Rounding in Recreation Demand Models: A Latent Class Count Model
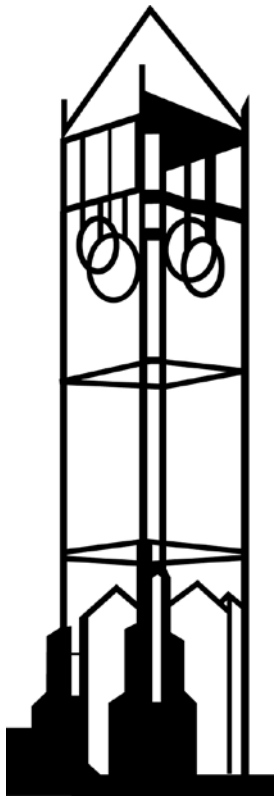
Keith Evans, Joseph A. Herriges

IOWA STATE UNIVERSITY
Department of Economics
Ames, Iowa, 50011-1070

Rounding in Recreation Demand Models:
A Latent Class Count Model

Keith Evans[1]   Joseph A. Herriges
Iowa State University   Iowa State University

June 2, 2010

Abstract

A commonly observed feature of visitation data, elicited via a survey instrument, is a greater propensity for individuals to report trip numbers that are multiples of 5's, relative to other possible integers (such as 3 or 6). One explanation of this phenomenon is that some survey respondents have difficulty recalling the exact number of trips taken and instead choose to round their responses. This paper examines the impact that rounding can have on the estimated demand for recreation and the bias that it may induce on subsequent welfare estimates. We propose the use of a latent class structure in which respondents are assumed to be members of either a *nonrounding* or a *rounding* class. A series of generated data experiments are provided to illustrate the range of possible impacts that ignoring rounding can have on the estimated parameters of the model and on the welfare implications from site closure. The results suggest that biases can be substantial, particularly when then unconditional mean number of trips is in the range from two to four. An illustrative application is provided using visitation data to Saylorville Lake in central Iowa.

---

[1]Contact author information: 260 Heady Hall, Department of Economics, Iowa State University, Ames, IA 50011. email: *ksevans@iastate.edu*. Phone: 515-294-6173

1 Introduction

Models of recreation demand are used extensively to value both access to and potential changes in environmental amenities at recreation facilities, such as lakes, rivers and beaches. Analysts link visitation patterns to the cost of traveling to a site, consumer characteristics and the attributes of the available sites using a range of modeling frameworks, including discrete choice Random Utility Maximization (RUM) models, count data models, and the structural Kuhn-Tucker model. Key to all of these approaches, of course, are data on the numbers of trips to the sites of interest. Trip data most often take the form of counts of the trips taken over a fixed time horizon (e.g., a summer season or calendar year) elicited via a survey instrument, asking the individual to recall (or in some applications to forecast) their numbers of trips. A commonly observed feature of these counts is a greater propensity for individuals to report trip numbers that are multiples of 5's, relative to other possible integers (such as 3 or 6). One explanation of this phenomenon is that some survey respondents have difficulty recalling the exact number of trips taken and instead choose to round their responses.[2] While the apparent clumping of trip data around specific integers is a familiar pattern in recreation demand data, we are aware of no efforts in the literature to date that attempt to account for this pattern. Instead, practitioners treat the reported counts as an accurate reflection of the trips taken by the survey respondent. Even in the broader survey literature, attempts to account for rounding in survey data analyses are rare. Manksi and Molinari [8] provide one of the few exceptions, developing an approach to partially identify patterns in probabilistic expectations elicited via survey instruments.

