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

Trudy Ann Cameron
R.F. Mikesell Professor of Environmental and Resource Economics
Department of Economics, 435 PLC
1285 University of Oregon, Eugene, OR 97403

Sonja Kolstoe
Assistant Professor
Department of Economics & Finance and
Department of Environmental Studies
Salisbury University, 1101 Camden Avenue, Salisbury, MD 21801

September 26, 2020

JEL classification: C83, Q51, Q57, Q58

Keywords: citizen science, community science, sample selection correction, selection propensity, ordered probit selection, inverse Mills ratios, exogenous weighting, eBird, bird-watching, consideration sets, distance-decay, trip generation, market extent functions, scalability

*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, D.C. and the USDA Workshop Applications and Potential of Ecosystem Services Valuation within USDA—Advancing the Science. John Morehouse and Garrett Stanford provided helpful editorial suggestions. Any remaining errors should be attributed to the authors.

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

ABSTRACT

Citizen science (CS) projects (and some social media) offer selected samples with extensive information about human interactions with the natural world. We independently survey (1) members of the eBird CS project and (2) a general population sample, eliciting awareness and/or levels of engagement with the eBird project in each case. The general-population sample allows an ordered-probit model to explain propensities to engage with eBird, which we transfer to predict selection-correction terms for our independent sample of eBird members. We illustrate, using a direct question posed only to our eBird member survey sample about the radius of their spatial consideration set for a typical one-day birding excursion.

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 individual provides depends on their level of engagement with the platform, be it a citizen science project or social media platform. In addition, these data pertain only to the contributing set of users, be they citizen scientists, social media users, etc. These are "samples of convenience" rather than random samples from the overall population. Sample selection bias is therefore an obvious concern. Users' intensity of engagement or participation with the project or platform affects the likelihood 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 have the potential to be a valuable research resource for environmental economists seeking to provide scalable and policy-relevant inferences.

Non-representative voluntary surveys are often used by environmental economists to collect data. As a consequence, variety of different methods have been developed to correct for respondents' differing propensities to respond to the survey and therefore to be part of the estimating sample. These methods certainly include the the traditional method of Heckman (1979).[2] Or, alternative ad hoc approaches have been proposed, as in Cameron and DeS-

[1]"Citizen science" or "community science" (CS) 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 smart-phone applications. As of February 2020, there are now more than 2,000 active CS projects, according to the *Citizen Science Association* (see `CitizenScience.Org`), 448 are registered in the federal crowd-sourcing and CS registry (see `CitizenScience.Gov/Catalog/#`)

[2]Heckman (1979) is a foundational paper for the least-squares context, now cited more than 10,000 times in Web of Science.

hazo (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 mis-represent the influence of low- and high-intensity participants, 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 paper, we develop a new approach to sample-selection correction for CS/VGI data. The goal of this approach 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 citizen science project. The eBird project has already proven itself to be a valuable CS/VGI data source for both natural scientists and social scientists alike (e.g. Rosenberg et al. (2019), Kolstoe and Cameron (2017), Kolstoe et al. (2018), Roberts et al. (2017)). Furthermore, bird-watching is a very popular pastime. About 45.1 million people observed birds in the US, both around home and away from home, according to the U.S. Fish & Wildlife Service's 2016 survey on Fishing, Hunting, and Wildlife-Associated Recreation (FHWAR) report.

We use, in tandem, 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 U.S. Both of our samples include specific information from respondents about the *degree* to which they participate in the eBird project, so that we can distinguish both the extensive margin (whether an individual participates in eBird at all), and the intensive margin (the degree to which they engage with this CS project).

We propose three candidate strategies for sample-selection correction. At the most basic level (with its details therefore relegated to an 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 our 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 upon 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 our eBird member survey sample. Demeaned individual predicted engagement propensities in the eBird member survey sample, normalized on the mean engagement propensity in the qBus general population sample, are then allowed to shift both 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 three types of 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 et al. (2011), and (Glenk et al., 2020, section 3.1.2). Our study presents a unique opportunity to address consideration sets because we included in our 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 actually travelled 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 systematic variations across our sample of eBird members in their directly elicited individual 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 notion of a consideration-set radius for an individual is also somewhat related to other concepts in the revealed and stated preference literatures. 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

[3]The question explicitly excludes trips to destinations with reported rare-bird sightings. The relevant consideration-set radius for an individual, in those special cases, can be expected to be much larger. This would be consistent with the distinction between iconic and non-iconic destinations in the market-extent literature for such destinations.(Glenk et al., 2020, footnote 5).

and destination attributes). This TGF approach does not focus on the maximum distance willingly traveled by an individual. Nor does it emphasize heterogeneity in this maximum distance across individuals with different characteristics.[4]

The spatial stated-preference literature also 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 substitution destinations because the *area* of a circle around a given origin location increases much more quickly than the *radius* of that circle.[5] 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 paper 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.

In Section 2 of this paper, we generalize the familiar binary-probit first-stage selection equation to a six-level ordered-probit selection equation. We then explain how to use this ordered-probit selection equation, estimated for one sample and yielding an inverse Mills ratio *function*, can then be transferred to an independent sample. We also adapt to this two-sample context an ad hoc alternative selection-correction method based on demeaned selection propensities. We also develop *heterogeneous* sampling weights based on ordered-probit models to describe eBird engagement intensity (fitted separately to the general-population

[4]However, an estimated TGF model could presumably be solved for the average travel time (and therefore approximately the average distance) at which expected trips fall to zero, conditional on origin and destination attributes. But in such a case, the origin attributes would typically be medians or proportions of the population within an origin area, rather than individual characteristics.

[5]The average radius around an origin at which predicted demand drops to zero could also be solved from an estimated willingness-to-pay function that includes distance as an argument.

sample and the eBird member survey sample). We relegate the details of our weighting strategy to an online appendix. Section 3 briefly discusses our estimated selection models, and Section 4 discusses our "outcome" model for heterogeneous consideration-set radii for birding excursions. We compare parameter estimates and predicted consideration-set radii when the model is estimated both naively and with our different types of selection-correction strategies. Section 5 concludes and recommends strongly that future general-population surveys like the FHWAR be expanded to include questions about engagement with CS projects related to ecosystems services that influence the demand for recreational uses of the natural environment.

2 Strategies for Dealing with Systematic Sample Selection in eBird data

Our "eBird member survey" sample is self-selected, consisting only of eBird members who chose to respond to our survey. These birders are also likely to participate in the eBird project with a different mix of engagement levels than might be expected for members of the general population. Specifically for this study, over several waves of the Qualtrics Omnibus (qBus) survey, we also independently surveyed more than 4000 respondents from that general-population panel. Online Appendix A offers some further discussion of our qBus sample, and Table A2 contrasts a simple binary indicator for eBird citizen-science participation, CS, with the greater level of detail in our six ordered categories of engagement intensity, $CS6$, elicited from both the qBus data and our sample of eBird members.[6]

As noted in the introduction, standard selection-correction models use a *binary* selection model. In this paper, we increase the level of detail by switching to an ordered-probit

[6]The Qualtrics Omnibus surveys have been discontinued, but there remain numerous other Omnibus options. See Online Appendix A, Table A3.

selection model, using six categories for our "selection into citizen science project participation intensity" model where eBird is the specific citizen science project in question.[7] 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 nonparticiation. Our additional level of detail about engagement intensity provides unusual but valuable information about the "intensive margin" of participation in eBird, for both of our samples.[8]

2.1 Selection in the general-population qBus sample

For the $i = 1, ..., N$ individuals in our general population (qBus) sample, let citizen science 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 set of variables on sociodemographics and income, Z_i, that we will use to explain eBird participation intensity, where respondents $i = 1, ..., r$ participate in eBird at one of four different levels and respondents $i = s, ..., N$ do not participate in eBird (but may either have heard about eBird, or not). If we sort these observations in decreasing order of participation intensity, the data for

[7] We will 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. Unlike our previous research with eBird data, the current analysis is not affected by substantial shares of missing or out-of-date home address information needed to allow calculation of actual travel distances from each person's home to all of their relevant birding destinations.

[8] Practitioners should be aware that an adjustment may be necessary, to the intercept of the fitted propensity-to-engage with eBird, depending upon how the ordered probit algorithm has been parameterized. Comparability is necessary across specifications if one wishes to compare fitted "propensity" estimates across ordered-probit selection models with differing numbers of levels.

the selection model can be written as:

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

For the qBus sample, we can model an underlying continuous latent propensity (denoted with an asterisk) to be a member of eBird 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, $y_i = X_i\beta + \epsilon$. We could implement either a standard binary-probit selection model, or the six-level ordered-probit selection model we develop in this paper. 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.[9]

2.2 Outcome variable for eBird member survey sample

For the $j = 1, ..., J$ 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

[9] A researcher could attempt to collect all the variables provided by our eBird member survey and its linked eBird citizen-science observations from a large sample of respondents drawn from the general population. This would likely be impractical and inefficient, however, given the number of survey questions that would be required (and hence the cost of using a representative panel) as well as the potential for recall bias concerning the respondent's history of birding activity.

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 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:

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

For the $j = 1, ...J$ 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 $j = 1, ..., J$ observations in the eBird member survey sample alone, the relevant pair of equations for our selection-correction model would be:

$$\begin{aligned} CS6_j^* &= Z_j\gamma + \eta_j \qquad (1) \\ y_j &= X_j\beta + \epsilon_j \\ (\eta_j, \epsilon_j) &\sim BVN(0, 0, \sigma_\eta, \sigma_\epsilon, \rho) \end{aligned}$$

Of course, this complete joint model cannot be estimated using our eBird members survey sample alone, because there are no *non-members* of eBird in the $j = 1, ...J$ observations from that survey (i.e. there are no observations with $CS6_j = 1$ or $CS6_j = 2$.[10]

2.3 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 general population sample who happen to be eBird members. We have these variables only for our completely separate 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 likewise that the underlying statistical relationship $(CS6^*, y) \sim BVN\left(Z\gamma, X\beta, 1, \sigma_\epsilon, \rho\right)$ applies for both the $i = 1, ..., N$ members of our qBus sample and for $j = 1, ..., J$ for members of our eBird sample.[11]

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 fitted index for each member of the eBird member survey sample, $Z_j\hat{\gamma^q}$, that takes account of the potentially different pattern of Z_j characteristics in our eBird member survey sample. We can use this fitted 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,

[10]For readers who may wish to review the conventional Heckman two-step sample-selection correction procedure in more detail, we provide a summary in Online Appendix B.

[11]Mechanically, it would be possible to pool our two samples and use the combined dataset 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 dataset that no longer represents the general population. One could, potentially, weight the eBird member survey sample according to the proportion of eBird members in the qBus sample, but this would still leave a pooled sample for the selection equation that is not randomly selected from the general population.

noting the j subscripts for the eBird data.[12]

$$\begin{aligned} E\left[y_j|y_j\ observed\right] = E\left[y_j|CS6_j^* > -Z_j\hat{\gamma^q}\right] &= X_j\beta + \rho\sigma_\epsilon\lambda\left(-Z_j\hat{\gamma^q}\right) \\ &= X_j\beta + \beta_\lambda\lambda\left(-Z_j\hat{\gamma^q}\right) \\ Var\left[y_j|y_j\ observed\right] = Var\left[y_j|CS6_j^* > -Z_j\hat{\gamma^q}\right] &= \sigma_y^2\left[\left(1 - \rho^2\delta\left(-Z_j\hat{\gamma^q}\right)\right]\right. \end{aligned} \tag{2}$$

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 desired unconditional (i.e. non-systematically selected) expectation for y_j can be simulated, counterfactually, by setting $\rho = 0$, so that $E\left[y_j|y_j\ observed\right] = X_j\beta$ and $Var\left[y_j|y_j\ observed\right] = \sigma_y^2$.

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 calculated as $\phi(Z_j\hat{\gamma^q})/\Phi(Z_j\hat{\gamma^q})$. As in the case of a standard binary probit selection correction, this term can be appended to the list of regressors, X_j, in the outcome equation of interest.[13,14]

[12] We note that it 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 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 later in this paper. Here, therefore, we assume the qBus and eBird selection-equation error distributions are identical.

[13] It will be appropriate, in future research, to graduate to a 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 should be estimated as an ordered-probit model.

[14] In Online Appendix C, we provide a detailed discussion of the types of outcome models where may be appropriate to contemplate just adding an IMR term to correct for sample selection. Here, we simply note that this straightforward strategy is likely to be suitable only if the latent continuous dependent variable selection equation and the observed (or latent) continuous dependent variable in the outcome equation can be modeled as involving error terms that are plausibly jointly normally distributed.

2.4 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 employed. We use the estimated engagement propensity model from our first-stage selection model to calculate *fitted* propensities to engage with eBird at each of 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 then 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 of the β parameters in the outcome model. After estimation, this demeaned response propensity can be counterfactually set to zero, effectively dropping all of 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.[15]

[15] For identification, one or more exogenous explanatory variables should be *included* in the $Z_i\gamma$ 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.

3 Selection Model: Implementation

3.1 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 for the qBus and eBird member survey datasets). 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 the 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)
- 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.” As a consequence, 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 through 6 (where a qBus respondent is at least a member of eBird), then the relative frequencies for engagement levels 3 through 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). While 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.

Table 1: Descriptive statistics (proportions) for variables in first-stage Bird engagement intensity models. Availability indicators are proportions of total sample; group shares are proportions of the available data

	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 or less	0.125	0.018
Age: 25 to 34 years	0.224	0.065
Age: 35 to 44 years	0.196	0.089
Age: 45 to 54 years	0.135	0.146
Age: 55 to 64 years	0.175	0.311
Age: 65 years and up	0.145	0.370
Income data available	1.000	0.804
Income: Less than 25K	0.179	0.072
Income: 25 K to 50 K	0.219	0.203
Income: 50 K to 75 K	0.189	0.231
Income: 75 K to 100 K	0.141	0.173
Income: 100 K or more	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
Empl. status data available	1.000	0.849
Empl. status: Full time	0.473	0.359
Empl. status: Part time	0.132	0.080
Empl. status: Looking for work	0.057	0.008
Empl. status: Unemployed	0.145	0.066
Empl. 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: Masters degree	0.118	0.396
Education: Doctoral degree	0.038	0.121
Observations	4161	1081

3.2 Estimation results for selection model

3.2.1 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 all of the standard demographic variables used in our selection model are part of the "profile" data supplied for each qBus panelist, rather than being information we sought to elicit via our own questions. There are considerably more missing values for the Z_j variables from our eBird member survey, since all of the sociodemographic 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 our 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 non-missing data for that particular eBird respondent.[16] 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 Online Appendices 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 has been relegated to the Appendices.[17]

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.[18]

[16] In 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.

[17] See Online Appendix G for the models using eBird data. Our heterogeneous population weights for the eBird member survey sample are described in Online Appendix H. As a technical note regarding the correspondence between the qBus and eBird models, Stata's default parameterization of the thresholds between ordered-probit intervals is not comparable between the six-interval qBus models and these four-interval eBird member survey models. E.g. 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 same two levels. A simple change-of-origin restores comparability. The *polr* package in R handles ordered-probit models, with its "method" argument set to "probit."

[18] At the other end of the spectrum of data completeness in the eBird member survey sample, Online 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 employ 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.

Table 2: Estimated coefficients, ordered-probit engagement-level model; example with maximum heterogeneity (6-level model using full qBus sample; 4-level model for the *subset* of 572 eBird survey respondents with complete data for this least-restricted specification[a])

	Ordered probit qBus data		Ordered probit eBird data	
Travel 1+ mile data available	0.136	(0.661)	- [b]	
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)	- [c]	
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)
Relative to omitted category: 45 to 54 years				
Age: 24 years or less	0.539***	(0.0991)	0.994**	(0.424)
Age: 25 to 34 years	0.545***	(0.0871)	0.293	(0.216)
Age: 35 to 44 years	0.360***	(0.0894)	0.325*	(0.192)
Age: 55 to 64 years	-0.207*	(0.110)	-0.117	(0.159)
Age: 65 years and up	-0.298**	(0.134)	0.146	(0.201)
Relative to omitted category: \$50K to \$75K				
Income: Less than \$25K	-0.0590	(0.0853)	-0.0751	(0.254)
Income: \$25K to \$50K	-0.0192	(0.0774)	0.143	(0.150)
Income: \$75K to \$100K	-0.0423	(0.0867)	-0.0150	(0.158)
Income: \$100K or more	-0.0118	(0.0784)	0.143	(0.138)
Relative to omitted category: West				
Region: Northeast	0.164**	(0.0737)	- [c]	
Region: Midwest	-0.0186	(0.0754)	- [c]	
Region: South	0.0552	(0.0660)	- [c]	
Relative to omitted category: Full time				
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)
Relative to omitted category: 4-year college degree				
Education: High school	-0.0294	(0.0754)	0.693**	(0.309)
Education: Some college	-0.128*	(0.0671)	0.00246	(0.170)
Education: Masters degree	0.269***	(0.0826)	0.259**	(0.121)

Education: Doctoral degree	0.165	(0.124)	0.00792	(0.165)
Ordered-probit thresholds				
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)	- [d]	
cut5	3.341***	(0.132)	- [d]	
Observations	4161		572	
Max. log-likelihood	-2390.32		-582.44	

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ [a] Specifications with fewer explanatory variables can use more observations in the eBird dataset. Other models to accommodate all of the patterns of missing variables in the eBird data are relegated to the Online Appendices. [b]Too few missing values. [c]Equals 0 for all. [d]Only 4 levels.

To predict participation intensities for respondents to our eBird member survey, we use the γ^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 non-response 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), as well as predicted engagement intensities and predicted IMR terms to be used for sample-selection corrections in outcome equations that rely upon 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 within the eBird member survey sample, an ingredient in our heterogeneous sampling weights described in Online Appendix H.[19]

One of the key innovations in this paper 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 still be estimated and used in an analogous manner, although it contains less

[19]To 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," i.e. *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.

information.[20]

In comparing the coefficient estimates for each sample in Table 2, we note numerous differences. These differences do not imply, however, 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. Our 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 last year, the more days per 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 or not 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 of age or younger. Older respondents in the

[20] Online Appendix I digresses to explore a variety of the intermediate components of our models. For example, 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).

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 U.S. 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 Masters degree increases eBird engagement propensity.[21]

3.2.2 Transferring qBus selection model to eBird member survey sample

We next use the assumption that for each individual in our eBird member survey sample, we can transfer the relevant set of $\hat{\gamma}^q$ parameters estimated using the qBus data. The qBus selection model to be transferred needs to be estimated using the same set of non-missing 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 possess about that individual eBird member's sociodemographic characteristics.[22]

[21] In our eBird member survey sample, only about 3.5% of respondents have just a high school education or less, so perhaps not much should 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 projects recruit 18-year-olds still in high school, however, this could account for the greater eBird engagement propensities in this group.

[22] Online Appendix 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. In this case, the ordered-probit coefficients tend to predict lower IMRs than do the binary-probit coefficients, at least over the upper half of the distribution. Keep in mind that the selection process for the eBird sample is actually the compound effect of selection into eBird and selection into our sample of survey respondents. To the extent that the sample from our eBird member survey does not represent the population of eBird members, there may be a second layer of selection to consider. We ignore that additional complexity in this paper.

