respondent has made such a trip, the greater their propensity to engage with eBird. These effects are statistically significant only in the eBird member survey sample, however. Past participation in the Audubon Christmas Bird Count increases engagement propensity in the qBus sample, but this effect is not apparent in the eBird member survey sample. Whether the respondent also hunts birds has no discernible effect on eBird engagement intensity in either sample, although the point estimate is positive for the qBus sample and negative in the eBird member survey sample.

Female qBus respondents have statistically lower eBird engagement intensities than males, but the same is not true for women in the eBird member survey sample. Individuals who are less than 44 years old have higher propensities to engage with eBird, with the largest effect for eBird members 24 years old

levels. A simple change-of-origin restores comparability. The *polr* package in R handles ordered-probit models, with its “method” argument set to “probit.”

¹⁶At the other end of the spectrum of data completeness in the eBird member survey sample, Appendix E contains an analogous table for the largest model that can be estimated for every respondent in the eBird member survey sample without being limited by missing data. This selection model can use all 1,081 eBird member survey respondents who answered the question about our outcome variable of interest, but must employ far fewer explanatory variables Z_j .

¹⁷To construct our weights to be applied to each observation in the eBird member survey sample, we require engagement-level probabilities for our eBird sample that are (a) “expected,” that is, predicted, based on parameter estimates transferred from the qBus sample, and (b) “observed” (i.e., fitted, based on parameter estimates directly from the eBird sample alone).

¹⁸Appendix I digresses to explore a variety of the intermediate components of our models. E.g., in terms of in-sample fitted engagement propensities, our ordered-probit selection model tracks the conventional binary-probit selection specification closely when each is applied to the same sample of qBus respondents (although the ordered-probit model predicts somewhat greater propensities at the low end of the range).

Table 2
 Estimated Coefficients, Ordered-Probit Engagement-Level Models
 Examples with Maximum Heterogeneity

	Ordered Probit qBus Data		Ordered Probit eBird Data	
Travel 1+ mile data available	0.136	(0.661)	— ^a	
Trips 1+ miles = 0	-0.870	(0.667)	-2.828***	(0.228)
Trips 1+ miles = [1,4)	-0.260	(0.678)	-3.113***	(0.268)
Trips 1+ miles = [4,7)	-0.554	(0.683)	-1.862***	(0.287)
Trips 1+ miles = [7,10)	-0.388	(0.681)	-2.243***	(0.372)
Trips 1+ miles = [10,21)	-0.0566	(0.667)	-1.820***	(0.250)
Trips 1+ miles = [21,41)	-0.0189	(0.668)	-1.496***	(0.262)
Trips 1+ miles = [41,72)	0.287	(0.665)	-1.354***	(0.254)
Trips 1+ miles = [72,124)	0.518	(0.663)	-0.711***	(0.258)
Trips 1+ miles = [124,174)	0.400	(0.663)	-0.644**	(0.293)
Trips 1+ miles = [174,238)	0.487	(0.662)	-0.582*	(0.321)
Trips 1+ miles = [238,364)	0.730	(0.661)	-0.422	(0.286)
Trips 1+ miles = 365	0.699	(0.687)	— ^b	
Has participated in CBC	1.916***	(0.0706)	0.170	(0.107)
Hunts birds	0.0640	(0.0965)	-0.0767	(0.181)
Gender: female	-0.169***	(0.0509)	-0.111	(0.107)
<i>Relative to Omitted Category: 45–54 years</i>				
Age: ≤24 years	0.539***	(0.0991)	0.994**	(0.424)
Age: 25–34 years	0.545***	(0.0871)	0.293	(0.216)
Age: 35–44 years	0.360***	(0.0894)	0.325*	(0.192)
Age: 55–64 years	-0.207*	(0.110)	-0.117	(0.159)
Age: ≥65 years	-0.298**	(0.134)	0.146	(0.201)
<i>Relative to Omitted Category: \$50K–\$75K</i>				
Income: <\$25K	-0.0590	(0.0853)	-0.0751	(0.254)
Income: \$25K–\$50K	-0.0192	(0.0774)	0.143	(0.150)
Income: \$75K–\$100K	-0.0423	(0.0867)	-0.0150	(0.158)
Income: ≥\$100K	-0.0118	(0.0784)	0.143	(0.138)
<i>Relative to Omitted Category: West</i>				
Region: Northeast	0.164**	(0.0737)	— ^b	
Region: Midwest	-0.0186	(0.0754)	— ^b	
Region: South	0.0552	(0.0660)	— ^b	
<i>Relative to Omitted Category: Full Time</i>				
Empl. status: part time	0.0409	(0.0751)	-0.148	(0.188)
Empl. status: looking for work	-0.173	(0.108)	-0.819	(0.650)
Empl. status: unemployed	-0.109	(0.0805)	-0.0110	(0.217)
Empl. status: retired	-0.148	(0.107)	-0.325**	(0.163)
<i>Relative to Omitted category: 4-Year College Degree</i>				
Education: high school	-0.0294	(0.0754)	0.693**	(0.309)
Education: some college	-0.128*	(0.0671)	0.00246	(0.170)
Education: master's degree	0.269***	(0.0826)	0.259**	(0.121)
Education: doctoral degree	0.165	(0.124)	0.00792	(0.165)
<i>Ordered-Probit Thresholds</i>				
Cut1	1.279***	(0.113)	-2.651***	(0.288)
Cut2	1.895***	(0.116)	-1.309***	(0.279)
Cut3	2.258***	(0.119)	-0.363	(0.274)
Cut4	2.690***	(0.123)	— ^c	
Cut5	3.341***	(0.132)	— ^c	

(table continued on following page)

Table 2
Estimated Coefficients, Ordered-Probit Engagement-Level Models
Examples with Maximum Heterogeneity (*continued*)

	Ordered Probit qBus Data	Ordered Probit eBird Data
Observations	4,161	572
Max. log likelihood	-2,390.32	-582.44

Note: The first model is a six-level model using the full qBus sample. The second is a four-level model for the subset of 572 eBird survey respondents with complete data for this least-restricted specification. Specifications with fewer explanatory variables can use more observations in the eBird data set. Other models to accommodate all of the patterns of missing variables in the eBird data are in the [Appendixes](#).