The purpose of this paper is to examine the impact that rounding can have on the estimated demand for recreation and the bias that it may induce on subsequent welfare estimates. In particular, we propose a latent class count data model of visitations to a single site in which respondents are assumed to be members of either a *nonrounding* or a *rounding* class, with the latter group providing censored responses to trip questions by rounding their trip counts to the nearest multiple of five. We are agnostic as to why the latter group chooses to round. As Manski and Molinari [8] suggest, "...[t]here are no established conventions for rounding survey responses. Hence, researchers cannot be sure how much rounding there may be in survey data. Nor can researchers be sure whether respondents round to simplify communication or to convey partial knowledge" (p. 219). We go on to suggest the use of an expectation-maximization (EM) algorithm for the estimation of the model. A series of generated data experiments are then provided to illustrate the range of possible impacts that ignoring rounding can have on the estimated parameters of the model and on the welfare implications from site closure. The results suggest that biases can be substantial,

---

[2]Similar phenomena have been observed in other survey settings. For example, Dominitz and Manski [4] note that in surveys eliciting probabilistic expectations (e.g., the probability of loosing one's job or living to a specific age), responses tend to bunch around multiples of 5%. See Manksi and Molinari [8] for additional discussion of this phenomenon.

particularly when then unconditional mean number of trips is in the range of two to four. Finally, an illustrative application is provided using data on the visitations to Saylorville Lake, a popular recreational site and reservoir in central Iowa. The paper closes with an overall summary of our findings and a discussion of possible extensions of the modeling framework.

2 The Model

We begin this section by formally defining the assumed latent class structure and developing the necessary notation. Latent class models have emerged in recent years as a popular approach to incorporating preference heterogeneity in discrete choice models, both in recreation demand (e.g., [2],[10],[11]) and in the broader literature (e.g., [5],[6],[7]). In our application, the heterogeneity lies in the individual's propensity to round. We then propose an EM algorithm for use in the estimation of the model.

2.1 The Latent Class Count Data Model

The starting point in our approach to examining the impact of rounding in the modeling of recreation demand is to assume that individuals fall into one of two *latent classes*: *Nonrounders* (N) or *Rounders* (R). Individual class membership (denoted by $C_i^* = N$ or R) is unknown to the analyst. For each class, the actual number of trips (y_i^*) taken to the site in question is assumed to be drawn from a Poisson distribution, though the underlying parameters of the Poisson distribution are allowed to vary by class. Specifically, we assume that:

$$Pr(y_i^* = k | C_i^* = c) = \frac{\exp(-\lambda_{ic})\lambda_{ic}^k}{k!} \quad i = 1, \dots, I; c = N, R, \tag{1}$$

where

$$\lambda_{ic} = \exp(\boldsymbol{X}_i' \boldsymbol{\beta}_c) \tag{2}$$

denotes the conditional mean trips for individuals in class c given characteristics $\boldsymbol{X}_i$ and the parameter vector $\boldsymbol{\beta}_c$. For individuals in the nonrounding class, the reported trips y_i are assumed to be the same as the actual number of trips (i.e., $y_i = y_i^*$). Thus, *conditional* on knowing that $C_i^* = N$, the individual's choice probability is simply:

$$L_{iN}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_N) = \frac{\exp(-\lambda_{iN})\lambda_{iN}^{y_i}}{y_i!} \tag{3}$$

In contrast, for individuals in the rounding class, reported trips are assumed to be rounded to the nearest multiple of five for trips greater than 2.[3] Let $\mathcal{I}_5$ denote the set of positive integers that

---

[3]We assume that when actual trips (y_i^*) equal 1 or 2, they are not rounded down to zero by the survey respondent, even in the case of the rounding class. It seems reasonable to us that, in reporting trips, the rounding individual distinguishes taking a trip from staying at home even when the number of trips is small.

are multiples of five. In this case, *conditional* on knowing that $C_i^* = R$, the individual's choice probability is simply:

$$L_{iR}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_R) = \begin{cases} \frac{\exp(-\lambda_{iR})\lambda_{iR}^{y_i}}{y_i!} & y_i = 0, 1, 2 \\ \sum_{j=-2}^{2} \frac{\exp(-\lambda_{iR})\lambda_{iR}^{y_i+j}}{(y_i+j)!} & y_i \in \mathcal{I}_5 \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Let $s_R \in [0, 1]$ denote the probability of being in the nonrounding class. Since class membership is not known, the unconditional choice probability for individual i becomes:

$$L_i(y_i, \boldsymbol{X}_i; \boldsymbol{\theta}) = s_R L_{iR}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_R) + (1 - s_R) L_{iN}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_N), \quad (5)$$

where $\theta = (\boldsymbol{\beta}_N', \boldsymbol{\beta}_R', s_R)'$ denotes the combined parameters of the model. The parameter vector θ can be obtained using standard maximum likelihood gradient based methods. However, latent class models are notorious for their difficulty in estimation, particularly since the class labels are themselves arbitrary. In our generated data experiments and application, we instead employ the EM algorithm described below.

2.2   The EM Algorithm

EM algorithms were introduced by Dempster, Laird and Rubin [3] as a means dealing with missing data and have subsequently been adapted to a variety of estimation problems in which some piece of information in a model is missing.[4] In the current application, the missing piece of information we focus on is the class membership variable C_i^*. As with all EM algorithms, the procedure is iterative. Let θ^t denote value of the parameter vector at iteration t. Following the notation in [13], the next iteration on θ (i.e., θ^{t+1}) is obtained for our latent class model by maximizing:

$$\mathcal{E}(\theta|\theta^t) = \sum_{i=1}^{I} \sum_{c=N}^{R} h_{ic}^t \ln \left[s_c L_{ic}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_c) \right] \quad (6)$$

where

$$h_{ic}^t = h(C_i^* = c|y_i, s^t) = \frac{s_c^t L_{ic}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_c^t)}{s_N^t L_{iN}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_N^t) + s_R^t L_{iR}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_R^t)} \quad (7)$$

denotes the probability that individual i belongs to class c conditional on the observed choice of the individual. Given the structure of the problem in (6), this is equivalent to separately maximizing:

$$\mathcal{E}(s|\theta^t) = \sum_{i=1}^{I} \left[h_{iR}^t \ln(s_R) + h_{iN}^t \ln(1 - s_R) \right] \quad (8)$$

---

[4]Train [13], chapter 14, provides an excellent overview of EM algorithms, while McLachlan and Krishnan [9] provide a review of applications.

with respect to s_R and (for both $c = R$ and N) maximizing

$$\mathcal{E}(\beta_c | \beta^t) = \sum_{i=1}^{I} h_{ic}^t \ln \left[L_{ic}(y_i, \boldsymbol{X}_i; \boldsymbol{\beta}_c) \right] \tag{9}$$

with respect to β_c. The solution to the maximization of (8) yields:

$$s_R^{t+1} = \frac{\sum_{i=1}^{I} h_{iR}^t}{\sum_{i=1}^{I} \left(h_{iN}^t + h_{iR}^t \right)}. \tag{10}$$

The specific steps involved in the EM algorithm are then:

1. With t denoting the current iteration, set $t = 0$ and specify the initial values for both the share of rounders (i.e., s_R^0) and the parameters of the two classes (i.e., β_N^0 and β_R^0). We set $s_R^0 = 0.5$. Similar to the approach suggested by [13] in the context of a latent class logit model (p. 360), the starting values for the parameters of our two latent classes are obtained by randomly partitioning the sample into two groups (N and R) and maximizing (9) for each subsample to obtain β_N^0 and β_R^0.

2. For each observation i and each class c, the probability h_{ic}^t that individual i belongs to class c conditional on the observed choice of the individual is computed using (7).

3. The updated class share of rounders (s_R^{t+1}) is obtained using (10).

4. The updated parameters for the two latent classes (β_N^{t+1} and β_R^{t+1}) are obtained by maximizing (9).

5. Check for convergence. If convergence has not been achieved, then t is incremented by 1 and the algorithm returns to step 2. Otherwise, the algorithm ends and the standard errors for the parameters can be calculated. We use bootstrapped standard errors, but an alternative approach would be to use the converged values from the EM algorithm as starting values in a maximum likelihood estimation of θ using (5).