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

4.1 Available variables for outcome model

This section illustrates the use of our predicted, rather than estimated, IMR terms in 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 than 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 both 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 available in the qBus sample. For example, our eBird member survey elicits income in much finer brackets than we could use in the

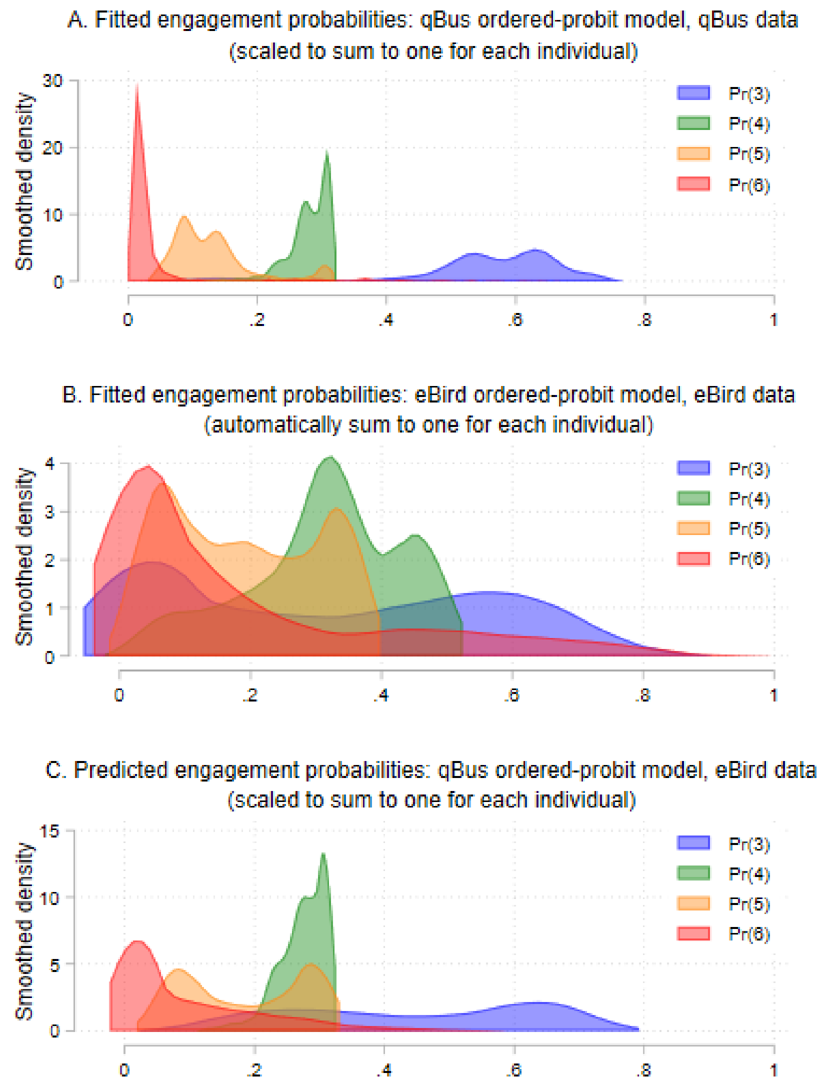

Figure 1: Predicted probabilities for engagement levels 3, 4, 5, and 6. Panel A: fitted engagement level probabilities for qBus sample; Panel B: fitted engagement probabilities for the eBird sample; Panel C: predicted engagement probabilities for the eBird sample, based on an ordered-probit model estimated using the qBus sample.

Table 3: Descriptive statistics: Variables for outcome model, elicited from eBird member survey sample for n = 1,081 respondents who answered the question about maximum one-way distance for a birding day-trip

	mean	sd
Dependent variable (market extent, lower bound, chosen interval):		
Self-reported maximum radius of travel in miles	83.283	58.104
Explanatory variables:		
Empl. 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: Less than 45 years	0.171	0.377
Age: More than 64 years	0.367	0.482
Education: Grad school	0.505	0.500
No Interest: Perching birds	0.057	0.233
No Interest: Other game birds	0.111	0.314
Selection-correction options:		
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	1081	

engagement intensity models, so we convert the income bracket data into an approximate continuous income variable.

We also take advantage of our eBird member survey data concerning eBirders' interests in different species categories. For various categories of bird species, between 6% and 11% of eBirders report that they have no interest in that category. The least popular category in our eBird member sample, for example, is "game birds other than waterfowl" (e.g. pheasants, turkeys, grouse or partridges). This information about the goals of individual birders ties our analysis to the notions explored in Swait et al. (2020), who find that benefit variations associated with distance depend upon 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 impacts 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 most bird-watchers, especially if climate change accelerates.)

4.2 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. We thus assume that these distances are strictly positive. Given that the boundaries for these brackets are known, a reasonable estimation method assumes that the latent continuous dependent variable is conditionally lognormally distributed. An interval-data regression model can then be estimated by maximum likelihood methods.[23]

Model 1 in Table 4 is a naive specification to explain consideration-set radius (maximum willingness to travel to see birds) with no corrections. The other columns show several alternative types of corrected models. Model 2 is an otherwise naive specification that employs only our constructed weights, as detailed in Online Appendix H. Models 3 through 5 continue to employ these weights. Model 3 includes an IMR variable based on a conventional binary-probit selection equation, and Model 4 employs our novel ordered-probit selection

[23]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."* Were we to resort to FIML estimation of a joint model for the pooled samples, we could possibly entertain an interval regression specification with extra probability at zero. Our survey includes a question that reads: "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 trips to closer destinations. We could also look for eBird members for whom every birding report is identically geo-located, although we would have to assume that this location was their home, to conclude that their maximum historical travel distance has been zero. Ultimately, the notion of consideration sets elicited from our eBird respondents is *prospective*, not revealed from their past behavior, so we do not attempt to implement a zero-inflated interval-regression specification in this study.

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).[24] 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 (RUM).

In the models reported 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, as well as whether they specifically express *no* interest in each of two categories of bird species.[25]

[24] For Models 2 through 5 in Table 4, the weights are designed to correct for differences in observable determinants of engagement intensities among eBird members, to make this particular sample of eBird members more representative of engagement intensities among eBird members in the general population (i.e. in the qBus sample, in this case).

[25] Our survey elicited levels of interest in five categories of species, but for only two categories does disinterest have statistically significant effects on consideration-set radii. We have also explored other potential explanatory variables, but exclude them because they have persistently statistically insignificant coefficients across a wide variety of specifications.

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 propensity. Dependent variable: logarithm of maximum one-way distance willingly traveled on a typical birdwatching day-trip.

	(1) Naive	(2) Weights only	(3) Binary probit IMR	(4) Ordered probit IMR	(5) Demeaned propensity
Main variables					
Empl. Status: Employed	-0.0351	-0.0613	-0.0872	-0.0934	-0.299*
	(0.0696)	(0.0973)	(0.0943)	(0.0936)	(0.154)
Income data available	-0.429***	-0.362**	-0.342*	-0.354**	-0.223
	(0.127)	(0.180)	(0.176)	(0.175)	(0.186)
ln(Income in 10K, if reported)	0.218***	0.171**	0.156**	0.158**	0.0628
	(0.0533)	(0.0730)	(0.0718)	(0.0710)	(0.0887)
Gender: Female	-0.127**	-0.123	-0.0374	-0.0410	0.000944
	(0.0590)	(0.0796)	(0.0787)	(0.0783)	(0.106)
Age: Less than 45 years	0.185**	0.225**	0.167	0.137	-0.317
	(0.0835)	(0.112)	(0.106)	(0.108)	(0.197)
Age: More than 64 years	-0.0489	-0.114	-0.0287	-0.103	-0.0110
	(0.0713)	(0.0982)	(0.100)	(0.0961)	(0.115)
Education: Grad school	0.113*	0.0132	-0.0649	-0.0642	-0.259***
	(0.0598)	(0.0751)	(0.0741)	(0.0748)	(0.0958)
No Interest: Perching birds	-0.811***	-0.862***	-0.847***	-0.829***	-0.749***
	(0.148)	(0.198)	(0.205)	(0.202)	(0.190)
No Interest: Other game birds	-0.634***	-0.660***	-0.577***	-0.578***	-0.549***
	(0.107)	(0.130)	(0.131)	(0.130)	(0.127)

Continued on next page

Constant	4.066*** (0.0972)	4.170*** (0.127)	4.494*** (0.135)	4.403*** (0.131)	4.102*** (0.140)
Selection-correction strategies					
Binary probit IMR			-0.169*** (0.0305)		
Ordered probit IMR				-0.386*** (0.0689)	
Interactions between main variables and demeaned engagement propensity (demeaned using qBus sample mean)					
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: Less than 45 years × Engagement prop. (demeaned)					-0.00451 (0.0733)
Age: More than 64 years × Engagement prop. (demeaned)					-0.0223 (0.0607)
Education: Grad school × Engagement prop. (demeaned)					0.0176 (0.0551)
No Interest: Perching birds × Engagement prop. (demeaned)					-0.231 (0.175)

Continued on next page

No Interest: Other game birds × Engagement prop. (demeaned)					0.160
					(0.107)
Constant × Engagement prop. (demeaned)					0.282***
					(0.0813)
lnsigma					
Constant	-0.0803***	-0.0583*	-0.0788**	-0.0808***	-0.105***
	(0.0237)	(0.0306)	(0.0313)	(0.0310)	(0.0310)
Observations	1081	1081	1081	1081	1081
Log Likelihood	-2411.41	-2429.96	-2409.82	-2408.03	-2382.37
AIC	4844.83	4881.92	4843.64	4840.05	4806.74
BIC	4899.67	4936.77	4903.47	4899.88	4911.44
Weighted?	No	Yes	Yes	Yes	Yes

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Compared to Model 1, 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 inverse Mills ratio. In two-stage methods, this coefficient is the estimate of $\rho\sigma_\epsilon = \beta_\lambda$, as in equation (3). 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 non-stochastic (thereby understating the amount of noise in the model). Nevertheless, these negative IMR coefficients are strongly strongly statistically significantly different from zero.[26]

A priori, we expected (if anything) that the propensity to participate in eBird would be positively associated with a respondent's consderation-set radius, since latent birding avidity

[26]FIML estimation of the joint model for engagement propensity and consideration-set radii could remedy this estimated-regressors problem and should be pursued in future applications where hypothesis testing is particularly important for policy.

could be a potentially important omitted variable. It is thus somewhat counter-intuitive to find negative coefficients on our sample-selection correction terms. Instead of birding avidity, the relevant unobserved heterogeneity might include the opportunity cost of time (or perhaps unobserved age-related technical sophistication in using the online eBird app, or for online surveys in general).

Employment status. None of Models 1 through 4 in Table 4 suggest that employment status has a statistically significant effect on consideration-set radii for birders. However, Model 5, using our ad hoc correction of interacting each of the regressors with the demeaned engagement propensity variable, suggests that for the general population, consideration-set radius is smaller by about 30% if the respondent is currently employed. This is not surprising. Employed individuals are likely to have less free time for all leisure activities, including birding day-trips.

Income data availability indicator and level of income, if known. For all but Model 5, compared to 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% to 43%. However, this negative effect in these specifications 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% to 0.22%. This may seem plausible because higher-income respondents likely have less-binding budget constraints for travel expenses.[27] However, Model 5 suggests that in the general population, income has no statistically discernible effect on consideration-set radii for birding trips.

Gender. The point estimate for the effect of being female on consideration-set radius is estimated to be negative in Models 1 through 4 (although the estimate is statistically

[27] Compared to 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 to $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.

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 radius 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 to the omitted category of 45 to 64 years old) is associated with a consideration-set radius that is larger by about 18% to 23%, but this effect disappears in Models 3 through 5 that explore alternative remedies for systematic selection.

Education. Relative to the omitted category of eBird members with college degrees 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 through 4 suggest that graduate school has no statistically significant effect on radius, but Model 5 implies that in the general population, graduate school is actually 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 effect are also similar across all specifications. Reporting a lack of interest in either of these categories of birds shrinks expected radius substantially, by 58% to 86%.

Model 5's interaction terms. Among the selection-correction models, Models 3 and 4 rely upon a strong assumption of bivariate normal errors and slope coefficient 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 explanatory variables that are uncontaminated by sample selection. The inverse Mills ratio strategy can be described as a structural approach to sample selection correction.

In contrast, Model 5 is ad hoc, unstructured and highly flexible. This approach makes a different, but perhaps equally strong assumption—that each of the parameters 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 have 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 our estimating sample is not representative of that population. In Table 3, note that the average demeaned engagement propensity in the estimating sample is positive, at about 0.657. The people who appear in our estimating sample from our eBird member survey understandably 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 non-interacted basic regressors in the outcome model 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 systematically with the individual's predicted engagement propensity, where that predicted engagement propensity is based on our adjusted ordered-probit engagement intensity model.

The interaction terms in Model 5 suggests that while 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. This interaction is just the demeaned engagement 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 has an average value that is positive, the eBird member survey sample, without correction, overstates the consideration-set radii for birders in the general population.

One important observation about the demeaned engagement propensity variable in Model 5 is that the ordered-probit inverse Mills ratio 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 $\rho\sigma_\epsilon = 0$). In Model 5, if we were to include just the demeaned engagement propensity variable without its interactions with the basic regressors, we would zero out the demeaned propensity variable itself to produce corrected predictions about consideration-set radii. Given the degree of correlation between the two variables, either type of 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 to permit not only the intercept to differ (with either the inverse Mills ratio or the demeaned engagement intensity), but all the slopes as well.

4.3 Predicted values for the outcome variable, with and without corrections

The point of correcting for systematic sample selection is to adjust the statistical relationships observed in the selected sample to better reflect the general population. In this section, we compare the predicted consideration-set radii for selected specifications.[28] The top graph in Figure 2 shows the 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 any correction for sample selectivity in the eBird citizen science sample. Model 4, with its ordered-probit inverse Mills ratio 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 inverse Mills ratio 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 characteristics would predict.

However, the effects of systematic selection on predicted consideration-set radii 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, likewise plotted 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. The pattern of clustering in these predicted

[28]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 $(\sigma^2/2)$ to recover the mean of the fitted conditional distribution, due to the skewness of the implied log-normal distribution.

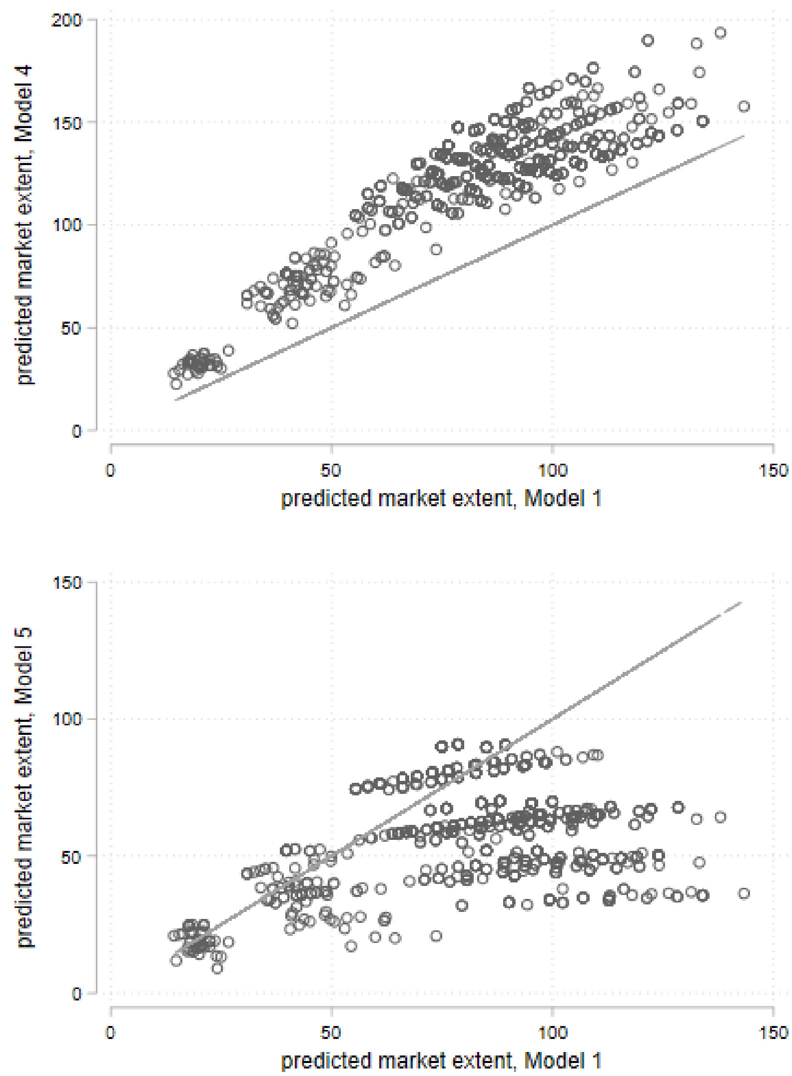

Figure 2: Predicted consideration set radii for 1,081 observations in the eBird sample based on Model 4 versus Model 5 in Table 4 plotted against the predicted radii from the naive Model 1 (45-degree lines added).

values stems from the fact that the demeaned engagement propensity is a function mostly of indicator variables, and the interaction terms in the outcome model for consideration-set radii are likewise mostly indicator variables. 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 notable, given that roughly 12% of the general population that 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]

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 tool-kit for using citizen science data—with improved confidence that any insights to be derived are more suitable for scaling to the general population or for use in benefits transfer exercises. The two main tasks in the paper are to (1) illustrate some new sample selection correction techniques we have developed to allow for the use of data from auxiliary general-population surveys to correct

[29]Nevertheless, we note that even consideration-set radii of zero, for actual travel away from home to enjoy "active use" of wild birds, do not preclude the possibility of "passive use" values (option, bequest, or existence values) for wild birds in the region. Birds are also mobile, not just people. The presence of wild birds within any given radius will also affect the probability that these birds may be viewed in one's backyard, without the necessity of travel.

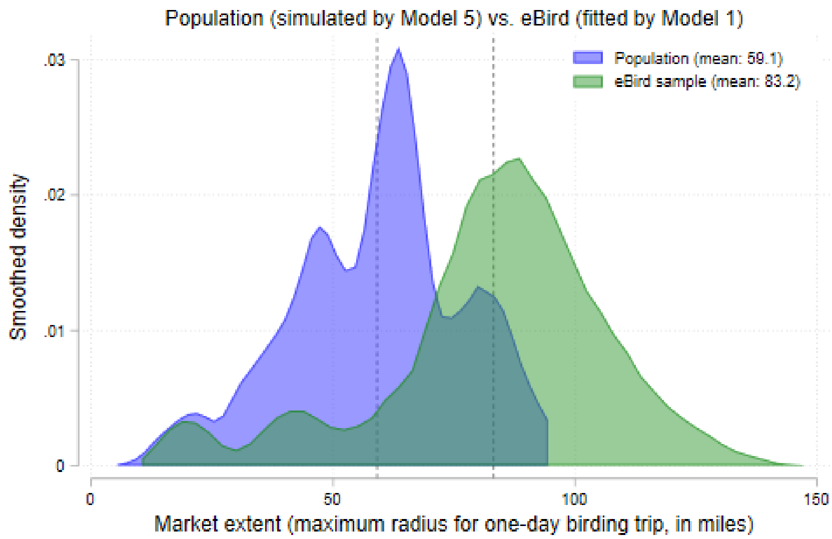

Figure 3: For 1,081 eBird member survey observations: Marginal distribution of predicted consideration-set radii according to our naive Model 1 overlaid by selection-corrected predictions of these radii for (birders in) the general population based on Model 5.

for sample selection present in citizen science data, and (2) model heterogeneity in the radius of consideration sets for regular bird-watching 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 very important for scaling to the general population, or transferring to other contexts, any results derived from citizen science data.