^aToo few missing values to identify this coefficient.

^bThese indicators are zero for all observations in the eBird data.

^cThere are only four levels of engagement in the eBird data.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

or younger. Older respondents in the qBus survey have significantly lower eBird engagement propensities.

Income, aggregated into five brackets, does not appear to influence eBird engagement propensity in either sample. However, engagement propensities are statistically significantly higher in the Northeast region of the United States than elsewhere. In the qBus sample, employment status seems to have no effect on eBird engagement propensities, but in the eBird member survey sample, being retired (as opposed to being employed full time, the omitted category) decreases eBird engagement propensities (where these estimates control for age group and annual frequencies of trips of more than one mile to see birds).

In the qBus sample, compared to individuals with a four-year college degree (the omitted category), those with only some college have lower engagement propensities. For both samples, having a master's degree increases eBird engagement propensity.¹⁹

Transferring qBus Selection Model to eBird Member Survey Sample

We use the assumption that for each individual in our eBird member survey sample, we can transfer the relevant set of $\hat{\gamma}^q$ parameters

¹⁹In our eBird member survey sample, only about 3.5% of respondents have just a high school education or less, so perhaps not much ought to be read into the statistically significantly positive effect of lower educational attainment on eBird participation propensities in the eBird member survey sample (where respondents were required to be 18 years or older). If school-based eBird assignments recruit 18-year-olds still in high school, this could account for the greater eBird engagement propensities in this group.

estimated using the qBus data. The qBus selection model to be transferred needs to be estimated using the same set of nonmissing regressors, so that we have exactly the necessary information to calculate a predicted propensity index, $Z_j \hat{\gamma}^q$, that exploits as much information as we have about that individual eBird member's sociodemographic characteristics.²⁰

Figure 1 displays smoothed densities for the marginal distributions across the relevant sample (i.e., the degree of heterogeneity) across respondents in the fitted (or predicted) probabilities of being at each of the four engagement levels (3, 4, 5, and 6), conditional on the individual being a member of eBird. Panel A shows the fitted individual probabilities of being at each engagement level for the qBus sample. Panel B shows the same for the eBird member survey sample. Panel C shows the predicted probabilities of being at each engagement level for the eBird member survey sample, calculated by transferring the parameters of the relevant ordered probit model estimated using the qBus sample.

4. Outcome Model: Consideration-Set Radii

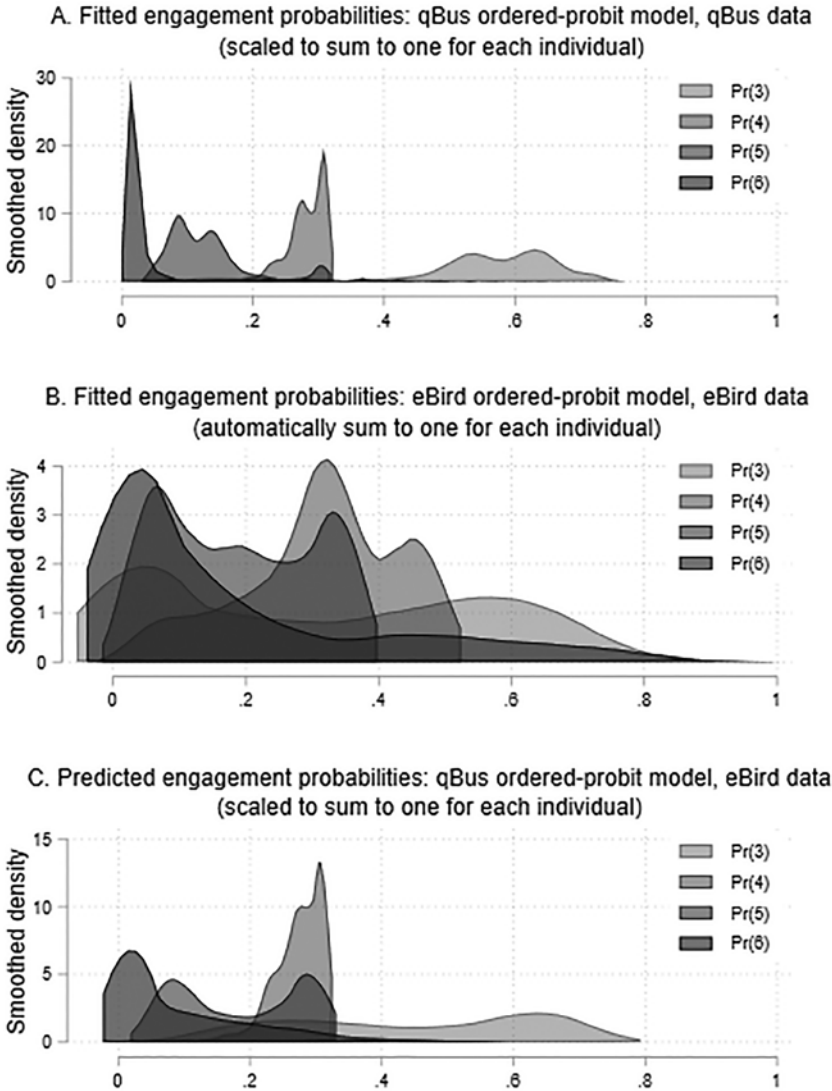
Available Variables for Outcome Model

This section illustrates the use of our predicted, rather than estimated, IMR terms in

²⁰Appendix I also includes a comparison of the predicted inverse Mills ratios based on the binary-probit and ordered-probit selection models when the parameters estimated for the qBus sample are transferred to the eBird member survey sample.