3 Generated Data Experiments

In order to investigate to potential impact that rounding can have on both parameter estimates and subsequent welfare calculations, we conduct a series of generated data experiments. In all of the experiments, the conditional mean number of trips for each class is assumed to be a linear exponential function of travel cost to the site (denoted by P_i), individual income (denoted by Y_i), and a demographic variable (Z_i) . Specifically, we assume that:

$$\lambda_{ic} = \exp(\beta_{c0} + \beta_{cP} P_i + \beta_{cY} Y_i + \beta_{cZ} Z_i), \quad c = N, R. \tag{11}$$

Across the experiments we vary two factors: (1) the share of rounders (i.e., s_R) and (2) the mean trips for the two classes (by varying β_{N0} and β_{R0}).[5] The emphasis on s_R is obvious, with our model reducing to the standard count data model when $s_R = 0$. Five values for s_R were considered ($s_R = 0.1$, 0.25, 0.5, 0.75, and 0.9). We focus on mean trips by class, since the propensity for individuals to round will depend, in part, on their trip frequencies. If the vast majority of the sample takes 0, 1, or 2 trips, then there will be little room for rounding to occur. Along these lines, we consider two basic experiments. In experiment #1, we fix $\phi_0 \equiv \beta_{R0}/\beta_{N0} = 2$ (making rounders correspond to somewhat more avid trip takers) and vary β_{N0} from 0.5 to 1.5 in steps of 0.25 (increasing the overall trip taking by the two groups). The resulting intercepts are listed part a of Table 1.[6] In experiment #2, we fix $\bar{\beta}_0 = \frac{1}{2}(\beta_{N0} + \beta_{R0})$ (i.e., the simple average of the two type intercepts), varying ϕ_0 from 0.5 to 2.0. When $\phi_0 = 0.5$, rounders are less frequent trip takers than non-rounders, whereas when $\phi_0 = 2$ the opposite is true. The resulting intercepts are listed in part a of Table 2. In both experiments, we assume that the price, income and demographic coefficients are the same across the two classes (i.e., $\beta_{cP} = \beta_P$, $\beta_{cY} = \beta_Y$, and $\beta_{cZ} = \beta_Z$ for $c = N, R$). In total, twenty-five generated data settings were analyzed for each experiment. In all of the experiments, the sample size was set at I=5000.

Formally, 100 generated data sets (with 5000 observations in each data set) were constructed for each experiment/setting as follows:

1. Vectors of travel cost $(P_1, \ldots, P_I)$, income $(Y_1, \ldots, Y_I)$ and the demographic variable $(Z_1, \ldots, Z_I)$ were drawn from uniform distributions (i.e., $P_i \sim U[0,1]$, $Y_i \sim U[0,1]$, and $Z_i \sim U[0,1]$).

2. Using β_{R0} and β_{N0} for the given setting, along with $\beta_P = -0.75$, $\beta_Y = 0.25$, and $\beta_Z = -0.25$, λ_{ic} was computed for each individual and latent class using (11).

3. Each individual in the sample was randomly assigned to either the nonrounding ($C_i^* = N$) or rounding ($C_i^* = R$) latent class with probabilities s_N and s_R, respectively, using a draw from uniform distribution; i.e.,

$$C_i^* = \begin{cases} N & u_i < s_N \\ R & \text{otherwise,} \end{cases} \tag{12}$$

where $u_i \sim U[0,1]$.

4. Using

$$\lambda_i = 1(C_i^* = N)\lambda_{iN} + 1(C_i^* = R)\lambda_{iR}, \tag{13}$$

---

[5]We also investigated the impact of varying both the overall price coefficient and differences between β_{NP} and β_{RP}, but found that this had relatively little impact on the bias induced by rounding.

[6]Variations in these intercepts will induce variations in a group's unconditional mean number of trips. Table 1 also provides (in square brackets) the corresponding unconditional mean trips for each group and parameter setting given the assumed data generating process. For example, with $\beta_{0R} = 1$, the corresponding unconditional mean trips would be 1.40.

where $1(\cdot)$ is the indicator function, the individual's actual trips (y_i^*) were drawn from a Poisson distribution with conditional mean λ_i.