The key takeaway from our illustrative consideration-set radius "outcome" model is the potential importance of non-random selection into citizen science 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 because land-cover change and climate change present significant threats to wild bird populations. Changes in bird populations affect bird-watcher welfare (see Kolstoe et al.

(2018) for an illustrative example). 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 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 not to have been sufficient, given that avian biodiversity remains a concern (in light of 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 (a concern also explored by Bakhtiari et al. (2018) and Vogdrup-Schmidt et al. (2019)). These market extents are dictated by the consideration-set radii of individual birders.

The need for a qBus-type sample to permit sample selection corrections in this instance highlights a potential supplementary role for broad-based surveys of bird-watching trip behavior and citizen-science 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. However, as of 2016, the information began to be reported only at the census division level, rather than the state level, as had been the case for prior waves of the survey. The loss of geographic resolution due to this decision limits the usefulness of FHWAR information for city-, county- and state-level government agencies.

The FHWAR survey is perhaps the most appropriate existing survey to which a detailed question could be added about participation in outdoor-based CS projects (assuming U.S. Fishing, Hunting & Wildlife Watching Survey continues in the future). Also, given that the federal registry now documents more than 400 CS projects (see `www.citizenscience.gov`), general-population information on CS participation would benefit other agencies, such as the National Oceanic Atmospheric Administration (NOAA) or US 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 could (and should) include questions about citizen science engagement in projects related to ecosystems services that are valued for active recreational activities. This general-population engagement information would be a vital complement to any special-purpose surveys fielded to members of CS projects to help researchers understand both 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.

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. DeShazo (2013). Demand for health risk reductions. *Journal of Environmental Economics and Management 65*(1), 87–109.

Cameron, T. A. and J. Englin (1997). Respondent experience and contingent valuation of environmental goods. *Journal of Environmental Economics and Management 33*(3), 296–313.

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), 209–242.

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), 215–242.

Greene, W. H. (2012). *Econometric Analysis (7 th)*. Upper Saddle River, NJ: Prentice Hall.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.

Holland, B. M. and R. J. Johnston (2017). Optimized quantity-within-distance models of spatial welfare heterogeneity. *Journal of Environmental Economics and Management 85*, 110–129.

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), 246–267.

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), 302–322.

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), 779–782.

Olsen, R. J. (1980). A least-squares correction for selectivity bias. *Econometrica 48*(7), 1815–1820.

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*.

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), 233–249.

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), 243–270.

Terza, J. V. (1998). Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *Journal of Econometrics 84*(1), 129–154.

Terza, J. V. (2009). Parametric nonlinear regression with endogenous switching. *Econometric Reviews 28*(6), 555–580.

Vella, F. (1998). Estimating models with sample selection bias: a survey. *Journal of Human Resources*, 127–169.

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*, 156–164.

Walsh, P. J., J. W. Milon, and D. O. Scrogin (2011). The spatial extent of water quality benefits in urban housing markets. *Land Economics 87*(4), 628–644.

Walters, K., D. A. Christakis, and D. R. Wright (2018). Are mechanical turk worker samples representative of health status and health behaviors in the u.s.? *PLOS ONE 13*(6), e0198835.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* MIT Press.-2002.

Yuan, Y., K. J. Boyle, and W. You (2015). Sample selection, individual heterogeneity, and regional heterogeneity in valuing farmland conservation easements. *Land Economics 91*(4), 627–649.

Complete Table of Contents for paper, plus Online Appendices to accompany:

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

Contents

List of Figures

List of Tables

Appendices
(Online supplementary material)

A Birding activity variables for qBus and engagement intensity options for both qBus and eBird samples

As detailed in Table 1 in the body of the paper, there are a number of notable differences between our two samples. For the annual number of days with trips of more than one mile to observe birds, we defined bins roughly according to decides of the qBus distribution between 1 and 364 days. Only 44% of the qBus sample responds to this question, but we can construct this variable for 77% of the eBird sample.[30] Our eBird respondents are less likely to have taken zero such trips, and more likely to claim to have traveled to see birds all 365 days of the year.

Our eBird member survey respondents are more likely to have participated in the Audubon Christmas Bird Count, and they are much less likely to hunt birds. They are somewhat more likely to be female and to be older.[31] A considerably larger share of the eBird member survey sample did not provide any income data (29.6 percent). Everyone in the eBird member survey sample is from the states of Washington and Oregon, whereas the qBus sample is nationwide. Compared to qBus respondents, more than twice as many eBird member survey respondents are retired. Finally, the eBird member survey sample reports higher educational attainment. All of these differences point to empirical evidence of systematic selection on observables, so that selection on unobservable factors is also likely to be a concern.

A summary of the engagement levels elicited in our two samples is provided in Table A2.

We note in the body of this paper that the Qualtrics Omnibus (qBus) survey we used to gather our general-population data has been discontinued. This appendix also includes Table A3, which lists selected survey research firms currently offering Omnibus surveys. The prices quoted are accurate for August of 2020, but may be adjusted over time by these firms. Surveys can be distributed to Mechanical Turk to a wide range of people, but mTurk samples are understood not to be representative of the general population, as confirmed recently by Walters et al. (2018), for example.

[30] To construct an analogous distribution for our eBird sample, we combine their actual number of days with submitted birding checklists over the preceding twelve months with their self-report as to what fraction of their bird sightings they report to eBird. The documented information about their actual trips distinguishes this constructed explanatory variable from the engagement intensities that form the outcome variable.

[31] They are also more likely to identify as White. However, the proportions of Black and Asian and Hispanic eBird respondents are all less than 1 percent, so we will not use the Race or Ethnicity indicators in our specifications.

Table A1: Share of 1050 qBus respondents reporting at least one day over the last year of engagement with the following wild-bird-related activities (response format: slider with labels at 0, 61, 122, 183, 243, 304 and 365 days); mean days per year and lower and upper quartiles of days. We find these counts to be rather high. This may be an artifact of using the Qualtrics' sliders to elicit numbers of days.

Description	N	At least 1 day	Mean days/ year	Lower quartile (days)	upper quartile (days)
Positive non-consumptive engagement with wild birds:					
Pause what you are doing to observe wild birds	999	0.881	92.8	10	161
Put out food for wild birds	960	0.783	98.4	3	181
Seek opportunities to learn more about wild birds	940	0.728	69.7	0	112
Photograph wild birds	945	0.747	68.5	0	107
Visit public parks/areas less than one mile from home to see, photograph or feed wild birds	926	0.703	67.2	0	109
Travel more than one mile from home to see, photograph or feed wild birds	914	0.667	62.4	0	90
— Any days, any of the above?	1050	0.878	-	-	-
Other interactions with wild birds:					
Employ measures to keep wild birds from harming your garden or property	910	0.624	60.3	0	92
Hunt wild birds for sport or for food	895	0.517	49.7	0	48

Table A2: Definitions and values of two eBird engagement variables for our two samples. CS is the dependent variable for the binary probit selection model; $CS6$ is the dependent variable for the six-level ordered-probit selection model (where $CS6$ degenerates to a four-level model for the eBird sample used alone).

eBird engagement bins	CS values	$CS6$ values	Observed for qBus general population sample?	Observed for eBird citizen science sample?
Does not know eBird	0	1	Y	N
Knows eBird, not a member	0	2	Y	N
Member, reports rarely	1	3	Y	Y
Member, reports $<$ half	1	4	Y	Y
Member, reports $>$ half	1	5	Y	Y
Member, reports almost all	1	6	Y	Y

Table A3: Selected Omnibus-type surveys (alphabetical), other than the Qualtrics Omnibus, available as of 8/2020

Survey firm	Product	Pricing (ca. 2020)	Available sociodemographic variables
Abacus Data (Canada)	National Omnibus Survey; At least 1,500 Canadian adults interviewed monthly; representative samples from large panels, statistical weighting according to the Census. Customization available: provincial oversamples, target audiences, etc.	1 to 3 questions: $1,000 per question	Included: Demographics (age, gender, education), household income, employment status and union membership, community type (urban, suburban, rural), Federal vote intention, 2015 federal vote choice.

Continued on next page

Table A3 – continued from previous page

AmeriSpeak (every two weeks)	AmeriSpeak Omnibus: 1000 responses; data file including weights and AmeriSpeak Omnibus Profile variables	Minimum of 3 survey question units: Per question: $1,000 (e.g. checklist with up to 10 response categories; grid questions using a rating scale with up to four attribute statements)	Included: One standard demographic banner table: age, gender, education, race/ethnicity, HH income (4 categories). For $300 extra: an additional profile variable (contact for available variables) All profile variables: Gender*, Age, Age (4 categories)*, Age (7 categories), Education (4 categories)*, Education (14 categories), Race/Ethnicity*, Household Size, Housing Type, Ownership of Living Quarters, Household Income (18 categories)*, Marital Status, Internet Access, Metropolitan Statistical Area Status, Region (U.S. Census - 4 categories), Region (U.S. Census - 9 categories), State, Household members, age 0-1, Household members, age 2-5, Household members, age 6-12, Household members, age 13-17, Household members, age 18+, Current Employment Status, Survey Start (date/time), Survey End (date/time), Survey Duration (minutes), Survey Mode (online/phone), Device Type (used to take survey). Note: * = Demographics variables included on the standard banner table
Drive Research	Census-representative sample; Excel file of raw responses in csv or SPSS format	1 to 2 questions, n=1000, $3,500; n-2000, $5,000; Additional questions at $200/question	Included: Up to four sociodemographic add-ons at no additional charge. Standard menu of sociodemographic questions: age, gender, household income, marital status, ethnicity, children in the household, education, employment, or region/state in the U.S. (If there is a question not on this list, they can check their respondent database/profiles to see if they can access it.)

Continued on next page

Table A3 – continued from previous page

Ipsos	KP Weekly Omnibus consists of 1,000 adults ages 18+; English only, fields weekly.	$1,000 each for questions 1-5; $800 each for questions 6-10; $400 each for questions 11+	Standard Profiling Variables (provided at no additional cost): Age, Education (highest degree received), Race/Ethnicity, Gender, Household Head, Household Size, Housing type, HH income - profile and imputed, Marital status, MSA Status (live in metro area or rural), Region 4 - Based On State Of Residence, Ownership status of living quarters, State, Total number of HH members age 1 or younger, Total number of HH members age 2 to 5, Total number of HH members age 6 to 12, Total number of HH members age 13 to 17, Total number of HH members age 18 or older.
QuestionPro	?	?	?
SurveyMonkey	?	?	?
YouGov (every 2 weeks, nationally representative)	Academic Omnibus: 1000 responses; codebook, dataset, with appended profile and weight variable	Setup: $500; Each single choice question: $500; Each 3-item matrix question: $750	Included: Birth year; Gender; Race; Education; Employment; Marital Status; Household Income; State of residence. For $500 extra, political demographics: vote registration, 2016 vote, party id, ideology, news interest, and the Pew religion battery

B Review of the usual context for Heckman's two-stage binary selection correction

For the $i = 1, ..., N$ individuals in our general population (qBus) sample, we have observations for some people who are members of eBird and other observations for other people who are not. For everyone, we have conformable variables on sociodemographics and income, Z_i, that we will use to explain eBird participation or non-participation, where respondents $i = 1, ..., r$ participate in eBird and respondents $i = s, ..., N$ do not:

$$CS_i = \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, Z_i = \begin{bmatrix} Z_{11} & \dots & Z_{k1} \\ \vdots & & \vdots \\ Z_{1r} & \dots & Z_{kr} \\ Z_{1s} & \dots & Z_{ks} \\ \vdots & & \vdots \\ Z_{1N} & \dots & Z_{kN} \end{bmatrix}$$

For the qBus sample, we model the latent propensity to be a member of eBird as $CS_i^* = Z_i\gamma + \eta_i$. We only have this data for the qBus sample, so we cannot estimate a second model to explain the outcome variable of interest, $y_i = X_i\beta + \epsilon$, because there are no data for our y_i variable of interest in the qBus sample.[32]

B.0.1 Binary selection and the eBird CS sample

For the $j = 1, ..., J$ observations from our eBird member survey sample, we have Z_j variables that conform to the Z_i variables in the qBus sample, but we have no information about anyone for whom $CS_j = 0$ (i.e. everyone in this sample is a member of eBird). In this case, the selection process 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 this case, the radius of the individual's consideration set, namely their maximum one-way distance for a one-day birding trip), and regressors, X_j, to explain this outcome, where this information is not available for the qBus sample:

$$CS_j = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, Z_j = \begin{bmatrix} Z_{11} & \dots & Z_{k1} \\ \vdots & & \vdots \\ Z_{1J} & \dots & Z_{kJ} \end{bmatrix}, y_j = \begin{bmatrix} y_1 \\ \vdots \\ y_J \end{bmatrix}, X_j = \begin{bmatrix} X_{11} & \dots & X_{m1} \\ \vdots & & \vdots \\ Z_{1J} & \dots & Z_{mJ} \end{bmatrix}$$

For our eBird member survey sample, we assume the underlying relationship between CS and the Z variables is identical to the analogous relationship in the qBus sample. Our proposed selection-correction method will be appropriate if the identical γ and β parameters *would* apply in this eBird sample (and the same $\sigma_\eta, \sigma_\epsilon, \rho$, as well). If the selection equation *could* be estimated for the $j = 1, ..., J$ observations in the eBird member survey sample, the

[32]That variable is collected only in our separate eBird member survey sample. A researcher could simply pose all the questions on our eBird member survey to a large sample of respondents from the general population. Then this would become a standard sample selection story which we outline in Appendix B.

relevant pair of equations would be:

$$\begin{aligned} CS_j^* &= Z_j\gamma + \eta_j \\ y_j &= X_j\beta + \epsilon_j \\ (\eta_j, \epsilon_j) &\sim BVN(0, 0, \sigma_\eta, \sigma_\epsilon, \rho) \end{aligned} \tag{3}$$

Of course, this joint model cannot be estimated because the observed CS_j variable is constant (at 1) in the data from our eBird member survey. For the qBus general population panel, consider a binary indicator for eBird participation CS_i and a set of available regressors Z_i, for a representative sample. If we also had data, for these eBird participants, on an outcome variable of interest, y_i, and a set of regressors, X_i for a subset of this same sample, we would proceed as follows. Suppose the latent propensity to participate in eBird in this qBus sample is a linear-in-parameters function of the Z_i variables, $CS_i^* = Z_i\hat{\gamma} + \eta_i$, then the standard Heckman two-step sample-selection correction procedure involves two terms constructed from $Z_i\hat{\gamma}$.[33] Define:

$$\begin{aligned} \lambda\left(\alpha_{CS_i}\right) &= \lambda\left(-Z_i\hat{\gamma}\right) = \frac{\phi\left(-Z_i\hat{\gamma}\right)}{1 - \Phi\left(-Z_i\hat{\gamma}\right)} = \frac{\phi\left(Z_i\hat{\gamma}\right)}{\Phi\left(Z_i\hat{\gamma}\right)} \\ \delta\left(\alpha_{CS_i}\right) &= \delta\left(-Z_i\hat{\gamma}\right) = \lambda\left(-Z_i\hat{\gamma}\right)\left[\lambda\left(-Z_i\hat{\gamma}\right) - \left(-Z_i\hat{\gamma}\right)\right] \end{aligned} \tag{4}$$

With sample selection, the conditional expected value and the error variance of the outcome variable y_i are no longer given simply by $E\left[y_i\right] = X_i\beta$ and $Var\left[y_i\right] = \sigma_y^2$. Instead, we need the expected value and variance of the *marginal* distribution of y_i *conditional* on y_i being observed (i.e. when $CS_i = 1$). If we can assume that the latent propensity variable CS_i^* is distributed bivariate normal with the outcome variable y_i, but the joint distribution is truncated below at $-Z_i\hat{\gamma}$ in the CS_i^* dimension, the formulas for the expected value and variance of the relevant marginal distribution of y_i for this singly truncated bivariate normal distribution are as follows, as in Greene (2012, p. 836):

$$\begin{aligned} E\left[y_i|y_i\ observed\right] &= E\left[y_i|CS_i^* > -Z_i\hat{\gamma}\right] = X_i\beta + \rho\sigma_\epsilon\lambda\left(-Z_i\hat{\gamma}\right) = X_i\beta + \beta_\lambda\lambda\left(-Z_i\hat{\gamma}\right) \\ Var\left[y_i|y_i\ observed\right] &= Var\left[y_i|CS_i^* > -Z_i\hat{\gamma}\right] = \sigma_y^2\left[\left(1 - \rho^2\delta\left(-Z_i\hat{\gamma}\right)\right)\right] \end{aligned} \tag{5}$$

These formulas provide the rationale for the Heckman two-step approach and why, once this augmented second-stage model has been estimated, we would have unbiased estimates of the expected value of y_i when y_i is observed under the counterfactual conditions where the correlation between the errors in these two equations is zero. For *uncorrelated* bivariate normal variables, the conditional distributions are everywhere equal to the marginal distribution, so we want to *simulate* the absence of any such error correlation. Based on the

[33]Typically, however, attention is focused primarily on the λ term.

augmented regression model, therefore, we can set $\rho = 0$ to get:

$$\begin{aligned} E\left[y_i|y_i\ observed\right] &= X_i\beta + (0\sigma_\epsilon)\lambda\left(-Z_i\hat{\gamma}\right) = X_i\beta \\ Var\left[y_i|y_i\ observed\right] &= \sigma_y^2\left[(1-(0)^2\delta\left(-Z_i\hat{\gamma}\right)\right] = \sigma_y^2 \end{aligned} \tag{6}$$

B.0.2 Sample-Selection Correction in the Related Literature

To date, the non-market valuation literature for environmental goods has focused mostly on correction strategies for when some surveys are not returned at all (called "unit" non-response), or for when the researcher cannot use some responses because those surveys are incomplete and one or more key variables are missing (called "item" non-response).

Standard econometric sample selection correction methods are familiar in the case of continuous outcome variables, as reviewed by Vella (1998) and Wooldridge (2002). However, sample-selection correction methods for multiple discrete outcomes are not particularly well developed in the environmental literature.[34] For conditional logit discrete-choice outcome models, Johnston and Abdulrahman (2017) use an ad hoc approach that builds on earlier work by Cameron and DeShazo (2013) to adjust for response propensity. Kolstoe and Cameron (2017) and Kolstoe et al. (2018) also use this approach, but employ the method to correct only for the individual's propensity to be in the estimating sample drawn from the population of eBird members, not the propensity to be an eBird member in the first place (see the Online Appendix from Kolstoe and Cameron (2017) for details).

Yuan et al. (2015) use a binary probit model to explain systematic selection into their estimating sample and compute a Heckman-style inverse Mills ratio (IMR). This IMR is used as a regressor in their second-stage conditional logit choice model, to shift the coefficient on the status-quo alternative in their choice sets.[35] However, a simple IMR term is appropriate only when the latent selection propensity variable and the (possibly transformed) outcome variable have a bivariate normal distribution. [36]

[34]Terza (2009) proposes a strategy for multinomial (multi-index) models, but does not illustrate his approach for the conditional logit models relevant to destination choice models or stated-preference choice experiments common in the environmental literature.