Figure 1

Predicted Probabilities, Engagement Levels 3–6. A: qBus Sample; B: eBird Sample; C: eBird Sample Based on an Ordered-Probit Model Estimated Using the qBus Sample



a model that explains the maximum distance that people state they would typically consider traveling for a one-day birding excursion. This model is estimated using only our eBird member survey sample. As noted in the introduction, consideration sets for destination-choice models are related to the concepts of market extent, trip-generating functions, and distance decay. The summary statistics for the eBird-only data available for these models are given in Table 3.

The variables available to use as regressors in our consideration-set-radius model are different from those used in our ordered-probit models to explain levels of eBird engagement intensity. For our engagement-intensity models, we were limited to variables that were available and could be measured conformably for the qBus sample and the eBird member survey sample, given our need to perform a “model transfer.” We have richer data from the eBird member survey that was not

Table 3
Descriptive Statistics

	Mean	Std. Dev.
<i>Dependent Variable (Consideration-Set-Radius, Lower Bound of Chosen Interval)</i>		
Self-reported maximum radius (i.e., extent) of travel in miles	83.28	58.10
<i>Explanatory Variable</i>		
Employment status: employed	0.373	0.484
Income data available	0.804	0.397
Income in 10k, if reported	7.011	5.891
Gender: female	0.570	0.495
Age: <45 years	0.171	0.377
Age: >64 years	0.367	0.482
Education: graduate school	0.505	0.500
No interest: perching birds	0.057	0.233
No interest: other game birds	0.111	0.314
<i>Selection-Correction Options</i>		
Binary-probit IMR	2.035	1.212
Adjusted ordered probit IMR	1.668	0.978
Engagement propensity (demeaned using qBus mean)	0.657	1.261
Observations	1,081	

Note: Variables are for the outcome model, elicited from the eBird member survey sample for $n = 1,081$ respondents who answered the question about maximum one-way distance for a birding day trip.

available in the qBus sample. For example, our eBird member survey elicits income in much finer brackets than we could use in the engagement-intensity models, so we convert the income bracket data into an approximate continuous income variable.²¹

We also take advantage of our eBird member survey data concerning eBirders' interests in different species categories. For various categories of bird species, 6%–11% of eBirders report that they have no interest in that category. The least popular category in our eBird member sample is “game birds other than waterfowl” (e.g., pheasants, turkeys, grouse, partridges). This information about the goals of individual birders ties our analysis to the notions explored in Swait, Franceschinis, and Thiene (2020), who find that benefit variations

²¹ Given the vagaries of just what income measure is the relevant measure of income, we simply use the midpoint of each income bracket. For the under \$20,000 interval, we arbitrarily assign \$18,000. For the over \$200,000 interval, we arbitrarily assign \$225,000.

associated with distance depend on people's goals in their recreational pursuits.

In the specific context of birding, the question of the relevant consideration sets for birders has bearing on the potential “active use” versus “passive use” (option, bequest, or existence) values of environmental projects to protect or enhance local wild bird populations. It is likewise relevant to calculation of the welfare effects of wholesale shifts in the geographic ranges of different bird species in response to climate change. (Birds are highly mobile and are likely to relocate more quickly than are most birdwatchers.)

Estimation Results for the Outcome Model

Our dependent variable for these models, the maximum distance considered for a typical one-day birding trip, is elicited in distance brackets in our eBird survey. The exact wording of the question is: “If you are NOT making a special trip to try to see a reported rare bird, what is the greatest distance you would consider traveling, one way, for a regular single-day birding trip?” The lowest category is “10 miles or less,” so no answers of exactly zero are observed. A reasonable estimation method assumes that the latent continuous dependent variable is conditionally log-normally distributed. An interval data regression model can then be estimated by maximum likelihood methods.²²

Model 1 in Table 4 is a naive specification to explain consideration-set radius (maximum willingness to travel to see birds) with no corrections. Model 2 is an otherwise naive specification that uses only our constructed weights, as detailed in [Appendix H](#). Models 3–5 continue to employ these weights. Model 3 includes an IMR variable based on a conventional binary-probit selection equation, and model 4 uses our novel ordered-probit

²² Stata's *intreg* estimator is available for such models. The *survival* package in R, with its *survreg* function, would appear to handle similar interval-data regressions, with its “dist” argument set to “gaussian.” Our survey also includes the question: “If you travel more than one or two miles from home to go birding, what is your most frequent mode of travel for these birding trips?” Less than 6% of respondents selected the answer: “I never travel more than one or two miles for birding.” Still, this admits for shorter excursions.