5. Reported trips (y_i) were then constructed as

$$
y_i = \begin{cases} y_i^* & y_i = 0, 1, 2 \\ 1(C_i^* = N)y_i^* + 1(C_i^* = R)\mathrm{rnd}_5(y_i^*) & \text{otherwise,} \end{cases} \tag{14}
$$

where $\mathrm{rnd}_5(x)$ is the function censoring x to the nearest integer of five.

For each experiment, two models were estimated: (1) The latent class count data model (LCCM) outlined above and (2) the standard (single class) (SCCM) count data model in which no rounding is assumed. The resulting parameter estimates are available from the authors upon request. However, as the ultimate goal of the recreation demand model is typically for use in policy analysis, we focus our attention here on the potential bias on subsequent welfare calculations that results from ignoring respondent rounding. Specifically, we consider the welfare impact of the complete loss of access to the site as measured by consumer surplus (CS).[7] In a Poisson count data model with the linear exponential representation of mean trips, the change in consumer surplus resulting from the elimination of the site is given by:

$$
CS_i = \frac{\lambda_i}{\beta_{iP}}, \tag{15}
$$

where

$$
\lambda_i = \exp(\beta_{i0} + \beta_{iP}P_i + \beta_{iY}Y_i + \beta_{iZ}Z_i) \tag{16}
$$

denotes the mean number of trips and $\boldsymbol{\beta}_i = (\beta_{i0}, \beta_{iP}, \beta_{iY}, \beta_{iZ})$ denotes individual i's true parameter vector. The true welfare loss measures for individual i in the generated data sample are computed using equation (15). Averaged across the individuals yields the mean true welfare loss for the sample (denoted $\overline{CS}^{Tr}$) for the rth generated data set.

For the single class count data model, the estimated welfare loss measures were computed for each individual i using the fitted parameter vector from the SCCM specification for the rth generated data set. Averaged across the individuals yields the mean welfare loss for the sample predicted using the SCCM specification (denoted $\overline{CS}^{Sr}$).

For the latent class count data model, the predicted welfare loss for individual i is a weighted average of the welfare loss predicted for each latent class; i.e.,

$$
\widehat{CS}_i^{Lr} = (1 - \hat{s}_R^r)\widehat{CS}_i^{Nr} + \hat{s}_R^r\widehat{CS}_i^{Rr} \tag{17}
$$

[7]Similar results are obtained if either compensating variation (CV) or equivalent variation (EV) are used instead to measure the welfare impact.

where

$$\widehat{CS}_i^{cr} = \frac{\exp(\hat{\beta}_{c0}^r + \hat{\beta}_{cP}^r P_i + \hat{\beta}_{cY}^r Y_i + \hat{\beta}_{cZ}^r Z_i)}{\hat{\beta}_{cP}^r}, \quad \text{for } c = N,\ R \tag{18}$$

and $\hat{\beta}_{ck}^r$ ($k = 0, P, Y, Z$) and $\hat{s}_R^r$ denote the fitted parameter estimates from the LCCM using the rth generated data set. Averaged across the individuals yields the mean welfare impact for the sample predicted using the LCCM specification (denoted $\overline{CS}^{Lr}$). For each experiment/setting, we compute the percentage error of each model in predicting the true consumer surplus. Tables 1b and 2b provide a summary of our findings for experiments 1 and 2, respectively.

Starting with experiment 1, several patterns emerge. First, as we would expect, the LCCM model does well in predicting the mean welfare loss stemming from the elimination of the site, since it is the correct specification of the data generating process. In general, the average error is less than one percent. The errors are typically larger when the share of rounders s_R is small, leaving relatively few observations with which to estimate parameters for the rounding class. Second, the bias in welfare predictions from ignoring rounding (and using the standard SCCM) can be substantial. Consumer surplus is overstated by as much as 37%. Indeed, the extent to which the SCCM consumer surplus measure overstates the overall welfare loss appears to increase with the latent percentage of rounders in the sample, but does not increase monotonically as the average number of trips increase. Indeed, the largest bias occurs when the unconditional mean number of trips for the rounding class is just over two. This may simply be because, when few trips are taken by the rounding class, there is little opportunity for rounding, whereas when the rounding class takes many trips (e.g., with an unconditional mean of 10.32), the percentage error in reported trips is smaller (e.g., rounding 7 trips to 5 is a larger percentage error than when rounding 47 trips to 45).