[35]Given that the IMR derived from the selection model is individual-specific but does not vary across alternatives, including it in the utility-difference "index" that underpins a conditional logit model requires that the IMR term be interacted with at least one regressor that actually *does* vary across alternatives. A status-quo indicator is one such variable.

[36]The Heckman logic for using an IMR thus does not apply when the outcome model is a conditional logit specification—one cannot appeal to the usual bivariate normality assumption for the errors in the two equations to argue that the inclusion of this IMR variable in the outcome equation precisely solves the problem of selection bias. Given that the bivariate normality assumption is untenable in the case of a conditional logit outcome model, there is no good argument for converting the selection propensity into an IMR term. In the present paper, however, we have a latent outcome variable that is plausibly normally distributed. In the observable data, however, this variable is reported in brackets, so an interval-data model based on a normal distribution for the logarithm of the dependent variable is reasonable.

C Legitimate use of simple Inverse Mills Ratios to correct for sample selection in two-step methods

Over the last several decades, empirical researchers have become accustomed to the idea that estimating a sample-selection model via maximum likelihood methods, calculating the IMR, and including that estimated IMF into the desired "outcome" equation of interest will (somehow) purge the parameters of that outcome equation of any bias due to sample selection. However, it is crucial to remember that the IMR offers an appropriate correction for sample-selection bias only under some very specific conditions. Confidence that "including an IMR term" will "fix" selection bias hinges on the assumption that the selection equation and the outcome equation have error terms that are jointly normally distributed.

The joint normality assumption is critical because the IMR correction derives entirely from the formula for the expected value of a singly truncated bivariate normal distribution. If the conditional distribution latent variable in the selection equation is not normal or the conditional distribution of the dependent variable in the outcome equation is not normal (either observed or censored in some way, perhaps after some transformation), then the needed expected value of the singly truncated joint distribution of the errors in the selection equation and the outcome equation cannot automatically be assumed to be given by the usual IMR formulas.

Ideally, selection and outcome equations should be estimated jointly, in which case a wide variety of joint distributions for the two error terms can be assumed/employed, provided that the joint density can be derived and written down. In some cases, it is convenient to write the conditional joint distributions of the selection propensity and the outcome variable as the product of a conditional distribution and a marginal distribution.[37]

This insight is especially relevant for researchers who wish to estimate conditional logit "outcome" models based on people's choices across alternatives with different attributes. Nothing stops the analyst from estimating a binary probit sample selection model and calculating the usual IMR term from the fitted parameters. However, there is no rigorous statistical rationale for including this fitted IMR term like other respondent characteristics as a variable that might shift one or more slope characteristics or the coefficient on the status quo indicator variable, as is done in Yuan et al. (2015). Some types of joint models where IMR correction terms can make sense, statistically, include the following:

- The usual OLS outcome regression with a continuous dependent variable that is con-

[37]Stata now includes the "heckpoisson" estimator, following Terza (1998). Appropriately, this estimator is available only as a FIML estimator, not as a two-step estimator that relies on an IMR term. Jointly distributed variables that are not both normal have also been used in a FIML model that combines a participation/experience variable (that is distributed either Poisson or zero-inflated-Poisson) with a censored-normal outcome variable is estimated jointly in Cameron and Englin (1997).

ditionally normal, perhaps after some transformation

- A Tobit outcome model (censored anywhere—at the bottom, the top, or both) which involves a partially censored normal propensity variable
- An interval-data outcome variable censored between known thresholds (used in the present paper)
- An ordered-probit outcome model with a normally distributed latent propensity variable
- A censored normal outcome model with different censored points across observations

Simply appending an estimated IMR variable to a second-step outcome equation of interest cannot be assumed to be correct in any of the following cases:

- Count data models: Poisson, negative binomial
- Conditional logit models: fixed or random parameters
- Any other statistical model for the "outcome" equation, where the (perhaps latent) dependent is not conditionally normally distributed (even after transformation)

We note, however, the insights provided in Terza (2009), who describes a general approach to endogenous switching models, endogenous treatment models, and sample selection models. These techniques are extended versions of an approach proposed in Olsen (1980), suitable for within-sample corrections, but they seem not yet to have been widely employed in the empirical literture, especially in environmental economics. Should they become available as pre-coded general commands in commonly used software, these methods would likely be popular. However, they would need to be tailored specially for selection-model transfer exercises such as the two-step illustration in this paper.

D Complementary method: Weights based on predicted engagement intensities

In this section, we focus on the actual eBird members in the qBus sample, comparing their different participation levels to those in the entire eBird sample. With our eBird member survey data, there is the added concern that the *intensity* with which these survey respondents engage with the eBird citizen science project may not be representative of the distribution of eBird engagement propensities in the general population of the U.S. To address this issue, we consider how to develop weights for each of the four levels of participation intensity among these eBird members. We base our weights on the *fitted probabilities* of a respondent being each of the four engagement intensity bins in each sample.

For the qBus sample, we estimate an ordered probit model for all six possible bins and calculate a set of fitted probabilities for each bin for each person, conditional on the Z_i vector for that person. Call these probabilities $\hat{p}_{ki}$ for engagement levels $k = 1, ..., 6$.

We then make two calculations for each respondent in the eBird member survey sample. In the first calculation, we use the six-level engagement-intensity model estimated using the qBus data to predict (for the eBird member survey sample) the individual-specific set of six probabilities associated with each of the six engagement-intensity bins (even through nobody in the eBird sample is in non-participation bins 1 or 2). Call these fitted probabilities $\hat{p}^*_{kj}$ for engagement levels $k = 1, ..., 6$,

In the second calculation, we use the eBird sample independently, with its four possible participation-intensity bins. We estimate a four-level ordered-probit model using just the eBird member survey sample and calculate four fitted probabilities, which we will call $\hat{q}_{kj}$, for engagement levels $k = 3, ..., 6$ represented in that sample.

The next step is to assign weights to each respondent in the eBird dataset. These weights serve to scale the fitted probability of an individual being in their observed engagement-intensity bin to match the fitted probability in the population (i.e. the qBus sample). First, consider a hypothetical case where everyone in the qBus and eBird samples has been drawn from the same general population and people in both samples thus shared the same mixes of characteristics (i.e. had identical joint distributions for their Z variables). Then we would expect, across the two samples, to have roughly the same proportions of people in each engagement-intensity bin. However, since nobody in the eBird sample is observed in bins 1 or 2, we must focus on the portion of the engagement-intensity distribution corresponding to eBird membership. For the qBus sample, we should consider the probabilities of being in bins 3 through 6 for the qBus sample, *conditional* on the probability of being in at least one of those four bins. Thus we define $\hat{p}_{kj} = \hat{p}^*_{kj} / \left(\hat{p}^*_{3j} + \hat{p}^*_{4j} + \hat{p}^*_{5j} + \hat{p}^*_{6j}\right)$.[38]

When we allow for potentially very different joint distributions of the explanatory variables Z_i and Z_j for the engagement-intensity model in the qBus and eBird samples, it is readily apparent that we should not use simply the differing observed *proportions* of people in each bin in the two samples to construct weights to be used in estimating the outcome

[38] An alternative would be to attempt to fit a four-level ordered probit for only the qBus respondents, but there are relatively few eBird members in the qBus sample.

model. Our preferred approach would be more akin to the common method of constructing *exogenous* weights based on age brackets or gender. We wish to allow multiple exogenous factors to affect *expected* levels of engagement intensity for each eBird respondent. Consequently, we weight each observation in the eBird sample by $\hat{p}_{kj}/\hat{q}_{kj}, k = 3, ..., 6$, normalized so that these weights sum to the sample size for the eBird sample. Use of these fitted probabilities recruits all of the exogenous or predetermined factors that capture heterogeneity in response propensities (i.e. the Z_i and Z_j data) to build the empirical weights, rather than just 0/1 group membership indicators.

E Additional complications to estimating IMR: Dealing with missing values for Z_j variables in the eBird sample

E.1 qBus sociodemographic variables have few missing values

Any empirical application of this methodology may have to confront the problem of what to do when there are missing values of some variables in one sample or the other. If the correction is based upon the standard sociodemographic variables available for qBus panel members, the data for those variables can be expected to be relatively complete. Any missing values in the qBus sample might be expected to be missing at random.

If other key variables intended to serve as regressors, Z_i, used in our weighting strategy, are drawn from survey questions posed to qBus participants, it is entirely possible that there may be item non-response for some of those variables. Such is the case in the present study, where our own questions produced the data for the number of days per year on which the respondent traveled more than one mile to see birds. Our own questions also elicited the data for participation in the Christmas Bird count and whether the individual also hunts birds, but we assume in the case of these latter two variables that the few missing responses are quivalent to "no".

E.2 eBird sociodemographics match Census, but have more missing values

Missing values in the citizen-science eBird sample, for the sociodemographic variables that conform to the set available the qBus sample, are likely to be more of a problem. For example, due to time constraints for our survey of eBird members, we elected not to ask about individuals' political ideologies. Had we anticipated being able to employ qBus questions to build sampling weights and estimated response propensities, it would have been prudent to be sure that the citizen science members were asked *every* standard sociodemographic question, verbatim, that is available with the qBus responses.

For this first example of our procedure, we can assemble conformable measures for gender, race, ethnicity, broad income brackets, four regions of the U.S., employment status and educational attainment. Some aggregation of categories has been required in each sample to produce matching categories. In future applications of this method, it would be prudent to minimize this type of aggregation. In the eBird data, we used categories that matched the U.S. Census, which would facilitate more-conventional comparisons of marginal distributions in the eBird sample to marginal distributions in the general population. However, the U.S. Census does not provide any information about engagement in citizen science, so our special-purpose qBus general-population sample is much superior in that way.

E.3 Using maximal available Z_j regressors for each eBird observation

Suppose there were no data in the eBird sample on any of the same sociodemographic regressors, Z_i provided by the qBus sample. There would still be valuable information in the qBus sample that could help construct either probability weights or propensity corrections. If one runs an ordered probit model to explain the engagement outcome in the eBird data, but use *no* explanatory variables, the result is a set of estimates for only the three cut-points between the four outcome levels in in that eBird data. If one then calculates the predicted probabilities for each of the four participation intensities, the means of these probabilities, across the sample, match the proportions of the sample observed at each level.[39]

E.4 If there are no Z_j regressors available for some eBird respondents

If there were no Z_j regressors available for some (small) subset of observations in the eBird sample, the best available option for weighting the observations at each level of participation intensity would be derived solely from (a) the predicted probabilities for each of the four relevant participation-intensity levels in the qBus sample (also estimated without regressors) relative to (b) the analogous predicted probabilities for the same four participation intensity levels in the qBus sample. The implicit model being used to predict participation intensities, in that case, would have no Z regressors, so there would be no basis for observable systematic heterogeneity in these probabilities. The weights would then differ only across the four observed participation intensity levels, but would be the same for every person who had no available Z variables in the eBird sample.

E.5 If only some subset of Z_j regressors is available for some eBird respondents

The most-general approach to weighting by participation intensity level or correcting parameters for different-from-average participation intensity would exploit the maximum information available in both samples, on an observation-by-observation basis for the eBird sample. To simplify, assume that only three basic factors are available as explanatory variables. In practice, each factor may be captured by a set of indicators for the categories of that factor, but we will assume for now that there is one continuous variable per factor such that the universe of potential Z variables consists of Z_1, Z_2, and Z_3. All three variables (standing in for groups of indicator variables) are available for each qBus observation, but different

[39]For binary probit and logit models, the means of the fitted probabilities will be either extremely close to the observed proportions, or exactly equal to those proportions, as can be proven by the algebra of the first-order conditions for the maximum likelihood estimation algorithm.

Table A4: Ordered-probit engagement-level models with minimum heterogeneity required to accommodate entire eBird sample: estimated using qBus sample, full eBird sample. Region variable is "West" for all eBird respondents.

	Ordered probit qBus data		Ordered probit eBird data	
Has participated in CBC	2.202***	(0.0670)	0.487***	(0.0678)
Hunts birds	0.526***	(0.0504)	0.0400	(0.128)
Region: Northeast	0.170**	(0.0697)	- [a]	
Region: Midwest	-0.0668	(0.0713)	- [a]	
Region: South	-0.0335	(0.0623)	- [a]	
cut1	1.271***	(0.0525)	-0.0249	(0.0527)
cut2	1.814***	(0.0575)	0.715***	(0.0552)
cut3	2.136***	(0.0622)	1.320***	(0.0623)
cut4	2.527***	(0.0687)	- [b]	
cut5	3.142***	(0.0819)	- [b]	
Observations	4161		1081	
Max. log-likelihood	-2591.22		-1396.30	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ [a]Var = 0 for all. [b]Only 4 levels.

observations in the eBird sample have missing values for either one, two, or all three of these variables.

To fully exploit the available information, it is necessary to estimate an array of models for the qBus sample so that one of these models will be appropriate to transfer to every observation in the eBird sample. Suppose that we have indicators for the presence or absence of values for each of these three Z variables in the eBird sample. The number of necessary models using the qBus data could then be calculated using the sum of all the relevant combinations:

$$C_0^3 + C_1^3 + C_2^3 + C_3^3 = 1 + 3 + 3 + 1 = 8 \tag{7}$$

Of course, as the number of potential factors increases, the number of potentially relevant models to explain participation intensities in the qBus data can increase dramatically. In this study, we have six different factors with complete data in the qBus sample but missing data for at least some observations in the eBird sample: gender, age, income, region, employment status, and education. The number of potentially relevant models could be 64, but due to the correlation between missing values for some of these factors, the actual number of models required is only 30 in this study.

F Six-level ordered-probit eBird engagement-level models for all relevant subsets of regressors (using data from the qBus Sample); needed to compute weights, as well as to predict IMRs for eBird member survey sample

Table A5: qBus sample: Model 1-3 (of 30) to accommodate eBird missing values

	Model 1	Model 2	Model 3
Engagement-level indicator			
Has participated in CBC	2.202***	2.144***	2.120***
	(0.0670)	(0.0679)	(0.0682)
Hunts birds	0.526***	0.520***	0.526***
	(0.0504)	(0.0509)	(0.0514)
Region: Northeast	0.170**	0.162**	0.145**
	(0.0697)	(0.0705)	(0.0707)
Region: Midwest	-0.0668	-0.0596	-0.0580
	(0.0713)	(0.0721)	(0.0722)
Region: South	-0.0335	-0.0190	-0.0135
	(0.0623)	(0.0631)	(0.0633)
Empl. status: Part time		0.0240	0.0508
		(0.0674)	(0.0684)
Empl. status: Looking for work		-0.133	-0.106
		(0.101)	(0.102)
Empl. status: Unemployed		-0.151**	-0.120
		(0.0711)	(0.0737)
Empl. status: Retired		-0.632***	-0.638***
		(0.0800)	(0.0805)
Education: High school			0.0424
			(0.0688)
Education: Some college			-0.102
			(0.0628)
Education: Masters degree			0.244***
			(0.0784)
Education: Doctoral degree			0.209*
			(0.118)
/			

Continued on next page

Table A5 – continued from previous page

cut1	1.271***	1.150***	1.171***
	(0.0525)	(0.0581)	(0.0682)
cut2	1.814***	1.704***	1.728***
	(0.0575)	(0.0626)	(0.0722)
cut3	2.136***	2.032***	2.061***
	(0.0622)	(0.0670)	(0.0762)
cut4	2.527***	2.431***	2.465***
	(0.0687)	(0.0729)	(0.0816)
cut5	3.142***	3.054***	3.096***
	(0.0819)	(0.0852)	(0.0931)
Observations	4161	4161	4161
Max. log-likelihood	-2591.22	-2553.34	-2541.62

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A6: qBus sample: Model 4-6 (of 30) to accommodate eBird missing values

	Model 4	Model 5	Model 6
Engagement-level indicator			
Has participated in CBC	2.023***	2.166***	2.132***
	(0.0691)	(0.0676)	(0.0680)
Hunts birds	0.508***	0.535***	0.520***
	(0.0522)	(0.0510)	(0.0509)
Age: 24 years or less	0.534***		
	(0.0964)		
Age: 25 to 34 years	0.560***		
	(0.0847)		
Age: 35 to 44 years	0.389***		
	(0.0873)		
Age: 55 to 64 years	-0.239**		
	(0.107)		
Age: 65 years and up	-0.328**		
	(0.131)		
Region: Northeast	0.162**	0.158**	0.166**
	(0.0718)	(0.0701)	(0.0706)
Region: Midwest	-0.0513	-0.0621	-0.0517
	(0.0735)	(0.0714)	(0.0721)
Region: South	0.00627	-0.0248	-0.0109
	(0.0643)	(0.0625)	(0.0632)
Empl. status: Part time	0.00864		0.0495
	(0.0719)		(0.0681)
Empl. status: Looking for work	-0.194*		-0.110
	(0.104)		(0.101)
Empl. status: Unemployed	-0.166**		-0.112
	(0.0755)		(0.0725)
Empl. status: Retired	-0.138		-0.633***
	(0.104)		(0.0799)
Education: High school	-0.0213	0.0227	
	(0.0709)	(0.0663)	
Education: Some college	-0.144**	-0.105*	
	(0.0644)	(0.0617)	
Education: Masters degree	0.269***	0.208***	

Continued on next page

Table A6 – continued from previous page

	(0.0798)	(0.0775)	
Education: Doctoral degree	0.204*	0.162	
	(0.119)	(0.116)	
Gender: Female		-0.0946**	-0.135***
		(0.0472)	(0.0484)
/			
cut1	1.441***	1.226***	1.097***
	(0.0967)	(0.0680)	(0.0609)
cut2	2.021***	1.772***	1.652***
	(0.100)	(0.0720)	(0.0652)
cut3	2.365***	2.098***	1.981***
	(0.103)	(0.0758)	(0.0694)
cut4	2.780***	2.495***	2.382***
	(0.108)	(0.0812)	(0.0750)
cut5	3.421***	3.120***	3.007***
	(0.118)	(0.0928)	(0.0868)
Observations	4161	4161	4161
Max. log-likelihood	-2477.40	-2578.51	-2549.44

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A7: qBus sample: Model 7-9 (of 30) to accommodate eBird missing values

	Model 7	Model 8	Model 9
Engagement-level indicator			
Has participated in CBC	2.111***	2.050***	2.015***
	(0.0683)	(0.0685)	(0.0689)
Hunts birds	0.525***	0.488***	0.505***
	(0.0514)	(0.0516)	(0.0522)
Gender: Female	-0.116**	-0.212***	-0.188***
	(0.0487)	(0.0484)	(0.0488)
Region: Northeast	0.149**	0.187***	0.168**
	(0.0708)	(0.0716)	(0.0719)
Region: Midwest	-0.0513	-0.0368	-0.0382
	(0.0723)	(0.0733)	(0.0735)
Region: South	-0.00677	0.0173	0.0193
	(0.0633)	(0.0641)	(0.0643)
Empl. status: Part time	0.0715		
	(0.0690)		
Empl. status: Looking for work	-0.0876		
	(0.102)		
Empl. status: Unemployed	-0.0878		
	(0.0749)		
Empl. status: Retired	-0.638***		
	(0.0805)		
Education: High school	0.0415		-0.0593
	(0.0688)		(0.0690)
Education: Some college	-0.0957		-0.149**
	(0.0629)		(0.0638)
Education: Masters degree	0.239***		0.266***
	(0.0784)		(0.0799)
Education: Doctoral degree	0.190		0.176
	(0.118)		(0.120)
Age: 24 years or less		0.492***	0.542***
		(0.0926)	(0.0937)
Age: 25 to 34 years		0.574***	0.578***
		(0.0837)	(0.0841)
Age: 35 to 44 years		0.407***	0.399***