Table 4
 Consideration-Set-Radius Models without and with Engagement-Intensity
 Weights and Either Sample-Selection Corrections or Interactions between
 All Regressors and Demeaned Ordered-Probit Selection Propensities

Model	(1) Naive	(2) Weights Only	(3) Binary Probit IMR	(4) Ordered Probit IMR	(5) Demeaned Propensity
<i>Main Variable</i>					
Empl. Status: employed	-0.0351 (0.0696)	-0.0613 (0.0973)	-0.0872 (0.0943)	-0.0934 (0.0936)	-0.299* (0.154)
Income data available	-0.429*** (0.127)	-0.362** (0.180)	-0.342* (0.176)	-0.354** (0.175)	-0.223 (0.186)
ln(Income in 10K, if reported)	0.218*** (0.0533)	0.171** (0.0730)	0.156** (0.0718)	0.158** (0.0710)	0.0628 (0.0887)
Gender: female	-0.127** (0.0590)	-0.123 (0.0796)	-0.0374 (0.0787)	-0.0410 (0.0783)	0.000944 (0.106)
Age: <45 years	0.185** (0.0835)	0.225** (0.112)	0.167 (0.106)	0.137 (0.108)	-0.317 (0.197)
Age: >64 years	-0.0489 (0.0713)	-0.114 (0.0982)	-0.0287 (0.100)	-0.103 (0.0961)	-0.0110 (0.115)
Education: graduate school	0.113* (0.0598)	0.0132 (0.0751)	-0.0649 (0.0741)	-0.0642 (0.0748)	-0.259*** (0.0958)
No interest: perching birds	-0.811*** (0.148)	-0.862*** (0.198)	-0.847*** (0.205)	-0.829*** (0.202)	-0.749*** (0.190)
No interest: other game birds	-0.634*** (0.107)	-0.660*** (0.130)	-0.577*** (0.131)	-0.578*** (0.130)	-0.549*** (0.127)
Constant	4.066*** (0.0972)	4.170*** (0.127)	4.494*** (0.135)	4.403*** (0.131)	4.102*** (0.140)
<i>Selection-Correction Strategy</i>					
Binary-probit IMR			-0.169*** (0.0305)		
Ordered probit IMR				-0.386*** (0.0689)	
<i>Interaction between Main Variables and Demeaned Engagement Propensity (Demeaned Using qBus Sample Mean)</i>					
Empl. Status: employed × engagement prop. (demeaned)					0.185 (0.158)
Income data avail. × engagement prop. (demeaned)					-0.120 (0.0756)
ln(Income in 10K, if reported) × engagement prop. (demeaned)					0.134* (0.0692)
Gender: female × engagement prop. (demeaned)					0.224* (0.117)
Age: <45 years × engagement prop. (demeaned)					-0.00451 (0.0733)
Age: >64 years × engagement prop. (demeaned)					-0.0223 (0.0607)
Education: graduate school × engagement prop. (demeaned)					0.0176 (0.0551)
No interest: perching birds × engagement prop. (demeaned)					-0.231 (0.175)
No interest: other game birds × engagement prop. (demeaned)					0.160 (0.107)
Constant × engagement prop. (demeaned)					0.282*** (0.0813)
$\ln(\sigma_e)$	-0.0803*** (0.0237)	-0.0583* (0.0306)	-0.0788** (0.0313)	-0.0808*** (0.0310)	-0.105*** (0.0310)

(table continued on following page)

Table 4
 Consideration-Set-Radius Models without and with Engagement-Intensity
 Weights and Either Sample-Selection Corrections or Interactions between
 All Regressors and Demeaned Ordered-Probit Selection Propensities (*continued*)

Model	(1) Naive	(2) Weights Only	(3) Binary Probit IMR	(4) Ordered Probit IMR	(5) Demeaned Propensity
Observations	1,081	1,081	1,081	1,081	1,081
Log likelihood	-2,411.41	-2,429.96	-2,409.82	-2,408.03	-2,382.37
AIC	4,844.83	4,881.92	4,843.64	4,840.05	4,806.74
BIC	4,899.67	4,936.77	4,903.47	4,899.88	4,911.44
Weighted?	No	Yes	Yes	Yes	Yes

Note: Standard errors are in parentheses. Dependent variable: log of maximum one-way distance willingly traveled on a typical birdwatching day trip.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

selection equation. Finally, model 5 shows the results from our ad hoc selection-correction strategy that interacts each main determinant of consideration-set radius extent with a demeaned predicted engagement propensity (based on our adjusted ordered-probit selection specification estimated on the qBus sample and transferred to the eBird member survey sample).²³

For the models in Table 4, our explanatory variables include whether the eBird member is currently employed, whether they were willing to report their income in our eBird survey, the level of that income, their gender, their membership in three broad age brackets and two educational attainment categories, and whether they specifically express no interest in each of two categories of bird species.²⁴

Model 2 employs our calculated sample weights—based on relative fitted engagement-level probabilities in the general population, as opposed to this eBird member survey sample. Recall that women represent about 57% of the eBird sample, but only 51% of the qBus general-population sample. The only notable difference in the estimates, with the inclusion of weights, is that the coefficient on the female indicator, which was negative and statistically significant at the 5% level in

the unweighted model, becomes statistically insignificant in all other specifications. Given this difference, we retain these weights in subsequent specifications and consider the ways in which the results for models 3, 4, and 5 are different from those for models 1 and 2.

IMR Coefficients

Models 3 and 4 in Table 4 are the two IMR-based selection-corrected models that rely on the strong assumption of bivariate normal errors for the latent engagement-intensity variable and the interval-censored outcome variable. The coefficient of interest is that on the relevant fitted IMR. In two-stage methods, this coefficient is the estimate of $\rho\sigma_\varepsilon = \beta\lambda$. Given that the error standard deviation, σ_ε , must be positive, the sign of this compound parameter implies the sign of ρ , the correlation between the errors in the selection and outcome equations.