Turning to the second experiment in Table 2b, we again see that percentage error resulting from ignoring rounding increases with the size of the rounding class, with the bias being largest when the unconditional mean trips for the rounding class is in the range from 2 to 3. Even when the coefficients are identical for the two latent classes (i.e., $\phi_0 = 1$), the SCCM welfare measures are biased; consumer surplus is overstated by as much 36%. This is due to the fact that, within the rounding class, there will be a larger percentage of individuals rounding up than rounding down (e.g., a larger percentage of the population will have actual trips of 3 and 4 relative to those having actual trips of 6 and 7). Thus, reported trips will be a biased indicator of actual trips for the rounding class, with $E(y_i|\boldsymbol{X}_i, C_i = R) > E(y_i^*|\boldsymbol{X}_i, C_i = R)$.

4 Application

As illustration of our proposed methods, we employ data from the Iowa Lakes Valuation Project. The Iowa Lakes Project, funded by the Iowa Department of Natural Resources and the US EPA, was a four year effort to gather panel data on the recreational lake usage patterns of Iowa households. Beginning in 2002, trip counts for the 132 primary recreational lakes in the state were elicited from a random sample of 8000 state residents. After accounting for nondeliverables, the overall response rate to the mail survey was approximately 62%.[8] In the current paper, we limit our attention to visits to a single site, Saylorville Lake, a reservoir in central Iowa locate just north of the state capital, Des Moines. We also restrict our attention to households within a 100 mile radius of the site, leaving a total of I=1395 observations for use in our analysis. Table 3 provides basic summary statistics for the sample.[9] As Table 3 indicates, the mean number of trips taken to Saylorville Lake in 2002 is 1.66, with approximately 69.4% of the sample choosing not visiting the site that year.

Table 4 provides the parameter estimates for the SCCM and two versions of the LCCM specification. In version 1 of the LCCM, we constrain the parameters of the rounding and non-rounding groups to be the same (and in doing so focus on rounding alone as a source of bias), whereas version 2 relaxes restriction. In general, all of the parameter estimates are statistically significant. The SCCM model finds, as expected, that travel cost negatively impacts the mean number of visits to the site. The results also suggest that trips increase with income, but decrease with the individual's age and education. Similar results are found in the constrained LCCM model, though the impact of education is now somewhat larger. The estimated share of rounders is approximately one-third of the population. The mean consumer surplus associated with closure of Saylorville Lake is approximately 5.3% higher using the SCCM model ($22.49) compared to estimates based on the constrained LCCM specification ($21.36), which is line with our generated data experiment. With the unconditional mean number of trips of 1.66, there is relatively little room for rounding to impact the results.

Turning to the unconstrained LCCM specification, while the general sign of the marginal effects are similar to the other two specification, the parameters differ somewhat between the nonrounding and rounding latent class. Trips are more responsive to age and education for the rounding class, but less responsive to price and income. As was the case in the constrained LCCM model, under forty percent of the population is found to belong to the rounding class. Despite the similarities with the other two model, the unconstrained specification yields a substantially higher estimate of the consumer surplus loss due to the closure of the site ($43.41). While this is certainly possible,

---

[8]Additional details regarding the Iowa Lakes Project can be found in [1]

[9]Travel cost P_i is computed assuming an out-of-pocket trip cost of $0.25 per mile times the individual's round trip distance to the site and a time cost of one-third the individual's hourly times the round trip travel time to the site. Travel distance and travel time were computed using the software package *PCMiler*.

we believe that some caution would be appropriate in using the unconstrained LCCM specification. Examining the summary statistics in Table 3, it is clear that the data exhibits a form of overdispersion, since the unconditional mean number of trips (1.66) is much less than the corresponding unconditional variance (24.2). Intuitively, it seems possible that the unconstrained LCCM results may be using the rounding class to compensate for overdispersion. A generalization of the LCCM specification (e.g., using the negative binomial as the base distribution) could be used to examine this issue further.