Continued on next page

Table A7 – continued from previous page

		(0.0866)	(0.0870)
Age: 55 to 64 years		-0.270***	-0.269**
		(0.105)	(0.105)
Age: 65 years and up		-0.375***	-0.422***
		(0.113)	(0.114)
/			
cut1	1.125***	1.404***	1.390***
	(0.0707)	(0.0894)	(0.0980)
cut2	1.683***	1.979***	1.970***
	(0.0746)	(0.0929)	(0.101)
cut3	2.016***	2.320***	2.316***
	(0.0784)	(0.0963)	(0.104)
cut4	2.422***	2.730***	2.733***
	(0.0835)	(0.101)	(0.109)
cut5	3.055***	3.362***	3.375***
	(0.0947)	(0.111)	(0.119)
Observations	4161	4161	4161
Max. log-likelihood	-2538.76	-2489.65	-2474.54

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8: qBus sample: Model 10-12 (of 30) to accommodate eBird missing values

	Model 10	Model 11	Model 12
Engagement-level indicator			
Has participated in CBC	2.034***	2.006***	2.016***
	(0.0690)	(0.0693)	(0.0691)
Hunts birds	0.495***	0.507***	0.505***
	(0.0517)	(0.0522)	(0.0522)
Gender: Female	-0.192***	-0.176***	-0.184***
	(0.0494)	(0.0497)	(0.0492)
Age: 24 years or less	0.510***	0.543***	0.551***
	(0.0959)	(0.0966)	(0.0947)
Age: 25 to 34 years	0.569***	0.572***	0.581***
	(0.0846)	(0.0849)	(0.0846)
Age: 35 to 44 years	0.400***	0.392***	0.399***
	(0.0871)	(0.0875)	(0.0870)
Age: 55 to 64 years	-0.251**	-0.252**	-0.271***
	(0.107)	(0.107)	(0.105)
Age: 65 years and up	-0.307**	-0.361***	-0.426***
	(0.130)	(0.132)	(0.114)
Region: Northeast	0.187***	0.168**	0.171**
	(0.0717)	(0.0720)	(0.0720)
Region: Midwest	-0.0398	-0.0396	-0.0349
	(0.0734)	(0.0736)	(0.0736)
Region: South	0.0146	0.0183	0.0225
	(0.0643)	(0.0645)	(0.0645)
Empl. status: Part time	0.00659	0.0390	
	(0.0719)	(0.0727)	
Empl. status: Looking for work	-0.204**	-0.169	
	(0.104)	(0.105)	
Empl. status: Unemployed	-0.162**	-0.121	
	(0.0744)	(0.0767)	
Empl. status: Retired	-0.143	-0.116	
	(0.103)	(0.104)	
Education: High school		-0.0276	-0.0472
		(0.0710)	(0.0731)
Education: Some college		-0.138**	-0.147**

Continued on next page

Table A8 – continued from previous page

		(0.0645)	(0.0654)
Education: Masters degree		0.263***	0.270***
		(0.0799)	(0.0810)
Education: Doctoral degree		0.175	0.178
		(0.120)	(0.122)
Income: Less than 25K			-0.0820
			(0.0815)
Income: 25 K to 50 K			-0.0137
			(0.0755)
Income: 75 K to 100 K			-0.0489
			(0.0845)
Income: 100 K or more			-0.0310
			(0.0762)
/			
cut1	1.367***	1.373***	1.367***
	(0.0908)	(0.0985)	(0.109)
cut2	1.943***	1.954***	1.948***
	(0.0942)	(0.102)	(0.112)
cut3	2.284***	2.300***	2.293***
	(0.0975)	(0.105)	(0.115)
cut4	2.696***	2.718***	2.711***
	(0.102)	(0.110)	(0.119)
cut5	3.331***	3.362***	3.352***
	(0.112)	(0.119)	(0.129)
Observations	4161	4161	4161
Max. log-likelihood	-2484.82	-2471.15	-2473.89

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A9: qBus sample: Model 13-15 (of 30) to accommodate eBird missing values

	Model 13	Model 14	Model 15
Engagement-level indicator			
Has participated in CBC	2.032***	2.008***	2.075***
	(0.0691)	(0.0694)	(0.0685)
Hunts birds	0.500***	0.506***	0.0780
	(0.0518)	(0.0523)	(0.0941)
Gender: Female	-0.186***	-0.175***	
	(0.0496)	(0.0499)	
Age: 24 years or less	0.526***	0.545***	
	(0.0966)	(0.0971)	
Age: 25 to 34 years	0.582***	0.572***	
	(0.0850)	(0.0854)	
Age: 35 to 44 years	0.399***	0.393***	
	(0.0871)	(0.0875)	
Age: 55 to 64 years	-0.258**	-0.254**	
	(0.107)	(0.107)	
Age: 65 years and up	-0.322**	-0.366***	
	(0.130)	(0.132)	
Income: Less than 25K	-0.0569	-0.0542	
	(0.0826)	(0.0840)	
Income: 25 K to 50 K	-0.0201	-0.0107	
	(0.0755)	(0.0758)	
Income: 75 K to 100 K	-0.00868	-0.0514	
	(0.0838)	(0.0846)	
Income: 100 K or more	0.0624	-0.0345	
	(0.0729)	(0.0763)	
Region: Northeast	0.188***	0.170**	0.158**
	(0.0718)	(0.0721)	(0.0715)
Region: Midwest	-0.0409	-0.0363	-0.0540
	(0.0736)	(0.0738)	(0.0731)
Region: South	0.0182	0.0203	0.000262
	(0.0644)	(0.0647)	(0.0638)
Empl. status: Part time	0.0255	0.0431	
	(0.0734)	(0.0737)	
Empl. status: Looking for work	-0.178*	-0.163	

Continued on next page

Table A9 – continued from previous page

	(0.106)	(0.106)	
Empl. status: Unemployed	-0.135*	-0.112	
	(0.0778)	(0.0790)	
Empl. status: Retired	-0.122	-0.114	
	(0.104)	(0.105)	
Education: High school		-0.0267	
		(0.0740)	
Education: Some college		-0.141**	
		(0.0658)	
Education: Masters degree		0.269***	
		(0.0810)	
Education: Doctoral degree		0.178	
		(0.122)	
Travel 1+ mile data available			0.354
			(0.688)
Trips 1+ miles = 0			-1.072
			(0.693)
Trips 1+ miles = [1,4)			-0.674
			(0.703)
Trips 1+ miles = [4,7)			-0.892
			(0.707)
Trips 1+ miles = [7,10)			-0.600
			(0.705)
Trips 1+ miles = [10,21)			-0.309
			(0.693)
Trips 1+ miles = [21,41)			-0.284
			(0.695)
Trips 1+ miles = [41,72)			0.0561
			(0.692)
Trips 1+ miles = [72,124)			0.293
			(0.690)
Trips 1+ miles = [124,174)			0.212
			(0.689)
Trips 1+ miles = [174,238)			0.341

Continued on next page

Table A9 – continued from previous page

			(0.689)
Trips 1+ miles = [238,364)			0.634
			(0.689)
Trips 1+ miles = 365			0.579
			(0.714)
/			
cut1	1.386***	1.349***	1.188***
	(0.102)	(0.110)	(0.0558)
cut2	1.963***	1.930***	1.769***
	(0.105)	(0.113)	(0.0610)
cut3	2.304***	2.276***	2.112***
	(0.108)	(0.115)	(0.0658)
cut4	2.718***	2.694***	2.518***
	(0.113)	(0.120)	(0.0722)
cut5	3.354***	3.337***	3.143***
	(0.122)	(0.129)	(0.0849)
Observations	4161	4161	4161
Max. log-likelihood	-2483.54	-2470.79	-2490.38

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A10: qBus sample: Model 16-18 (of 30) to accommodate eBird missing values

	Model 16	Model 17	Model 18
Engagement-level indicator			
Travel 1+ mile data available	0.299	0.242	0.183
	(0.674)	(0.673)	(0.689)
Trips 1+ miles = 0	-1.017	-0.963	-0.921
	(0.679)	(0.678)	(0.694)
Trips 1+ miles = [1,4)	-0.625	-0.485	-0.313
	(0.689)	(0.688)	(0.705)
Trips 1+ miles = [4,7)	-0.852	-0.787	-0.590
	(0.693)	(0.693)	(0.709)
Trips 1+ miles = [7,10)	-0.551	-0.528	-0.442
	(0.691)	(0.691)	(0.708)
Trips 1+ miles = [10,21)	-0.264	-0.207	-0.105
	(0.679)	(0.678)	(0.694)
Trips 1+ miles = [21,41)	-0.240	-0.144	-0.0796
	(0.681)	(0.679)	(0.696)
Trips 1+ miles = [41,72)	0.0970	0.132	0.229
	(0.678)	(0.676)	(0.692)
Trips 1+ miles = [72,124)	0.356	0.415	0.461
	(0.676)	(0.674)	(0.690)
Trips 1+ miles = [124,174)	0.272	0.295	0.335
	(0.675)	(0.674)	(0.690)
Trips 1+ miles = [174,238)	0.382	0.441	0.451
	(0.675)	(0.673)	(0.689)
Trips 1+ miles = [238,364)	0.666	0.711	0.703
	(0.674)	(0.673)	(0.689)
Trips 1+ miles = 365	0.623	0.630	0.649
	(0.700)	(0.698)	(0.714)
Has participated in CBC	2.049***	2.003***	1.959***
	(0.0689)	(0.0696)	(0.0699)
Hunts birds	0.0977	0.0874	0.0524
	(0.0943)	(0.0949)	(0.0962)
Region: Northeast	0.143**	0.137*	0.175**
	(0.0718)	(0.0725)	(0.0731)

Continued on next page

Table A10 – continued from previous page

Region: Midwest	-0.0537	-0.0401	-0.0356
	(0.0733)	(0.0740)	(0.0749)
Region: South	0.00426	0.0225	0.0340
	(0.0640)	(0.0647)	(0.0654)
Education: High school	-0.00150	0.0284	
	(0.0677)	(0.0702)	
Education: Some college	-0.103	-0.0925	
	(0.0631)	(0.0643)	
Education: Masters degree	0.222***	0.250***	
	(0.0792)	(0.0801)	
Education: Doctoral degree	0.176	0.205*	
	(0.119)	(0.120)	
Empl. status: Part time		0.0552	-0.0301
		(0.0698)	(0.0725)
Empl. status: Looking for work		-0.127	-0.246**
		(0.105)	(0.105)
Empl. status: Unemployed		-0.116	-0.212***
		(0.0752)	(0.0747)
Empl. status: Retired		-0.612***	-0.209**
		(0.0832)	(0.105)
Age: 24 years or less			0.491***
			(0.0977)
Age: 25 to 34 years			0.526***
			(0.0861)
Age: 35 to 44 years			0.364***
			(0.0888)
Age: 55 to 64 years			-0.186*
			(0.109)
Age: 65 years and up			-0.198
			(0.131)
/			
cut1	1.185***	1.096***	1.354***
	(0.0684)	(0.0717)	(0.0917)
cut2	1.769***	1.691***	1.963***
	(0.0728)	(0.0759)	(0.0954)
cut3	2.115***	2.043***	2.319***

Continued on next page

Table A10 – continued from previous page

	(0.0770)	(0.0800)	(0.0989)
cut4	2.527***	2.463***	2.741***
	(0.0826)	(0.0853)	(0.104)
cut5	3.160***	3.104***	3.381***
	(0.0941)	(0.0964)	(0.113)
Observations	4161	4161	4161
Max. log-likelihood	-2480.49	-2446.81	-2410.56

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A11: qBus sample: Model 19-21 (of 30) to accommodate eBird missing values

	Model 19	Model 20	Model 21
Engagement-level indicator			
Travel 1+ mile data available	0.127	0.166	0.146
	(0.671)	(0.666)	(0.670)
Trips 1+ miles = 0	-0.867	-0.905	-0.885
	(0.677)	(0.672)	(0.676)
Trips 1+ miles = [1,4)	-0.261	-0.295	-0.276
	(0.688)	(0.683)	(0.686)
Trips 1+ miles = [4,7)	-0.549	-0.577	-0.569
	(0.692)	(0.687)	(0.691)
Trips 1+ miles = [7,10)	-0.396	-0.408	-0.414
	(0.691)	(0.685)	(0.690)
Trips 1+ miles = [10,21)	-0.0601	-0.0922	-0.0801
	(0.677)	(0.672)	(0.676)
Trips 1+ miles = [21,41)	-0.0278	-0.0705	-0.0441
	(0.678)	(0.673)	(0.677)
Trips 1+ miles = [41,72)	0.275	0.243	0.260
	(0.675)	(0.670)	(0.673)
Trips 1+ miles = [72,124)	0.527	0.494	0.510
	(0.673)	(0.668)	(0.672)
Trips 1+ miles = [124,174)	0.397	0.363	0.381
	(0.673)	(0.667)	(0.671)
Trips 1+ miles = [174,238)	0.489	0.447	0.470
	(0.672)	(0.667)	(0.671)
Trips 1+ miles = [238,364)	0.731	0.690	0.711
	(0.671)	(0.666)	(0.670)
Trips 1+ miles = 365	0.697	0.688	0.685
	(0.697)	(0.692)	(0.696)
Has participated in CBC	1.929***	1.943***	1.930***
	(0.0703)	(0.0700)	(0.0704)
Hunts birds	0.0730	0.0678	0.0728
	(0.0964)	(0.0963)	(0.0964)
Age: 24 years or less	0.527***	0.534***	0.534***
	(0.0984)	(0.0964)	(0.0989)

Continued on next page

Table A11 – continued from previous page

Age: 25 to 34 years	0.530***	0.545***	0.536***
	(0.0864)	(0.0861)	(0.0869)
Age: 35 to 44 years	0.358***	0.365***	0.358***
	(0.0892)	(0.0887)	(0.0892)
Age: 55 to 64 years	-0.192*	-0.226**	-0.196*
	(0.110)	(0.107)	(0.110)
Age: 65 years and up	-0.262**	-0.358***	-0.269**
	(0.133)	(0.117)	(0.134)
Region: Northeast	0.156**	0.160**	0.159**
	(0.0734)	(0.0734)	(0.0735)
Region: Midwest	-0.0333	-0.0303	-0.0306
	(0.0751)	(0.0752)	(0.0753)
Region: South	0.0398	0.0453	0.0441
	(0.0656)	(0.0657)	(0.0659)
Empl. status: Part time	0.00414		0.0151
	(0.0733)		(0.0745)
Empl. status: Looking for work	-0.207*		-0.193*
	(0.106)		(0.108)
Empl. status: Unemployed	-0.165**		-0.148*
	(0.0770)		(0.0794)
Empl. status: Retired	-0.177*		-0.167
	(0.106)		(0.107)
Education: High school	-0.0321	-0.0444	-0.0177
	(0.0722)	(0.0743)	(0.0752)
Education: Some college	-0.137**	-0.140**	-0.130*
	(0.0657)	(0.0667)	(0.0671)
Education: Masters degree	0.275***	0.274***	0.273***
	(0.0812)	(0.0824)	(0.0824)
Education: Doctoral degree	0.201*	0.189	0.191
	(0.122)	(0.124)	(0.124)
Income: Less than 25K		-0.111	-0.0713
		(0.0825)	(0.0851)
Income: 25 K to 50 K		-0.0287	-0.0229
		(0.0770)	(0.0773)
Income: 75 K to 100 K		-0.0328	-0.0362

Continued on next page

Table A11 – continued from previous page

		(0.0865)	(0.0866)
Income: 100 K or more		0.00665	0.0000751
		(0.0779)	(0.0781)
/			
cut1	1.357***	1.380***	1.349***
	(0.0997)	(0.110)	(0.111)
cut2	1.970***	1.993***	1.963***
	(0.103)	(0.113)	(0.114)
cut3	2.332***	2.355***	2.324***
	(0.107)	(0.116)	(0.117)
cut4	2.761***	2.782***	2.754***
	(0.111)	(0.121)	(0.122)
cut5	3.409***	3.428***	3.402***
	(0.121)	(0.130)	(0.131)
Observations	4161	4161	4161
Max. log-likelihood	-2396.30	-2400.01	-2395.83

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A12: qBus sample: Model 22-24 (of 30) to accommodate eBird missing values

	Model 22	Model 23	Model 24
Engagement-level indicator			
Travel 1+ mile data available	0.343	0.295	0.241
	(0.679)	(0.668)	(0.666)
Trips 1+ miles = 0	-1.062	-1.013	-0.958
	(0.685)	(0.674)	(0.671)
Trips 1+ miles = [1,4)	-0.664	-0.621	-0.481
	(0.694)	(0.684)	(0.681)
Trips 1+ miles = [4,7)	-0.885	-0.850	-0.787
	(0.698)	(0.687)	(0.686)
Trips 1+ miles = [7,10)	-0.580	-0.539	-0.517
	(0.697)	(0.686)	(0.684)
Trips 1+ miles = [10,21)	-0.289	-0.251	-0.196
	(0.685)	(0.674)	(0.671)
Trips 1+ miles = [21,41)	-0.264	-0.228	-0.132
	(0.686)	(0.675)	(0.673)
Trips 1+ miles = [41,72)	0.0791	0.112	0.146
	(0.683)	(0.672)	(0.669)
Trips 1+ miles = [72,124)	0.303	0.358	0.415
	(0.681)	(0.670)	(0.667)
Trips 1+ miles = [124,174)	0.231	0.283	0.305
	(0.681)	(0.670)	(0.667)
Trips 1+ miles = [174,238)	0.357	0.391	0.448
	(0.680)	(0.669)	(0.667)
Trips 1+ miles = [238,364)	0.649	0.675	0.720
	(0.680)	(0.669)	(0.666)
Trips 1+ miles = 365	0.590	0.629	0.634
	(0.705)	(0.695)	(0.692)
Has participated in CBC	2.060***	2.038***	1.994***
	(0.0688)	(0.0691)	(0.0697)
Hunts birds	0.0756	0.0937	0.0816
	(0.0940)	(0.0943)	(0.0949)
Gender: Female	-0.128***	-0.104**	-0.120**
	(0.0478)	(0.0484)	(0.0498)

Continued on next page

Table A12 – continued from previous page

Region: Northeast	0.162**	0.147**	0.141*
	(0.0716)	(0.0719)	(0.0726)
Region: Midwest	-0.0460	-0.0472	-0.0326
	(0.0732)	(0.0733)	(0.0741)
Region: South	0.00866	0.0109	0.0300
	(0.0639)	(0.0640)	(0.0648)
Education: High school		0.00516	0.0277
		(0.0679)	(0.0703)
Education: Some college		-0.0931	-0.0858
		(0.0632)	(0.0643)
Education: Masters degree		0.215***	0.244***
		(0.0793)	(0.0801)
Education: Doctoral degree		0.156	0.184
		(0.119)	(0.120)
Empl. status: Part time			0.0766
			(0.0705)
Empl. status: Looking for work			-0.108
			(0.105)
Empl. status: Unemployed			-0.0833
			(0.0765)
Empl. status: Retired			-0.610***
			(0.0831)
/			
cut1	1.128***	1.140***	1.051***
	(0.0599)	(0.0714)	(0.0740)
cut2	1.711***	1.725***	1.646***
	(0.0647)	(0.0756)	(0.0781)
cut3	2.054***	2.071***	1.999***
	(0.0691)	(0.0796)	(0.0820)
cut4	2.462***	2.484***	2.421***
	(0.0751)	(0.0849)	(0.0871)
cut5	3.091***	3.120***	3.063***
	(0.0872)	(0.0960)	(0.0979)
Observations	4161	4161	4161
Max. log-likelihood	-2486.77	-2478.16	-2443.92