Our negative IMR coefficients in models 3 and 4 imply that unobserved factors that make a respondent more likely to be intensely engaged with eBird also make them willing to travel less far on a typical one-day birding trip. We must acknowledge that models 3 and 4 treat these second-stage predicted IMR variables as nonstochastic (thereby understating the amount of noise in the model). Nevertheless, these negative IMR coefficients are strongly statistically significantly different from zero.²⁵

²³ Ad hoc correction specifications like model 5 are potentially helpful in contexts where the error term in the outcome equation does not have an explicit (or at least an underlying) normal distribution—as is the case with destination choice models estimated using the standard random utility method.

²⁴ Our survey elicited levels of interest in five categories of species, but for only two categories do interest levels have statistically significant effects on consideration-set radii.

²⁵ FIML estimation of the joint model for engagement propensity and consideration-set radii could remedy this

We had expected (if anything) that the propensity to participate in eBird would be positively associated with a respondent's consideration-set radius, since latent birding avidity could be an important omitted variable. It is somewhat counterintuitive to find negative estimates for the IMR coefficients. Perhaps the relevant unobserved heterogeneity includes the opportunity cost of time (or unobserved age-related technical sophistication in using the online eBird app, or for online surveys in general).

Employment status. None of models 1–4 in Table 4 suggest that employment status has a statistically significant effect on consideration-set radii. However, model 5, using our ad hoc correction of interacting each regressor with the demeaned engagement propensity, suggests that consideration-set radii for the general population are smaller by about 30% if the respondent is currently employed. Employed people are likely to have less free time for all leisure activities, including birding excursions.

Income data availability indicator and level of income, if known. For all but model 5, compared with respondents who decline to provide such data in the eBird member survey, those who do provide income data report consideration-set radii that are smaller by about 34%–43%. However, this negative effect is offset by the positive effect of income (when reported) on consideration-set radius—a 1% higher income corresponds to a radius that is larger by about 0.15%–0.22%. Higher-income respondents likely have less-binding budget constraints for travel expenses.²⁶ Model 5, though, suggests that income has no statistically discernible effect on consideration-set radii in the general population.

Gender. The point estimate for the effect of being female on consideration-set radius

is negative in models 1–4 (although the estimate is statistically significantly negative only in model 1). The estimated effect of gender changes sign in model 5, but remains insignificant, suggesting that gender has no effect on consideration-set radii in the general population.

Age. Models 1 and 2, which do not correct for systematic selection, suggest that being less than 45 years old (compared with the omitted category of 45–64 years old) is associated with a consideration-set radius that is larger by about 18%–23%, but this effect disappears in models 3–5, which explore alternative remedies for systematic selection.

Education. Relative to the omitted category with a college degree or less, model 1 implies that having attended at least some graduate school increases expected consideration-set radius by 11%, significant at the 10% level. Models 2–4 suggest that graduate school has no statistically significant effect on radius, but model 5 implies that in the general population, graduate school is associated with a strongly statistically significant 26% smaller radius.

Disinterest in particular categories of species. Across all five specifications, respondents to our eBird member survey who reveal that they are not interested in perching birds or not interested in “other game birds” (i.e., game birds other than waterfowl), have statistically significantly smaller consideration-set radii. The magnitudes of these effects are also similar across all specifications. Reporting a lack of interest in either of these categories of birds shrinks expected radius substantially, by 58%–86%.

Model 5's Interaction Terms

Among the selection-correction models, models 3 and 4 rely on strong assumptions of bivariate normal errors and slope coefficients that are identical in the eBird sample and the general population. Under these specific conditions, adding to the model a single (appropriate) IMR term, with an unrestricted coefficient, would yield slope coefficients for the other variables assumed not to be biased by sample selection. These IMR strategies can be described as structural approaches to sample-selection correction.

estimated-regressors problem and may be pursued in future applications.

²⁶ Compared with respondents who withhold their income data, the positive effect of greater income overcomes the negative effect associated with the provision of any income data when income reaches roughly \$21,400–26,200. Mean reported household income in the sample is about \$87,200, and the minimum reported income is \$18,000, so the effect of additional income on consideration-set radius is positive for most of the sample.

In contrast, model 5 is ad hoc, unstructured, and highly flexible. This approach makes a different but perhaps equally strong assumption—that each parameter of the outcome model varies linearly with the respondent's predicted propensity to engage with eBird, the latent continuous variable that drives people's intensive margin of participation in eBird at various engagement levels. The linear relationship between each estimated coefficient and the demeaned predicted engagement propensity may be positive or negative or statistically zero. The counterfactual we wish to simulate is the set of outcome model parameters that would obtain if everyone in the estimating sample shared the mean engagement propensity in the general population (i.e., the qBus sample).

Prior to estimation, we transformed each respondent's predicted engagement propensity (the $Z_j \hat{\gamma}^q$ "index") by taking its deviation from the population mean (i.e., from its mean in the qBus sample), the average value of $Z_i \hat{\gamma}^q$. In the population, the demeaned engagement propensity variable would be zero, but in Table 3, note that the average demeaned engagement propensity in the estimating sample is about 0.657. People from our eBird member survey do have a higher than average propensity to engage with eBird.

When we include in model 5 the interaction terms between each basic regressor and our demeaned engagement propensity variable, the coefficients on the noninteracted basic regressors can be interpreted as the simulated values of those coefficients at the mean engagement propensity in the population. In the bottom half of Table 4, for model 5, we show the estimated coefficients on the interaction terms, which reveal how the effects of each basic regressor vary with the individual's predicted engagement propensity.