5   Summary and Possible Extensions

The objective of this paper was to illustrate the potential bias that rounding can have on both the characterization of trip demand and the subsequent welfare estimates derived from a count data model of recreation demand. We propose a latent class model to allow for rounding by a subset of the population. Both our generated data experiments and an application to recreation demand at Saylorville Lake in central Iowa suggest that the potential bias can be substantial.

There are a number of possible directions for future research. First, our latent class model assumes that the share of rounders (s_R) is a constant. However, it seems reasonable that the propensity for individuals to round might depend upon their characteristic (e.g., age, gender, etc.), as well as the circumstance under which the survey is conducted (e.g., involving near-term recall versus recall for time periods further into past). Numerous authors have made the class membership probability in the latent class model of function of respondent attributes (e.g., using a logit specification). Second, our latent class model allows for only one type of rounding (i.e., to the near integer multiple of five). The framework could readily be generalized to allow for a variety of rounding behaviors (e.g., rounding to multiples of ten) by introducing addition latent classes. Finally, the Poisson count model underlying our latent class model specification carries with it the often criticized assumption of equidispersion (with the conditional mean of the trips being equal to the conditional variance). The latent class approach used above could, however, be readily generalized by assuming the each class has actual trips that are from a more general count data distribution allowing for overdispersion (e.g., the negative binomial).

References

[1] Azevedo, C., K. Egan, J. Herriges, and C. Kling (2003). *Iowa Lakes Valuation Project: Summary and Findings from Year One.* Final Report to the Iowa Department of Natural Resources, August.

[2] Boxall, A., and W. Adomowicz (2002). "Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach," *Environmental and Resource Economics*, Vol. 23, pp. 421-442.

[3] Dempster, A., N. Laird, and D. Rubin (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society B* Vol. 39, pp. 138.

[4] Dominitz, J., and C. Manski (1997). "Perceptions of Economic Insecurity: Evidence From the Survey of Economic Expectations," *Public Opinion Quarterly*, Vol. 61, pp. 261-287.

[5] d'Uva (2005). "Latent Class Models for Use of Primary Care: Evidence from a British Panel," *Health Economics*, Vol. 14, pp. 873-892.

[6] d'Uva (2006). "Latent Class Models for Utilisation of Health Care," *Health Economics*, Vol. 15, pp. 329-343.

[7] Greene, W., and D. Hensher (2003). "A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit," *Transportation Research, Part B*, Vol. 37, pp. 681-698.

[8] Manski, C., and F. Molinari (2010). "Rounding Probabilistic Expectations in Surveys," *Journal of Business & Economic Statistics*, Vol. 28, No. 2, pp. 219-231.

[9] McLachlan, G. and T. Krishnan (1997), *The EM Algorithm and Extensions*, John Wiley and Sons, New York

[10] Morey, E., J. Thacher, and W. Breffle (2006). "Using Angler Characteristics and Attitudinal Data to Identify Environmental Preference Classes: A Latent-Class Model," *Environmental & Resource Economics*, Vol. 34, pp. 91-115.

[11] Provencher, B., K. Baerenklau, and R. Bishop (2002). "A Finite Mixture Logit Model of Recreational Angling with Serially Correlated Random Utility," *American Journal of Agricultural Economics*, Vol. 84, No. 4, pp. 1066-1075.

[12] Train, K. (2008). "EM Algorithms for Nonparametric Estimation of Mixing Distributions," *Journal of Choice Modelling*, Vol. 1, No. 1, pp. 40-69.

[13] Train, K. (2009). *Discrete Choice Methods with Simulation*, Cambridge University Press, 2nd edition.