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A13: qBus sample: Model 25-27 (of 30) to accommodate eBird missing values

	Model 25	Model 26	Model 27
Engagement-level indicator			
Travel 1+ mile data available	0.125	0.170	0.122
	(0.659)	(0.676)	(0.662)
Trips 1+ miles = 0	-0.858	-0.904	-0.856
	(0.665)	(0.681)	(0.668)
Trips 1+ miles = [1,4)	-0.250	-0.296	-0.249
	(0.676)	(0.692)	(0.679)
Trips 1+ miles = [4,7)	-0.527	-0.571	-0.537
	(0.681)	(0.696)	(0.683)
Trips 1+ miles = [7,10)	-0.350	-0.412	-0.373
	(0.679)	(0.695)	(0.682)
Trips 1+ miles = [10,21)	-0.0324	-0.0784	-0.0411
	(0.665)	(0.681)	(0.668)
Trips 1+ miles = [21,41)	-0.0168	-0.0510	-0.00610
	(0.666)	(0.683)	(0.669)
Trips 1+ miles = [41,72)	0.300	0.260	0.300
	(0.663)	(0.679)	(0.665)
Trips 1+ miles = [72,124)	0.533	0.471	0.531
	(0.661)	(0.677)	(0.664)
Trips 1+ miles = [124,174)	0.412	0.359	0.413
	(0.661)	(0.677)	(0.663)
Trips 1+ miles = [174,238)	0.495	0.469	0.501
	(0.660)	(0.676)	(0.663)
Trips 1+ miles = [238,364)	0.741	0.724	0.745
	(0.659)	(0.676)	(0.662)
Trips 1+ miles = 365	0.723	0.664	0.708
	(0.686)	(0.702)	(0.688)
Has participated in CBC	1.923***	1.941***	1.913***
	(0.0701)	(0.0701)	(0.0704)
Hunts birds	0.0591	0.0436	0.0642
	(0.0963)	(0.0963)	(0.0965)
Gender: Female	-0.184***	-0.187***	-0.171***
	(0.0498)	(0.0503)	(0.0507)

Continued on next page

Table A13 – continued from previous page

Age: 24 years or less	0.534***	0.501***	0.535***
	(0.0957)	(0.0980)	(0.0987)
Age: 25 to 34 years	0.549***	0.538***	0.542***
	(0.0857)	(0.0863)	(0.0866)
Age: 35 to 44 years	0.367***	0.366***	0.360***
	(0.0889)	(0.0890)	(0.0894)
Age: 55 to 64 years	-0.232**	-0.198*	-0.204*
	(0.107)	(0.110)	(0.110)
Age: 65 years and up	-0.376***	-0.233*	-0.292**
	(0.117)	(0.132)	(0.134)
Region: Northeast	0.162**	0.180**	0.162**
	(0.0734)	(0.0732)	(0.0735)
Region: Midwest	-0.0204	-0.0222	-0.0212
	(0.0751)	(0.0750)	(0.0752)
Region: South	0.0522	0.0478	0.0520
	(0.0656)	(0.0656)	(0.0658)
Education: High school	-0.0723		-0.0377
	(0.0702)		(0.0723)
Education: Some college	-0.143**		-0.130**
	(0.0651)		(0.0658)
Education: Masters degree	0.271***		0.269***
	(0.0813)		(0.0814)
Education: Doctoral degree	0.170		0.170
	(0.122)		(0.122)
Empl. status: Part time		0.00313	0.0336
		(0.0732)	(0.0740)
Empl. status: Looking for work		-0.219**	-0.183*
		(0.106)	(0.107)
Empl. status: Unemployed		-0.162**	-0.120
		(0.0760)	(0.0782)
Empl. status: Retired		-0.184*	-0.154
		(0.105)	(0.106)
/			
cut1	1.311***	1.285***	1.292***
	(0.101)	(0.0934)	(0.101)
cut2	1.926***	1.896***	1.908***

Continued on next page

Table A13 – continued from previous page

	(0.104)	(0.0971)	(0.105)
cut3	2.289***	2.254***	2.271***
	(0.108)	(0.100)	(0.108)
cut4	2.719***	2.679***	2.703***
	(0.112)	(0.105)	(0.113)
cut5	3.368***	3.322***	3.354***
	(0.121)	(0.114)	(0.122)
Observations	4161	4161	4161
Max. log-likelihood	-2394.41	-2403.62	-2390.64

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A14: qBus sample: Model 28-30 (of 30) to accommodate eBird missing values

	Model 28	Model 29	Model 30
Engagement-level indicator			
Travel 1+ mile data available	0.192	0.148	0.174
	(0.667)	(0.658)	(0.671)
Trips 1+ miles = 0	-0.929	-0.882	-0.912
	(0.673)	(0.664)	(0.677)
Trips 1+ miles = [1,4)	-0.318	-0.269	-0.301
	(0.684)	(0.675)	(0.688)
Trips 1+ miles = [4,7)	-0.579	-0.553	-0.576
	(0.688)	(0.680)	(0.693)
Trips 1+ miles = [7,10)	-0.406	-0.376	-0.414
	(0.687)	(0.678)	(0.691)
Trips 1+ miles = [10,21)	-0.0974	-0.0572	-0.0898
	(0.673)	(0.664)	(0.677)
Trips 1+ miles = [21,41)	-0.0826	-0.0376	-0.0588
	(0.674)	(0.665)	(0.678)
Trips 1+ miles = [41,72)	0.250	0.279	0.263
	(0.671)	(0.662)	(0.675)
Trips 1+ miles = [72,124)	0.455	0.510	0.470
	(0.669)	(0.660)	(0.673)
Trips 1+ miles = [124,174)	0.352	0.390	0.367
	(0.669)	(0.660)	(0.673)
Trips 1+ miles = [174,238)	0.443	0.472	0.463
	(0.668)	(0.659)	(0.672)
Trips 1+ miles = [238,364)	0.700	0.715	0.718
	(0.667)	(0.658)	(0.671)
Trips 1+ miles = 365	0.667	0.707	0.665
	(0.693)	(0.685)	(0.697)
Has participated in CBC	1.947***	1.923***	1.937***
	(0.0700)	(0.0703)	(0.0703)
Hunts birds	0.0422	0.0595	0.0497
	(0.0961)	(0.0963)	(0.0963)
Gender: Female	-0.189***	-0.178***	-0.178***
	(0.0499)	(0.0502)	(0.0506)

Continued on next page

Table A14 – continued from previous page

Age: 24 years or less	0.523***	0.547***	0.521***
	(0.0960)	(0.0966)	(0.0987)
Age: 25 to 34 years	0.566***	0.556***	0.554***
	(0.0858)	(0.0863)	(0.0867)
Age: 35 to 44 years	0.372***	0.367***	0.365***
	(0.0885)	(0.0889)	(0.0890)
Age: 55 to 64 years	-0.237**	-0.235**	-0.208*
	(0.107)	(0.107)	(0.110)
Age: 65 years and up	-0.334***	-0.379***	-0.253*
	(0.115)	(0.117)	(0.132)
Income: Less than 25K	-0.103	-0.0873	-0.0628
	(0.0809)	(0.0829)	(0.0839)
Income: 25 K to 50 K	-0.0368	-0.0224	-0.0296
	(0.0768)	(0.0771)	(0.0771)
Income: 75 K to 100 K	0.00562	-0.0399	0.000688
	(0.0858)	(0.0866)	(0.0859)
Income: 100 K or more	0.0929	-0.00666	0.0838
	(0.0746)	(0.0782)	(0.0748)
Region: Northeast	0.184**	0.166**	0.182**
	(0.0732)	(0.0736)	(0.0734)
Region: Midwest	-0.0215	-0.0177	-0.0234
	(0.0751)	(0.0753)	(0.0752)
Region: South	0.0557	0.0566	0.0525
	(0.0656)	(0.0658)	(0.0658)
Education: High school		-0.0511	
		(0.0745)	
Education: Some college		-0.134**	
		(0.0668)	
Education: Masters degree		0.270***	
		(0.0825)	
Education: Doctoral degree		0.162	
		(0.124)	
Empl. status: Part time			0.0268
			(0.0748)
Empl. status: Looking for work			-0.188*

Continued on next page

			(0.108)
Empl. status: Unemployed			-0.129
			(0.0795)
Empl. status: Retired			-0.155
			(0.106)
/			
cut1	1.345***	1.299***	1.313***
	(0.104)	(0.112)	(0.105)
cut2	1.956***	1.914***	1.925***
	(0.107)	(0.115)	(0.108)
cut3	2.315***	2.277***	2.284***
	(0.110)	(0.118)	(0.111)
cut4	2.740***	2.708***	2.711***
	(0.115)	(0.123)	(0.116)
cut5	3.383***	3.357***	3.356***
	(0.124)	(0.132)	(0.125)
Observations	4161	4161	4161
Max. log-likelihood	-2405.49	-2393.71	-2401.66

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

G Four-level ordered-probit eBird engagement-level models for all relevant subsets of regressors (using data from the eBird member survey sample); needed to compute weights

Table A15: eBird sample: Model 1-3 (of 30) to accommodate missing values

	Model 1	Model 2	Model 3
Engagement-level indicator			
Has participated in CBC	0.487***	0.556***	0.525***
	(0.0678)	(0.0744)	(0.0759)
Hunts birds	0.0400	-0.0868	-0.0443
	(0.128)	(0.139)	(0.140)
Empl. status: Part time		-0.255*	-0.210
		(0.140)	(0.143)
Empl. status: Looking for work		-0.587	-0.562
		(0.425)	(0.428)
Empl. status: Unemployed		-0.154	-0.0841
		(0.157)	(0.161)
Empl. status: Retired		-0.501***	-0.480***
		(0.0807)	(0.0820)
Education: High school			0.197
			(0.219)
Education: Some college			-0.183
			(0.120)
Education: Masters degree			0.0652
			(0.0905)
Education: Doctoral degree			0.266**
			(0.125)
/			
cut1	-0.0249	-0.294***	-0.258***
	(0.0527)	(0.0759)	(0.0967)
cut2	0.715***	0.482***	0.527***
	(0.0552)	(0.0762)	(0.0972)
cut3	1.320***	1.086***	1.131***
	(0.0623)	(0.0818)	(0.102)
Observations	1081	918	899
Max. log-likelihood	-1396.30	-1162.68	-1135.31

t in parentheses

Continued on next page

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A16: eBird sample: Model 4-6 (of 30) to accommodate missing values

	Model 4	Model 5	Model 6
Engagement-level indicator			
Has participated in CBC	0.567***	0.413***	0.522***
	(0.0769)	(0.0700)	(0.0749)
Hunts birds	-0.0758	-0.0716	-0.214
	(0.142)	(0.131)	(0.141)
Age: 24 years or less	0.963***		
	(0.352)		
Age: 25 to 34 years	0.439**		
	(0.172)		
Age: 35 to 44 years	0.330**		
	(0.157)		
Age: 55 to 64 years	-0.140		
	(0.125)		
Age: 65 years and up	-0.163		
	(0.152)		
Empl. status: Part time	-0.222		-0.181
	(0.147)		(0.141)
Empl. status: Looking for work	-0.693		-0.520
	(0.435)		(0.425)
Empl. status: Unemployed	-0.0923		-0.0523
	(0.164)		(0.159)
Empl. status: Retired	-0.260**		-0.469***
	(0.121)		(0.0812)
Education: High school	0.0937	0.102	
	(0.234)	(0.194)	
Education: Some college	-0.135	-0.159	
	(0.122)	(0.108)	
Education: Masters degree	0.126	0.0479	
	(0.0919)	(0.0833)	
Education: Doctoral degree	0.344***	0.227*	
	(0.127)	(0.116)	
Gender: Female		-0.357***	-0.373***
		(0.0717)	(0.0771)
/			
cut1	-0.141	-0.268***	-0.518***

Continued on next page

Table A16 – continued from previous page

	(0.126)	(0.0930)	(0.0889)
cut2	0.666***	0.498***	0.272***
	(0.127)	(0.0934)	(0.0879)
cut3	1.276***	1.113***	0.889***
	(0.131)	(0.0968)	(0.0918)
Observations	898	1053	916
Max. log-likelihood	-1121.46	-1344.47	-1149.42

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A17: eBird sample: Model 7-9 (of 30) to accommodate missing values

	Model 7	Model 8	Model 9
Engagement-level indicator			
Has participated in CBC	0.493***	0.534***	0.500***
	(0.0765)	(0.0700)	(0.0715)
Hunts birds	-0.168	-0.208	-0.162
	(0.143)	(0.133)	(0.133)
Gender: Female	-0.344***	-0.412***	-0.378***
	(0.0787)	(0.0706)	(0.0724)
Empl. status: Part time	-0.144		
	(0.144)		
Empl. status: Looking for work	-0.503		
	(0.427)		
Empl. status: Unemployed	0.0109		
	(0.163)		
Empl. status: Retired	-0.452***		
	(0.0824)		
Education: High school	0.115		0.00592
	(0.220)		(0.207)
Education: Some college	-0.200*		-0.129
	(0.121)		(0.110)
Education: Masters degree	0.0603		0.166*
	(0.0908)		(0.0852)
Education: Doctoral degree	0.198		0.336***
	(0.126)		(0.118)
Age: 24 years or less		0.632**	0.769***
		(0.255)	(0.268)
Age: 25 to 34 years		0.328**	0.389**
		(0.155)	(0.157)
Age: 35 to 44 years		0.172	0.187
		(0.141)	(0.142)
Age: 55 to 64 years		-0.305***	-0.306***
		(0.107)	(0.108)
Age: 65 years and up		-0.431***	-0.439***
		(0.105)	(0.106)
/			
cut1	-0.480***	-0.490***	-0.394***

Continued on next page

Table A17 – continued from previous page

	(0.110)	(0.107)	(0.124)
cut2	0.317***	0.305***	0.413***
	(0.109)	(0.106)	(0.124)
cut3	0.931***	0.939***	1.046***
	(0.112)	(0.109)	(0.127)
Observations	897	1071	1051
Max. log-likelihood	-1124.29	-1339.49	-1307.49

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A18: eBird sample: Model 10-12 (of 30) to accommodate missing values

	Model 10	Model 11	Model 12
Engagement-level indicator			
Has participated in CBC	0.568***	0.535***	0.491***
	(0.0761)	(0.0775)	(0.0795)
Hunts birds	-0.259*	-0.207	-0.186
	(0.143)	(0.144)	(0.140)
Gender: Female	-0.385***	-0.359***	-0.364***
	(0.0779)	(0.0796)	(0.0804)
Age: 24 years or less	0.875***	0.924***	0.789**
	(0.334)	(0.352)	(0.315)
Age: 25 to 34 years	0.391**	0.424**	0.317*
	(0.171)	(0.173)	(0.167)
Age: 35 to 44 years	0.317**	0.321**	0.125
	(0.156)	(0.157)	(0.151)
Age: 55 to 64 years	-0.154	-0.161	-0.376***
	(0.124)	(0.126)	(0.117)
Age: 65 years and up	-0.217	-0.232	-0.405***
	(0.151)	(0.154)	(0.118)
Empl. status: Part time	-0.180	-0.136	
	(0.147)	(0.149)	
Empl. status: Looking for work	-0.663	-0.635	
	(0.433)	(0.435)	
Empl. status: Unemployed	-0.0534	0.00968	
	(0.163)	(0.165)	
Empl. status: Retired	-0.222*	-0.188	
	(0.120)	(0.123)	
Education: High school		0.0161	0.145
		(0.235)	(0.237)
Education: Some college		-0.149	-0.135
		(0.122)	(0.122)
Education: Masters degree		0.125	0.108
		(0.0922)	(0.0943)
Education: Doctoral degree		0.279**	0.251*
		(0.128)	(0.133)
Income: Less than 25K			-0.133

Continued on next page

			(0.170)
Income: 25 K to 50 K			0.118
			(0.117)
Income: 75 K to 100 K			-0.127
			(0.122)
Income: 100 K or more			0.117
			(0.106)
/			
cut1	-0.464***	-0.383***	-0.478***
	(0.119)	(0.137)	(0.154)
cut2	0.347***	0.438***	0.378**
	(0.118)	(0.137)	(0.154)
cut3	0.972***	1.059***	1.017***
	(0.122)	(0.140)	(0.156)
Observations	914	896	853
Max. log-likelihood	-1134.81	-1109.77	-1076.62

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A19: eBird sample: Model 13-15 (of 30) to accommodate missing values

	Model 13	Model 14	Model 15
Engagement-level indicator			
Has participated in CBC	0.562***	0.534***	0.0305
	(0.0846)	(0.0867)	(0.0844)
Hunts birds	-0.290*	-0.241	0.106
	(0.148)	(0.150)	(0.152)
Gender: Female	-0.387***	-0.375***	
	(0.0867)	(0.0888)	
Age: 24 years or less	1.021***	0.986**	
	(0.396)	(0.403)	
Age: 25 to 34 years	0.341*	0.330*	
	(0.185)	(0.188)	
Age: 35 to 44 years	0.233	0.229	
	(0.167)	(0.168)	
Age: 55 to 64 years	-0.195	-0.217	
	(0.134)	(0.136)	
Age: 65 years and up	-0.103	-0.125	
	(0.167)	(0.170)	
Income: Less than 25K	-0.267	-0.198	
	(0.193)	(0.200)	
Income: 25 K to 50 K	0.0989	0.150	
	(0.124)	(0.127)	
Income: 75 K to 100 K	-0.204	-0.182	
	(0.131)	(0.132)	
Income: 100 K or more	0.0531	0.0303	
	(0.115)	(0.117)	
Empl. status: Part time	-0.181	-0.164	
	(0.160)	(0.161)	
Empl. status: Looking for work	-0.476	-0.482	
	(0.462)	(0.465)	
Empl. status: Unemployed	-0.163	-0.127	
	(0.188)	(0.190)	
Empl. status: Retired	-0.343**	-0.321**	
	(0.135)	(0.136)	
Education: High school		0.130	

Continued on next page

Table A19 – continued from previous page

		(0.259)	
Education: Some college		-0.142	
		(0.138)	
Education: Masters degree		0.0764	
		(0.102)	
Education: Doctoral degree		0.196	
		(0.144)	
Travel 1+ mile data available			0
			(.)
Trips 1+ miles = 0			-2.765***
			(0.187)
Trips 1+ miles = [1,4)			-2.923***
			(0.213)
Trips 1+ miles = [4,7)			-1.944***
			(0.222)
Trips 1+ miles = [7,10)			-1.587***
			(0.283)
Trips 1+ miles = [10,21)			-1.556***
			(0.202)
Trips 1+ miles = [21,41)			-1.371***
			(0.216)
Trips 1+ miles = [41,72)			-1.097***
			(0.207)
Trips 1+ miles = [72,124)			-0.695***
			(0.213)
Trips 1+ miles = [124,174)			-0.560**
			(0.231)
Trips 1+ miles = [174,238)			-0.426
			(0.265)
Trips 1+ miles = [238,364)			-0.399*
			(0.229)
Trips 1+ miles = 365			0
			(.)
/			
cut1	-0.599***	-0.555***	-2.506***