The interaction terms in model 5 suggest that although income has no statistically discernible effect on consideration-set radius at the mean engagement propensity in the general population, the effect of income on this radius increases systematically with the respondent's predicted engagement propensity. There is a similar effect for being female. The most statistically significant interaction term in model 5, however, is the (implicit)

interaction between the demeaned engagement propensity variable and the intercept term in the basic specification (which is just the demeaned propensity variable itself). The strongly statistically significant positive coefficient on this term implies that expected consideration-set radius is larger as the demeaned engagement intensity increases. Given that demeaned engagement intensity in the estimating sample is positive, on average, the uncorrected eBird member survey sample overstates the consideration-set radii in the general population.

One relevant observation about the demeaned engagement propensity variable in model 5 is that the ordered-probit IMR is very close to being a linear transformation of this propensity variable over the relevant range in our data. The correlation between the two variables is -0.9955 . For corrected predictions about consideration-set radii in model 4, we eliminate the IMR term by setting its coefficient to zero (i.e., by assuming that ρ is zero, so that $\rho \sigma_\varepsilon = 0$). In model 5, if we were to include just the demeaned engagement propensity variable without its interactions with the basic regressors, we would set to zero the demeaned propensity variable itself to produce corrected predictions about consideration-set radii. Given the degree of correlation between the two variables, either correction would be expected to have about the same effect on the vector of coefficients on the basic variables. Thus, we can view model 5 as being, in effect, a generalization of model 4 with additional flexibility. Model 5 also permits all the slopes to differ, not just the intercept.

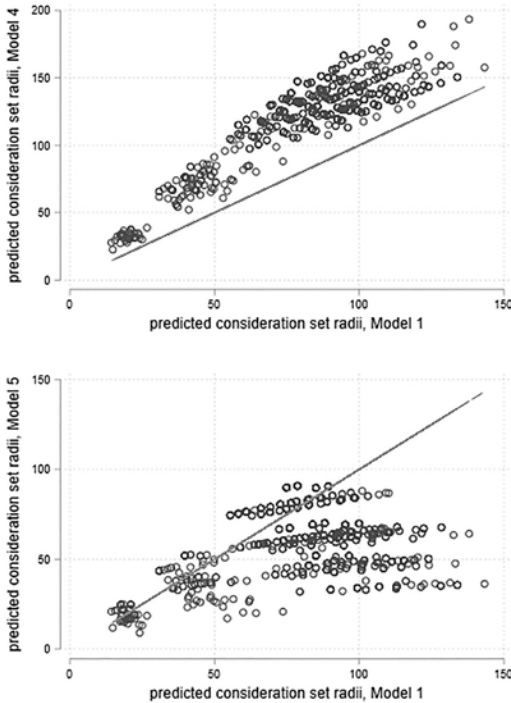
Predicted Values for the Outcome Variable with and without Corrections

The point of sample-selection correction is to adjust the statistical relationships observed in the selected sample to better reflect the general population. We can compare the predicted consideration-set radii for selected specifications.²⁷ The top graph in Figure 2 shows the

²⁷ Recall that the dependent variable in the specifications in Table 4 is in log form. Exponentiation of a fitted log value yields the median of the fitted level. One must multiply by the fitted value of $(\sigma_\varepsilon^2 / 2)$ to recover the mean of the fitted

Figure 2

Predicted Consideration-Set Radii for 1,081 Observations in the eBird Sample for Models 4 and 5 (in Table 4), Plotted against Predicted Radii from the Naive Model 1



predicted radii from model 4 in Table 4 plotted against the predicted radii for the same observations under the naive model 1 with no weighting or correction for sample selectivity. Model 4, with its ordered-probit IMR term, predicts consideration-set radii that are uniformly larger than those predicted by model 1. This difference arises because of the negative error correlation between the selection equation and the outcome equation, as implied by the negative coefficient on the IMR term. An individual who is more likely to show up in the eBird sample than their observed characteristics would predict also tends to have a smaller consideration-set radius than their observed characteristics would predict.

However, the effects of systematic selection on predicted consideration-set radii

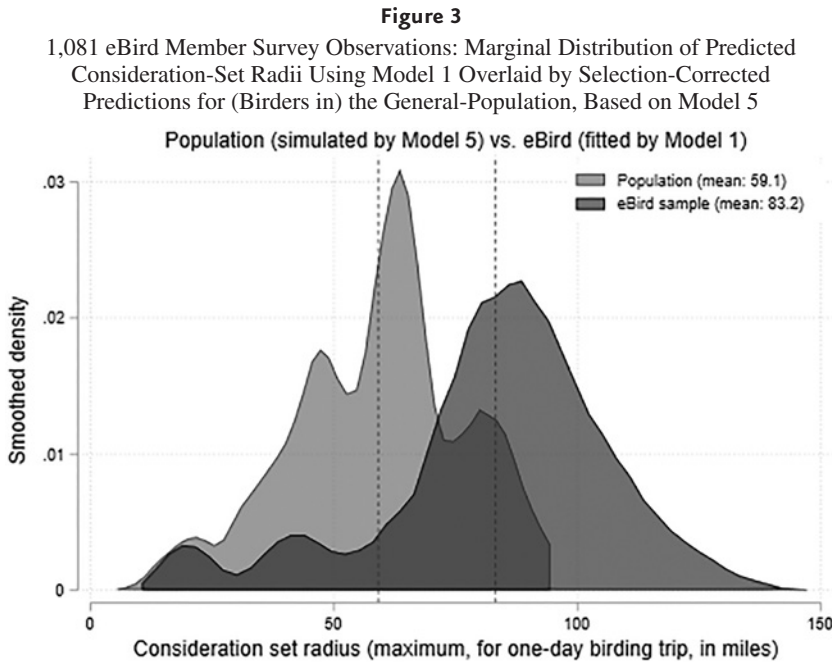
conditional distribution, due to the skewness of the implied log-normal distribution.