	Table 1: Welfare Performance - Experiment #1									
	a. Parameter Setting[a]									
β_{0R}	1.00 [1.40]		1.50 [2.30]		2.00 [3.80]		2.50 [6.26]		3.00 [10.32]	
β_{0N}	0.50 [0.85]		0.75 [1.09]		1.00 [1.40]		1.25 [1.79]		1.50 [2.30]	
	b. Mean Percentage Error in CS									
s_R	LCCM	SCCM	LCCM	SCCM	LCCM	SCCM	LCCM	SCCM	LCCM	SCCM
0.10	2.4	2.9	1.8	7.9	1.2	5.9	1.3	2.0	-0.1	-0.2
0.25	0.6	6.1	1.9	16.8	0.5	12.0	0.1	2.9	0.3	0.5
0.50	0.5	11.0	0.6	26.0	0.2	19.2	-0.4	2.0	0.1	1.3
0.75	0.9	15.6	-0.2	32.9	0.5	23.8	0.2	3.2	0.3	0.4
0.90	1.2	17.3	0.0	36.7	0.1	25.6	0.1	2.7	0.0	0.1

[a]Corresponding unconditional group mean trips in square brackets.

	Table 2: Welfare Performance - Experiment #2									
	a. Parameter Setting[b]									
ϕ_0	0.50		0.75		1.00		1.50		2.00	
β_{0R}	1.00 [1.40]		1.29 [1.86]		1.50 [2.30]		1.80 [3.11]		2.00 [3.80]	
β_{0N}	2.00 [3.80]		1.71 [2.85]		1.50 [2.30]		1.20 [1.71]		1.00 [1.40]	
	b. Mean Percentage Error in CS									
s_R	LCCM	SCCM	LCCM	SCCM	LCCM	SCCM	LCCM	SCCM	LCCM	SCCM
0.10	0.9	1.1	1.3	2.0	1.2	3.6	1.4	6.4	1.2	5.9
0.25	0.1	1.7	1.0	6.0	1.0	9.3	1.5	14.1	0.5	12.0
0.50	0.0	5.0	0.7	12.5	0.8	19.0	1.4	24.8	0.2	19.2
0.75	0.3	9.6	-0.1	19.9	0.3	28.8	0.1	31.1	0.5	23.8
0.90	0.4	14.7	0.8	27.0	0.7	35.8	0.6	35.9	0.1	25.6

[b]Corresponding unconditional group mean trips in square brackets.

Table 3: Summary Statistics					
Variable	Model Variable	Mean	Std. Dev.	Min	Max
Total Day Trips (2002)	y_i	1.664	4.924	0.000	60.000
Travel Cost ($10's)	P_i	3.057	2.095	0.185	17.067
Income ($10000's)	Y_i	6.300	5.904	0.500	32.500
Age (10 years)	$Z_{Age,i}$	5.290	1.744	1.550	8.750
College	$Z_{Educ,i}$	0.396	0.489	0.000	1.000

Table 4: Parameter Estimates							
Model	Class	$\hat{\beta}_0$	$\hat{\beta}_P$	$\hat{\beta}_Y$	$\hat{\beta}_{Z,Age}$	$\hat{\beta}_{Z,Educ}$	$\hat{s}_R$
SCCM		2.664	-0.740	0.033	-0.170	-0.030	n.a.
		(0.006)	(0.002)	(0.003)	(0.001)	(0.004)	
Constrained LCCM		2.626	-0.741	0.033	-0.171	-0.469	0.361
		(0.084)	(0.023)	(0.004)	(0.014)	(0.046)	(0.097)
LCCM	$C_i = N$	3.193	-0.816	0.032	-0.079	-1.621	
		(0.017)	(0.006)	(0.001)	(0.003)	(0.033)	
LCCM	$C_i = R$	0.866	-0.750	0.014	-0.258	4.093	0.386
		(0.125)	(0.039)	(0.002)	(0.018)	(0.062)	(0.008)