Continued on next page

Table A19 – continued from previous page

	(0.156)	(0.174)	(0.187)
cut2	0.285*	0.335*	-1.346***
	(0.156)	(0.173)	(0.178)
cut3	0.908***	0.954***	-0.401**
	(0.159)	(0.176)	(0.173)
Observations	740	727	831
Max. log-likelihood	-924.96	-907.59	-875.39

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A20: eBird sample: Model 16-18 (of 30) to accommodate missing values

	Model 16	Model 17	Model 18
Engagement-level indicator			
Travel 1+ mile data available	0	0	0
	(.)	(.)	(.)
Trips 1+ miles = 0	-2.761***	-2.837***	-2.855***
	(0.190)	(0.211)	(0.208)
Trips 1+ miles = [1,4)	-2.983***	-3.107***	-3.090***
	(0.218)	(0.242)	(0.240)
Trips 1+ miles = [4,7)	-1.977***	-1.959***	-1.955***
	(0.226)	(0.258)	(0.255)
Trips 1+ miles = [7,10)	-1.698***	-1.993***	-1.896***
	(0.290)	(0.314)	(0.309)
Trips 1+ miles = [10,21)	-1.546***	-1.678***	-1.692***
	(0.205)	(0.226)	(0.224)
Trips 1+ miles = [21,41)	-1.368***	-1.439***	-1.453***
	(0.219)	(0.240)	(0.237)
Trips 1+ miles = [41,72)	-1.105***	-1.196***	-1.225***
	(0.209)	(0.232)	(0.232)
Trips 1+ miles = [72,124)	-0.678***	-0.629***	-0.676***
	(0.216)	(0.238)	(0.236)
Trips 1+ miles = [124,174)	-0.592**	-0.629**	-0.622**
	(0.235)	(0.264)	(0.261)
Trips 1+ miles = [174,238)	-0.404	-0.546*	-0.568**
	(0.272)	(0.291)	(0.284)
Trips 1+ miles = [238,364)	-0.410*	-0.522**	-0.551**
	(0.234)	(0.258)	(0.255)
Trips 1+ miles = 365	0	0	0
	(.)	(.)	(.)
Has participated in CBC	0.0272	0.125	0.160*
	(0.0859)	(0.0935)	(0.0936)
Hunts birds	0.108	0.00874	-0.00447
	(0.153)	(0.170)	(0.170)
Education: High school	0.427*	0.462*	
	(0.235)	(0.255)	
Education: Some college	0.00506	-0.114	

Continued on next page

Table A20 – continued from previous page

	(0.134)	(0.150)	
Education: Masters degree	0.176*	0.183*	
	(0.0980)	(0.108)	
Education: Doctoral degree	0.147	0.0672	
	(0.133)	(0.144)	
Empl. status: Part time		-0.116	-0.226
		(0.169)	(0.173)
Empl. status: Looking for work		-0.718	-0.976*
		(0.587)	(0.589)
Empl. status: Unemployed		0.0905	0.0520
		(0.186)	(0.185)
Empl. status: Retired		-0.322***	-0.342**
		(0.0976)	(0.142)
Age: 24 years or less			0.755**
			(0.346)
Age: 25 to 34 years			0.206
			(0.194)
Age: 35 to 44 years			0.281
			(0.180)
Age: 55 to 64 years			-0.110
			(0.145)
Age: 65 years and up			0.130
			(0.178)
/			
cut1	-2.420***	-2.670***	-2.711***
	(0.202)	(0.231)	(0.235)
cut2	-1.243***	-1.434***	-1.457***
	(0.193)	(0.221)	(0.225)
cut3	-0.303	-0.496**	-0.510**
	(0.188)	(0.215)	(0.220)
Observations	810	693	707
Max. log-likelihood	-850.08	-711.03	-722.39

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A21: eBird sample: Model 19-21 (of 30) to accommodate missing values

	Model 19	Model 20	Model 21
Engagement-level indicator			
Travel 1+ mile data available	0	0	0
	(.)	(.)	(.)
Trips 1+ miles = 0	-2.867***	-2.810***	-2.850***
	(0.212)	(0.207)	(0.228)
Trips 1+ miles = [1,4)	-3.151***	-3.071***	-3.140***
	(0.245)	(0.242)	(0.267)
Trips 1+ miles = [4,7)	-2.024***	-2.120***	-1.884***
	(0.259)	(0.250)	(0.286)
Trips 1+ miles = [7,10)	-2.086***	-2.084***	-2.285***
	(0.318)	(0.346)	(0.370)
Trips 1+ miles = [10,21)	-1.707***	-1.800***	-1.851***
	(0.227)	(0.226)	(0.249)
Trips 1+ miles = [21,41)	-1.484***	-1.494***	-1.520***
	(0.241)	(0.238)	(0.261)
Trips 1+ miles = [41,72)	-1.247***	-1.341***	-1.363***
	(0.234)	(0.227)	(0.253)
Trips 1+ miles = [72,124)	-0.675***	-0.844***	-0.728***
	(0.239)	(0.233)	(0.258)
Trips 1+ miles = [124,174)	-0.677**	-0.696***	-0.654**
	(0.265)	(0.258)	(0.292)
Trips 1+ miles = [174,238)	-0.565*	-0.465	-0.569*
	(0.292)	(0.298)	(0.321)
Trips 1+ miles = [238,364)	-0.575**	-0.396	-0.424
	(0.260)	(0.260)	(0.286)
Trips 1+ miles = 365	0	0	0
	(.)	(.)	(.)
Has participated in CBC	0.150	0.103	0.183*
	(0.0948)	(0.0972)	(0.106)
Hunts birds	-0.0170	0.0233	-0.0400
	(0.171)	(0.164)	(0.177)
Age: 24 years or less	0.640*	0.902***	1.019**
	(0.370)	(0.334)	(0.423)
Age: 25 to 34 years	0.267	0.208	0.298

Continued on next page

Table A21 – continued from previous page

	(0.196)	(0.191)	(0.215)
Age: 35 to 44 years	0.301*	0.308*	0.325*
	(0.181)	(0.174)	(0.192)
Age: 55 to 64 years	-0.128	-0.300**	-0.117
	(0.148)	(0.137)	(0.159)
Age: 65 years and up	0.102	-0.184	0.161
	(0.181)	(0.137)	(0.200)
Empl. status: Part time	-0.189		-0.169
	(0.175)		(0.187)
Empl. status: Looking for work	-0.908		-0.825
	(0.593)		(0.650)
Empl. status: Unemployed	0.0780		-0.0457
	(0.189)		(0.214)
Empl. status: Retired	-0.307**		-0.343**
	(0.145)		(0.162)
Education: High school	0.392	0.606**	0.718**
	(0.275)	(0.280)	(0.308)
Education: Some college	-0.0874	0.0891	0.00245
	(0.152)	(0.150)	(0.169)
Education: Masters degree	0.222**	0.305***	0.259**
	(0.110)	(0.110)	(0.121)
Education: Doctoral degree	0.116	0.160	0.0245
	(0.147)	(0.150)	(0.164)
Income: Less than 25K		0.0219	-0.0785
		(0.209)	(0.254)
Income: 25 K to 50 K		0.154	0.154
		(0.138)	(0.150)
Income: 75 K to 100 K		0.0584	0.00504
		(0.144)	(0.157)
Income: 100 K or more		0.240*	0.167
		(0.123)	(0.136)
/			
cut1	-2.632***	-2.483***	-2.586***
	(0.250)	(0.249)	(0.282)
cut2	-1.370***	-1.211***	-1.246***
	(0.240)	(0.241)	(0.273)

Continued on next page

Table A21 – continued from previous page

cut3	-0.425*	-0.241	-0.302
	(0.236)	(0.236)	(0.268)
Observations	692	674	573
Max. log-likelihood	-704.13	-697.49	-583.38

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A22: eBird sample: Model 22-24 (of 30) to accommodate missing values

	Model 22	Model 23	Model 24
Engagement-level indicator			
Travel 1+ mile data available	0	0	0
	(.)	(.)	(.)
Trips 1+ miles = 0	-2.735***	-2.732***	-2.810***
	(0.188)	(0.191)	(0.211)
Trips 1+ miles = [1,4)	-2.886***	-2.952***	-3.082***
	(0.214)	(0.219)	(0.243)
Trips 1+ miles = [4,7)	-1.922***	-1.956***	-1.933***
	(0.223)	(0.226)	(0.258)
Trips 1+ miles = [7,10)	-1.545***	-1.659***	-1.948***
	(0.284)	(0.291)	(0.315)
Trips 1+ miles = [10,21)	-1.514***	-1.512***	-1.650***
	(0.203)	(0.206)	(0.227)
Trips 1+ miles = [21,41)	-1.334***	-1.341***	-1.416***
	(0.217)	(0.220)	(0.240)
Trips 1+ miles = [41,72)	-1.089***	-1.098***	-1.189***
	(0.206)	(0.209)	(0.232)
Trips 1+ miles = [72,124)	-0.668***	-0.658***	-0.611**
	(0.213)	(0.216)	(0.238)
Trips 1+ miles = [124,174)	-0.551**	-0.581**	-0.623**
	(0.231)	(0.235)	(0.265)
Trips 1+ miles = [174,238)	-0.429	-0.411	-0.563*
	(0.265)	(0.272)	(0.291)
Trips 1+ miles = [238,364)	-0.394*	-0.410*	-0.518**
	(0.229)	(0.234)	(0.258)
Trips 1+ miles = 365	0	0	0
	(.)	(.)	(.)
Has participated in CBC	0.0210	0.0183	0.115
	(0.0848)	(0.0862)	(0.0939)
Hunts birds	0.0512	0.0583	-0.0406
	(0.155)	(0.156)	(0.172)
Gender: Female	-0.160*	-0.132	-0.148
	(0.0830)	(0.0855)	(0.0940)
Education: High school		0.404*	0.437*

Continued on next page

Table A22 – continued from previous page

		(0.235)	(0.255)
Education: Some college		0.00943	-0.114
		(0.134)	(0.151)
Education: Masters degree		0.174*	0.183*
		(0.0981)	(0.108)
Education: Doctoral degree		0.115	0.0368
		(0.135)	(0.146)
Empl. status: Part time			-0.0868
			(0.170)
Empl. status: Looking for work			-0.716
			(0.584)
Empl. status: Unemployed			0.135
			(0.188)
Empl. status: Retired			-0.305***
			(0.0980)
/			
cut1	-2.587***	-2.490***	-2.738***
	(0.192)	(0.207)	(0.236)
cut2	-1.415***	-1.309***	-1.498***
	(0.182)	(0.198)	(0.225)
cut3	-0.466***	-0.365*	-0.557**
	(0.176)	(0.193)	(0.220)
Observations	826	808	691
Max. log-likelihood	-868.79	-847.83	-708.91

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A23: eBird sample: Model 25-27 (of 30) to accommodate missing values

	Model 25	Model 26	Model 27
Engagement-level indicator			
Travel 1+ mile data available	0	0	0
	(.)	(.)	(.)
Trips 1+ miles = 0	-2.750***	-2.820***	-2.838***
	(0.192)	(0.209)	(0.212)
Trips 1+ miles = [1,4)	-2.969***	-3.053***	-3.123***
	(0.221)	(0.241)	(0.245)
Trips 1+ miles = [4,7)	-2.031***	-1.928***	-1.997***
	(0.227)	(0.255)	(0.260)
Trips 1+ miles = [7,10)	-1.786***	-1.850***	-2.041***
	(0.294)	(0.310)	(0.320)
Trips 1+ miles = [10,21)	-1.559***	-1.658***	-1.678***
	(0.207)	(0.224)	(0.228)
Trips 1+ miles = [21,41)	-1.378***	-1.419***	-1.459***
	(0.221)	(0.238)	(0.241)
Trips 1+ miles = [41,72)	-1.157***	-1.210***	-1.235***
	(0.210)	(0.232)	(0.234)
Trips 1+ miles = [72,124)	-0.693***	-0.652***	-0.656***
	(0.216)	(0.236)	(0.239)
Trips 1+ miles = [124,174)	-0.612***	-0.617**	-0.669**
	(0.236)	(0.261)	(0.265)
Trips 1+ miles = [174,238)	-0.413	-0.573**	-0.578**
	(0.272)	(0.284)	(0.291)
Trips 1+ miles = [238,364)	-0.456*	-0.536**	-0.567**
	(0.235)	(0.255)	(0.260)
Trips 1+ miles = 365	0	0	0
	(.)	(.)	(.)
Has participated in CBC	0.0882	0.150	0.142
	(0.0880)	(0.0939)	(0.0952)
Hunts birds	-0.00231	-0.0533	-0.0639
	(0.158)	(0.173)	(0.174)
Gender: Female	-0.136	-0.154*	-0.138
	(0.0862)	(0.0927)	(0.0951)
Age: 24 years or less	0.649**	0.737**	0.628*

Continued on next page

Table A23 – continued from previous page

	(0.288)	(0.346)	(0.370)
Age: 25 to 34 years	0.194	0.216	0.271
	(0.178)	(0.194)	(0.197)
Age: 35 to 44 years	0.279*	0.288	0.304*
	(0.165)	(0.180)	(0.181)
Age: 55 to 64 years	-0.268**	-0.110	-0.131
	(0.127)	(0.145)	(0.148)
Age: 65 years and up	-0.223*	0.109	0.0808
	(0.124)	(0.179)	(0.182)
Education: High school	0.323		0.374
	(0.250)		(0.275)
Education: Some college	0.0255		-0.0841
	(0.137)		(0.152)
Education: Masters degree	0.269***		0.224**
	(0.101)		(0.110)
Education: Doctoral degree	0.203		0.0900
	(0.137)		(0.148)
Empl. status: Part time		-0.189	-0.156
		(0.175)	(0.177)
Empl. status: Looking for work		-0.978*	-0.908
		(0.587)	(0.590)
Empl. status: Unemployed		0.103	0.121
		(0.188)	(0.191)
Empl. status: Retired		-0.308**	-0.276*
		(0.144)	(0.147)
/			
cut1	-2.567***	-2.768***	-2.691***
	(0.227)	(0.238)	(0.254)
cut2	-1.345***	-1.510***	-1.425***
	(0.218)	(0.227)	(0.244)
cut3	-0.388*	-0.558**	-0.476**
	(0.213)	(0.222)	(0.239)
Observations	807	705	690
Max. log-likelihood	-833.29	-720.01	-702.08

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A24: eBird sample: Model 28-30 (of 30) to accommodate missing values

	Model 28	Model 29	Model 30
Engagement-level indicator			
Travel 1+ mile data available	0	0	0
	(.)	(.)	(.)
Trips 1+ miles = 0	-2.757***	-2.791***	-2.767***
	(0.204)	(0.208)	(0.224)
Trips 1+ miles = [1,4)	-2.928***	-3.047***	-3.002***
	(0.237)	(0.244)	(0.262)
Trips 1+ miles = [4,7)	-2.043***	-2.105***	-1.773***
	(0.246)	(0.251)	(0.281)
Trips 1+ miles = [7,10)	-1.792***	-2.052***	-1.943***
	(0.335)	(0.349)	(0.356)
Trips 1+ miles = [10,21)	-1.751***	-1.774***	-1.777***
	(0.226)	(0.229)	(0.247)
Trips 1+ miles = [21,41)	-1.450***	-1.475***	-1.450***
	(0.235)	(0.240)	(0.257)
Trips 1+ miles = [41,72)	-1.312***	-1.335***	-1.317***
	(0.224)	(0.227)	(0.251)
Trips 1+ miles = [72,124)	-0.841***	-0.830***	-0.714***
	(0.230)	(0.233)	(0.255)
Trips 1+ miles = [124,174)	-0.670***	-0.685***	-0.609**
	(0.253)	(0.258)	(0.287)
Trips 1+ miles = [174,238)	-0.466	-0.467	-0.540*
	(0.294)	(0.298)	(0.316)
Trips 1+ miles = [238,364)	-0.396	-0.392	-0.418
	(0.255)	(0.260)	(0.279)
Trips 1+ miles = 365	0	0	0
	(.)	(.)	(.)
Has participated in CBC	0.0778	0.0956	0.160
	(0.0954)	(0.0975)	(0.104)
Hunts birds	0.00672	-0.00191	-0.0502
	(0.166)	(0.167)	(0.179)
Gender: Female	-0.110	-0.0712	-0.125
	(0.0936)	(0.0965)	(0.104)
Age: 24 years or less	0.909***	0.893***	1.144***

Continued on next page

	(0.324)	(0.334)	(0.413)
Age: 25 to 34 years	0.119	0.207	0.218
	(0.188)	(0.191)	(0.212)
Age: 35 to 44 years	0.249	0.307*	0.284
	(0.172)	(0.174)	(0.190)
Age: 55 to 64 years	-0.267**	-0.298**	-0.0908
	(0.134)	(0.137)	(0.156)
Age: 65 years and up	-0.170	-0.186	0.188
	(0.135)	(0.137)	(0.197)
Income: Less than 25K	-0.0147	0.0242	-0.128
	(0.200)	(0.209)	(0.240)
Income: 25 K to 50 K	0.113	0.146	0.135
	(0.134)	(0.138)	(0.145)
Income: 75 K to 100 K	0.0129	0.0453	-0.0415
	(0.143)	(0.145)	(0.156)
Income: 100 K or more	0.191	0.225*	0.116
	(0.121)	(0.124)	(0.134)
Education: High school		0.590**	
		(0.281)	
Education: Some college		0.0934	
		(0.151)	
Education: Masters degree		0.303***	
		(0.111)	
Education: Doctoral degree		0.146	
		(0.151)	
Empl. status: Part time			-0.162
			(0.187)
Empl. status: Looking for work			-0.861
			(0.646)
Empl. status: Unemployed			-0.00203
			(0.213)
Empl. status: Retired			-0.369**
			(0.160)
/			
cut1	-2.706***	-2.525***	-2.749***
	(0.240)	(0.255)	(0.272)

Continued on next page

Table A24 – continued from previous page

cut2	-1.457***	-1.252***	-1.424***
	(0.230)	(0.246)	(0.262)
cut3	-0.488**	-0.281	-0.483*
	(0.225)	(0.241)	(0.257)
Observations	687	673	584
Max. log-likelihood	-718.20	-696.80	-599.55

t in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

H Calculating heterogeneous population weights for eBird member survey sample

Selection-correction models allow the analyst to accommodate the possibility that there is a correlation between the *unobserved error term* in the process that leads to an individual's presence in the estimating sample and the *unobserved error term* in the process that generates the outcome variable for that individual. However, it is possible that the *observable heterogeneity* across the individuals who show up in the estimating sample could also be different from the observable heterogeneity in the general population.

In such a case, researchers often consider the use of exogenous weights. Weights are used to scale the relative frequency of people of different types in the estimating sample so that group proportions more closely match the corresponding group proportions in the population. With a fully representative estimating sample, each observation represents an equal number of people in the population as a whole, so that average preferences in the sample (for example) should be the same as average preferences in the population. If each observation in the sample represents a very different number of people in the population, then estimated average preferences are less likely to scale up to the general population.