implied by model 4 are notably opposite from the effects implied by the results shown in the bottom graph in Figure 2. This second graph features the radii predicted by model 5 in Table 4, where each explanatory variable is also interacted with the demeaned engagement propensity variable. This demeaned propensity is then set to zero to simulate the expected consideration-set radius if everyone in the sample had a fitted engagement intensity equal to the average in the qBus sample. These fitted values plotted, likewise, against those for the naive model 1, show that the radii in the general population are smaller than they are in the selected sample of birding enthusiasts in the eBird sample.²⁸ Clearly, effect of the negative error correlation between the selection and outcome equations is more than offset by the heterogeneity in the slope coefficients that is a function of predicted engagement propensities.

Figure 3 compares the two marginal distributions of predicted consideration-set radii for birding trips with and without selection corrections. It is unsurprising that basing estimates of radii for one-day birding excursions on a sample of eBirders would likely overpredict the radii for such trips in a general-population sample with the same characteristics. However, the absence of a spike at zero in Figure 3 is also notable, given that roughly 12% of the general population does not report even incidental attention to wild birds over the past year. The absence of a point mass at zero for our eBird data on subjective consideration-set radii probably means that our corrected estimates cannot be scaled to 100% of the overall population. These radii will likely be relevant for a subset of the population.²⁹

²⁸ These predicted values show a pattern of clustering because the demeaned engagement propensity is a function mostly of indicator variables, and the interaction terms in the outcome model likewise involve indicator variables.

²⁹ Even a consideration-set radius of zero, for actual travel away from home to enjoy “active use” of wild birds, does not preclude the possibility of “passive use” values (option, bequest, or existence values) for wild birds in the region. Birds are also mobile. The presence of wild birds in any given radius will affect the probability that these birds may be viewed in one’s backyard, without the necessity of travel.



5. Conclusions and Recommendations

We intersect the sample-selection literature and the literatures on consideration sets for destination-choice models. Our goal is to augment the research toolkit for using CS data—with improved confidence that any derived insights are more suitable for scaling to the general population or for use in benefits transfer exercises. The two main tasks are to (1) illustrate some new sample-selection-correction techniques we have developed to allow for using data from auxiliary general-population surveys to correct for sample selection present in CS data and (2) model heterogeneity in the radius of consideration sets for regular birdwatching day trips in Oregon and Washington states as an illustration. Our contrasting results for the consideration-set radius model demonstrate that corrections for sample selection (and weighting for engagement intensity) may be important for scaling to the general population, or transferring to other contexts, any results derived from CS data.

The key takeaway from our illustrative consideration-set radius “outcome” model

is the potential importance of nonrandom selection into CS projects. Preferences in the general population are important if government agencies, for example, are to make good decisions about the efficient allocation of resources to protect wild birds, a public good. How to provide the appropriate amount of wild bird habitat is an increasingly relevant policy question. Changes in land cover and climate present significant threats to the declining wild bird populations documented in Rosenberg et al. (2019). Changes in bird populations affect birdwatcher welfare (as in Kolstoe, Cameron, and Wilsey 2018). To limit the loss of bird populations and bird biodiversity, multiple agencies at all levels of government will likely need to work together.

It is important to recognize—especially in the case of migratory species such as birds—that conservation-related actions in one location have the potential to affect outcomes at other locations. Existing programs, such as the National Wildlife Refuge System and the Urban Bird Treaty Program, make a good start but appear to have been insufficient, given that avian biodiversity remains a concern (especially given changes in land cover and the climate). Conservation solutions must account

for the fact that political jurisdictions may not align with the spatial “market extent” for non-market demands for conservation (as noted by Bakhtiari et al. 2018; Vogdrup-Schmidt et al. 2019). These market extents are related to the consideration-set radii of individual birders.

The need for a qBus-type sample to permit sample-selection corrections in this case highlights a potential supplementary role for broad-based surveys of birdwatching trip behavior and CS engagement. Information about trip-taking behavior has long been gathered by the U.S. Fish and Wildlife Service through their quinquennial general-population survey on Fishing, Hunting and Wildlife Watching for their national and state-level reports. For the 2016 survey, the U.S. Fish and Wildlife Service did not collect data at the state level; their survey information is reported primarily at the census division. Rockville Institute conducted a state-level survey.

National recreation surveys (e.g., FHWAR) in the future may consider adding a detailed question about participation in outdoor-based CS projects. A federal registry now documents more than 400 CS projects (www.citizenscience.gov), so general-population information on CS participation would benefit other agencies, such as the National Oceanic Atmospheric Administration (NOAA) or U.S. Geological Survey (USGS), which could also exploit data from CS projects on recreational behavior. Such CS projects include Watch for Whales (NOAA), Geocache for a Good Cause (NOAA), and Nature’s Notebook (USGS), for example.