We seek an appropriate set of weights to use when estimating our outcome equation, to explain consideration-set radius (i.e. the maximum one-way distance a respondent is willing to travel on a day-trip to see wild birds). Many weighting schemes employ the relative frequencies of people of different types in the population divided by the relative frequencies of people of those same types in the estimating sample. The ratios of relative frequencies are then scaled so that they sum to the size of the estimating sample. Typically, the researchers bins both the population and the estimating sample according to the values of some set of exogenous variables.

For rudimentary weights, we could use the observed undifferentiated proportions of respondents at each engagement level in the two samples, (qBus, eBird) = (0.273, 0.398), (0.252, 0.275), (0.265, 0.179), and (0.210, 0.146). However, we are also concerned that engagement intensity is not fully exogenous to the maximum distance variable we seek to model. Observed proportions do not allow for the possibility of systematically different mixes of people in the two samples. Thus we adapt the conventional exogenous weighting approach to express the *fitted probabilities* of an individual from each sample exhibiting the engagement intensity that they report. We compute our weights based on the within-sample fitted probabilities that each respondent participated in eBird at each of the four possible engagement levels, where these fitted probabilities are expressed as functions of the individual's exogenous characteristics, and any error term in the fitted probabilities is implicitly discarded, making each fitted probability a function of exogenous variables only.

For our eBird data, we are concerned that (a) the relative *proportions* of respondents in our eBird sample who engage with the project at different levels might differ from (b) the corresponding proportions in the population of eBird members who turn up in a random sample from the general population (i.e. our qBus sample). We again transfer our qBus parameter estimates for the six-level ordered-probit models with the sociodemographic characteristics of each person in our eBird member survey sample to calculate the *predicted* individual conditional engagement-level *probabilities* for respondents in our eBird member survey sample, as are shown in panel C of Figure 1 in the body of the paper. But then we also use our eBird member survey sample, *independently*, to estimate four-level ordered-probit models for engagement levels 3, 4, 5 and 6, and calculate predicted probabilities for these four engagement levels based on those parameters, where the distribution of these probabilities is shown in panel B of Figure 1 in the paper. We treat these two sets of predicted probabilities as the "expected" probabilities and the "observed" probabilities in the eBird member survey sample.

We construct our weights for each observation in the eBird sample by considering the observed engagement level for that person. We then generate a weight that reflects (a) the out-of-sample *predicted* probability that a person with these same characteristics would be at that level of engagement in the general population (qBus) sample, in ratio to (b) the within-sample *fitted* probability, estimated using the eBird member survey data, that they are at their observed level of engagement. As usual, we scale these weights so that they sum to the sample size for the eBird member survey. Figure A1 shows the smoothed density for the resulting distribution of heterogeneous weights for use in estimation of the outcome model that uses only the eBird member survey data (with a dotted line highlighting unit weights). For comparison, Figure A1 also shows what would be the four unique values of the set of homogeneous weights that would be calculated if we based the weight calculations only on the marginal distributions of engagement intensities, without reference to the heterogenity in respondent characteristics across the qBus general population sample and the eBird member survey sample.

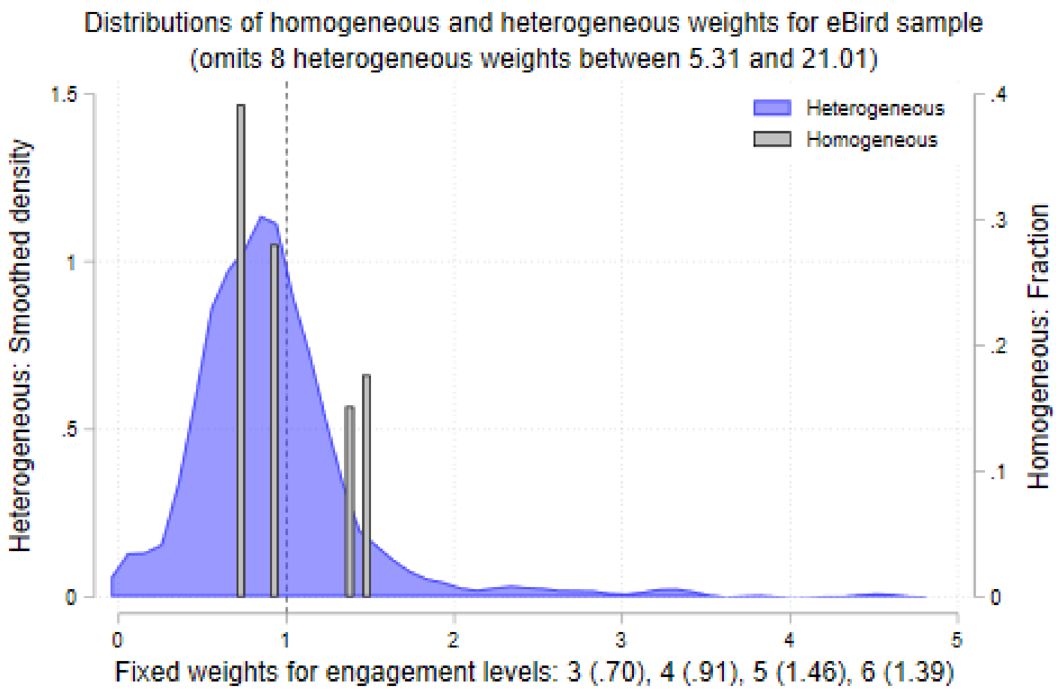

Figure A1: Distribution of weights across eBird member survey observations, where these weights serve to match engagement-intensity probabilities in the eBird member survey sample to engagement-intensity probabilities in the general-population qBus sample (six outlier weights, between 2.66 and 7.93, are not shown)

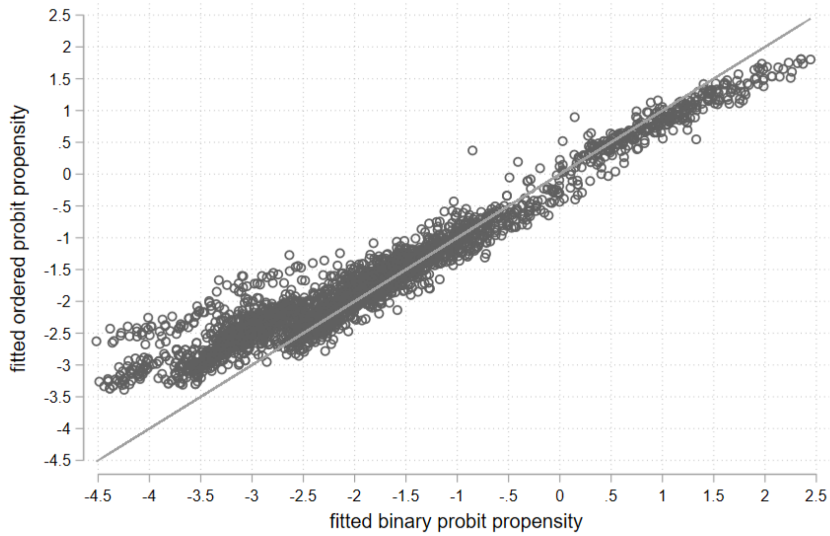

Figure A2: For 4,161 qBus general population respondents only: Fitted engagement propensities calculated from the higher-resolution ordered-probit model plotted against those from the simpler binary-probit model; equal values fall along the line

I Visualization of estimates of intermediate components

For the qBus dataset alone, which includes both members and non-members of eBird, Figure A2 shows the joint distribution of the (adjusted) fitted propensity index, $Z_i\hat{\gamma}^q$, for our new ordered-probit selection equation as well as the fitted propensity index, $Z_i\hat{\gamma}^q$ from a conventional binary-probit selection model using the same qBus data. The propensity index from our new ordered-probit selection model is somewhat higher than the index for the conventional binary model among people with low propensities to belong to eBird, but the upper part of the joint distribution coincides fairly closely. Our ordered-probit selection model recruits more information, with its multiple categories, from both non-members and members of eBird in the qBus sample, which likely accounts for the differences.

We can also consider the differences in the distributions of our two alternative IMR terms, calculated using parameter estimates from the qBus sample, *applied to individuals in our eBird member survey sample.* These two IMR variables are based on our two different selection models: (1) the binary-probit model and (2) the re-normalized (adjusted) ordered-probit model. Figure A3 shows the joint distribution of these two IMR terms.

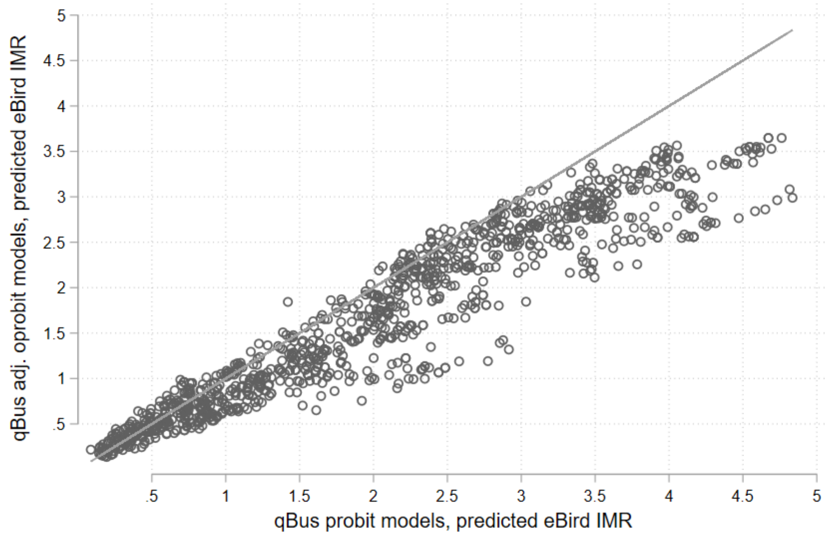

Figure A3: For 1,081 eBird member survey respondents: Predicted ordered-probit-based IMRs calculated with parameters estimated using qBus sample, plotted against predicted binary-probit-based IMRs calculated with parameters estimated using the qBus sample. For each individual observation in the eBird sample, we use the most-detailed qBus specification consistent with missing Z_j data for that observation.

Figure A4 illustrates the high degree of collinearity between the predicted inverse Mills ratio correction term and the predicted engagement propensity for the eBird sample. Had we devoted space in Table 4 to a model with only an interaction term between the demeaned engagement propensity and the intercept term in the model, the coefficient on the single additional would have adapted to the change of scale and sign in the propensity, as opposed to the inverse Mills ratio, and essentially the same values for the remaining parameters would result.

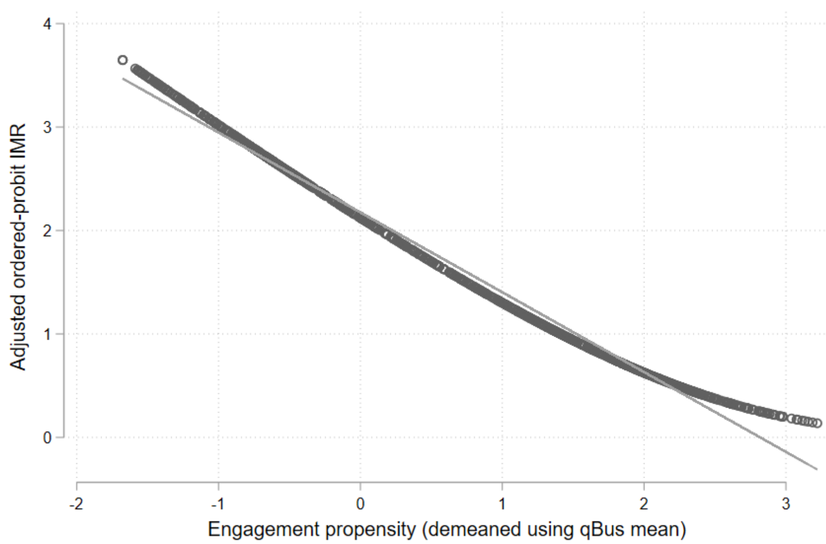

Figure A4: The relationship between our adjusted ordered-probit inverse Mills ratio term used in Model 4 in Table 4 and the underlying demeaned engagement propensity variable used to shift the basic coefficients in Model 5.

J Issues still to be addressed in the transfer of selection propensities

J.1 When the outcome equation is a conditional logit choice model, rather than a regression model

Our broader research project with the eBird sample involves destination choice models and inferred preferences for site attributes, employed to estimate non-market benefits associated with wild birds. However, Heckman selection-correction models are not appropriate when the outcome variable of interest is a discrete choice, because the latent choice propensity in a multiple discrete-choice model is *not* conditionally normally distributed.[40] When there is no bivariate normal error term to justify the use of a fitted IMR from a selection model as an additional regressor in an outcome equation, it is nevertheless still possible to explore the more ad hoc correction that accommodates systematic differences in selection propensities across respondents by allowing second-stage parameters to differ systematically with deviations of fitted respondent selection propensities from the average propensity in the general population. This ad hoc correction is used in Cameron and DeShazo (2013), Johnston and Abdulrahman (2017), Kolstoe and Cameron (2017) and Kolstoe et al. (2018). As mentioned in Appendix B, Terza (2009) offers models that hint at the possibility of selection-correction methods for conditional logit models, but his method would need to be adapted extensively to suit the case where the selection model is to be transferred to a different sample.

J.2 Estimated regressors and inference in a second-stage model

Of course, any two-step estimation process that does not account for the estimated property of the $\hat{\gamma}^q$ parameters embodied in the calculated IMR terms—as in Models 3 and 4 (or the fitted de-meaned response propensities, as in Model 5)—can risk some bias in the inferences to be drawn in the second step. The IMR term (or fitted de-meaned predicted response propensity) is an "estimated regressor" that likely overstates the amount of information in the data. It may be straightforward (if tedious), to implement an appropriate FIML estimator in the case where one does not need to contend with any missing data in the dataset to which the selection specification is to be transferred. Recall that in this case, it was necessary to estimate 30 different selection equation specifications using many combinations of none, some, or all of the categories of indicators in the full selection model.

If all variables for the selection equation were available for every observation in the CS sample having the outcome variable of interest, one could define the log-likelihood function over the full set of parameters: $\gamma, \beta, \sigma_\eta, \sigma_\epsilon, \rho$. The structure of the two-step model could be preserved, but the two equations could be estimated simultaneously, constraining the γ parameters to be the same in both the selection equation (using the qBus sample) and the

[40]Some researchers (e.g. Yuan et al. (2015)) have inserted a fitted inverse Mills ratio (IMR) into a second-stage discrete-choice model, although there seems to be no statistical justification for this particular transformation of the fitted selection propensity.

outcome equation (using the IMR term), where the index for the IMR variable is constructed using the γ parameters combined with the Z_j variables for the eBird dataset. The matter of how to construct the weights would need to be resolved, of course. We do not attempt this FIML estimation here, because of the significant amount of missing data for the selection equation applied to the eBird sample, and corresponding proliferation of different specifications necessary to provide predicted propensities, IMRs, and weights that maximize our use of the available data for each eBird respondent. While joint estimation of 30 different ordered-probit equations plus the outcome equations would likely be possible, we would not expect it to make any qualitative difference in our findings.

J.3 Other possible layers of selection

Our estimating sample for the illustrative market-extent model in this paper consists of respondents to our eBird member survey who provided complete data for all except the (typically sensitive) detailed income variable. The selection-correction strategies we feature in this paper presume that this group of eBird member survey respondents is representative of eBird members, an assumption we make to permit us to focus on the problem of systematic selection into eBird engagement at different levels of intensity.

Of course, there may be a variety of reasons why invited eBird members decide not to participate in our survey. In other research, we have sought to control for heterogeneous selection across all eBird members (rather than the general population) by linking the center of gravity of their birding trips to a specific census tract, which we impute to be their home census tract. We have employed census tract attributes as proxies for possible systematic variation across birder characteristics because there is no sociodemographic information available for eBird members who did not respond to our survey. We then employ these various census tract attributes to construct a "propensity of an invited eBird member to respond to our survey" and use this propensity in an attempt to control crudely for selection into our estimating sample for destination choice models. We do not attempt to overlay that correction procedure in addition to the strategies employed here. It is possible that the selection of eBird members into our survey, based on the attributes of their home zipcode relative to those of the general population, is dominated by the selection of the general population into eBird participation, but this is an empirical question that is beyond the scope of this study.

Even with the illustrative example concerning the the radii of consideration sets for birding excursions, it is possible that sample selection may occur along more than one dimension. Respondents to our eBird survey may have unobserved characteristics that make them simultaneously more likely (than otherwise expected) to be members of eBird and also more-likely than average to participate in travel of more than one mile from home to observe birds. In other research, it is this second behavior upon which we base our models of the "active" recreational use of opportunities to watch wild birds. In addition to the eBird engagement-level variable captured by our ordinal variable $CS6_i$, we can also distinguish between birders who do, or do not, travel more than one mile to see birds, in both the qBus sample and the eBird sample. Three categories of actual bird-watching might be distinguished: no birding trips, trips only less than 1 mile, and trips of one mile or more. Thus the ordinal eBird engagement levels might be supplemented by a second (presumably correlated) ordinal variable that captures heterogeneity in the actual bird-watching behavior of respondents in both surveys. Models with selection on two (correlated) latent variables would require working with trivariate normal joint distributions of the error terms. These models are also beyond the scope of the current paper, again because of the 30 different ordered-probit selection equations necessary to accommodate transfer of our selection model from the qBus sample to the eBird member survey sample with its missing values for different variables.

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. DeShazo (2013). Demand for health risk reductions. *Journal of Environmental Economics and Management 65*(1), 87–109.

Cameron, T. A. and J. Englin (1997). Respondent experience and contingent valuation of environmental goods. *Journal of Environmental Economics and Management 33*(3), 296–313.

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), 209–242.

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), 215–242.

Greene, W. H. (2012). *Econometric Analysis (7 th)*. Upper Saddle River, NJ: Prentice Hall.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*, 153–161.

Holland, B. M. and R. J. Johnston (2017). Optimized quantity-within-distance models of spatial welfare heterogeneity. *Journal of Environmental Economics and Management 85*, 110–129.

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), 246–267.

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), 302–322.

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), 779–782.

Olsen, R. J. (1980). A least-squares correction for selectivity bias. *Econometrica 48*(7), 1815–1820.

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*.

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), 233–249.

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), 243–270.

Terza, J. V. (1998). Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects. *Journal of Econometrics 84*(1), 129–154.

Terza, J. V. (2009). Parametric nonlinear regression with endogenous switching. *Econometric Reviews 28*(6), 555–580.

Vella, F. (1998). Estimating models with sample selection bias: a survey. *Journal of Human Resources*, 127–169.

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*, 156–164.

Walsh, P. J., J. W. Milon, and D. O. Scrogin (2011). The spatial extent of water quality benefits in urban housing markets. *Land Economics 87*(4), 628–644.

Walters, K., D. A. Christakis, and D. R. Wright (2018). Are mechanical turk worker samples representative of health status and health behaviors in the u.s.? *PLOS ONE 13*(6), e0198835.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data.* MIT Press.-2002.

Yuan, Y., K. J. Boyle, and W. You (2015). Sample selection, individual heterogeneity, and regional heterogeneity in valuing farmland conservation easements. *Land Economics 91*(4), 627–649.