To be most useful, existing general-population surveys may consider including questions about CS engagement in projects related to valued recreational services of ecosystems. This general-population engagement data would be a vital complement to surveys of CS project members to help researchers understand active and passive use values for a wide range of environmental public goods. Without general-population information, it will continue to be very difficult to scale to the general population any empirical findings based solely on surveys fielded to convenience samples of CS participants.

Acknowledgments

Much of this research was conducted while Cameron was the R. F. Mikesell Professor of Environmental and Resource Economics at the University of Oregon (until June 2021) and while Kolstoe was Assistant Professor of Economics, Department of Economics and Finance and Department of Environmental Studies, Salisbury University (until January 2021). Steve Kelling at the Cornell University Laboratory of Ornithology generously facilitated our survey of eBird members. This work has been supported in part by the endowment accompanying the Raymond F. Mikesell Chair in Environmental and Resource Economics at the University of Oregon. We are grateful for comments from participants at the Salisbury University Brown Bag Seminar, the 2018 Southern Economic Association in Washington, DC, and the USDA workshop entitled “Applications and Potential of Ecosystem Services Valuation within USDA: Advancing the Science.” John Morehouse, Garrett Stanford, and the editorial staff at *Land Economics* (especially the copyeditor, Laura Poole) have provided helpful expositional suggestions. The findings and conclusions are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. All remaining errors are the responsibility of the authors.

References

- Bakhtiari, F., J. B. Jacobsen, B. J. Thorsen, T. H. Lundhede, N. Strange, and M. Boman. 2018. “Disentangling Distance and Country Effects on the Value of Conservation across National Borders.” *Ecological Economics* 147: 11–20.
- Cameron, T. A., and J. R. DeShazo. 2013. “Demand for Health Risk Reductions.” *Journal of Environmental Economics and Management* 65 (1): 87–109.
- Dundas, S. J., and R. H. von Haefen. 2020. “The Effects of Weather on Recreational Fishing Demand and Adaptation: Implications for a Changing Climate.” *Journal of the Association of Environmental and Resource Economists* 7 (2): 209–42.
- Glenk, K., R. J. Johnston, J. Meyerhoff, and J. Sagebiel. 2020. “Spatial Dimensions of Stated

- Preference Valuation in Environmental and Resource Economics: Methods, Trends and Challenges." *Environmental & Resource Economics* 75 (2): 215–42.
- Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1): 153–61.
- Holland, B. M., and R. J. Johnston. 2017. "Optimized Quantity-within-Distance Models of Spatial Welfare Heterogeneity." *Journal of Environmental Economics and Management* 85: 110–29.
- Johnston, R. J., and A. S. Abdulrahman. 2017. "Systematic Non-response in Discrete Choice Experiments: Implications for the Valuation of Climate Risk Reductions." *Journal of Environmental Economics and Policy* 6 (3): 246–67.
- Kolstoe, S., and T. A. Cameron. 2017. "The Non-market Value of Birding Sites and the Marginal Value of Additional Species: Biodiversity in a Random Utility Model of Site Choice by eBird Members." *Ecological Economics* 137: 1–12.
- Kolstoe, S., T. A. Cameron, and C. Wilsey. 2018. "Climate, Land Cover, and Bird Populations: Differential Impacts on the Future Welfare of Birders across the Pacific Northwest." *Agricultural and Resource Economics Review* 47 (2): 272–310.
- Logar, I., and R. Brouwer. 2018. "Substitution Effects and Spatial Preference Heterogeneity in Single- and Multiple-Site Choice Experiments." *Land Economics* 94 (2): 302–22.
- Loomis, J. B. 1996. "How Large Is the Extent of the Market for Public Goods: Evidence from a Nationwide Contingent Valuation Survey." *Applied Economics* 28 (7): 779–82.
- Roberts, A. J., P. K. Devers, S. Knoche, P. I. Padding, and R. Raftovich. 2017. "Site Preferences and Participation of Waterbird Recreationists: Using Choice Modelling to Inform Habitat Management." *Journal of Outdoor Recreation and Tourism—Research Planning and Management* 20: 52–59.
- Rosenberg, K. V., A. M. Dokter, P. J. Blancher, J. R. Sauer, A. C. Smith, P. A. Smith, J. C. Stanton, A. Panjabi, L. Helft, M. Parr, and P. P. Marra. 2019. "Decline of the North American Avifauna." *Science* 366 (6461): 120–24.
- Sen, A., A. R. Harwood, I. J. Bateman, P. Munday, A. Crowe, L. Brander, J. Raychaudhuri, A. A. Lovett, J. Foden, and A. Provins. 2014. "Economic Assessment of the Recreational Value of Ecosystems: Methodological Development and National and Local Application." *Environmental & Resource Economics* 57 (2): 233–49.
- Swait, J., C. Franceschinis, and M. Thiene. 2020. "Antecedent Volition and Spatial Effects: Can Multiple Goal Pursuit Mitigate Distance Decay?" *Environmental & Resource Economics* 75 (2): 243–70.
- U.S. Fish and Wildlife Service. 2018. "2016 National Survey of Fishing, Hunting, and Wildlife-Associated Recreation." Available at https://www.fws.gov/wsfrprograms/subpages/nationalsurvey/nat_survey2016.pdf.
- Vogdrup-Schmidt, M., A. L. Abatayo, J. F. Shogren, N. Strange, and B. J. Thorsen. 2019. "Factors Affecting Support for Transnational Conservation Targeting Migratory Species." *Ecological Economics* 157: 156–64.
- Walsh, P. J., J. W. Milon, and D. O. Scroggin. 2011. "The Spatial Extent of Water Quality Benefits in Urban Housing Markets." *Land Economics* 87 (4): 628–44.