TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
2023, Vol. 2677(3) 1137–1150

DOI: 10.1177/03611981221121266
journals.sagepub.com/home/trr
SAGE

On Assignment to Classes in Latent Class Logit Models

Wangwei Wu[1] and Ricardo A. Daziano[2]

Abstract

Random parameter logit models address unobserved preference heterogeneity in discrete choice analysis. The latent class logit model assumes a discrete heterogeneity distribution, by combining a conditional logit model of economic choices with a multinomial logit (MNL) for stochastic assignment to classes. Whereas point estimation of latent class logit models is widely applied in practice, stochastic assignment of individuals to classes needs further analysis. In this paper we analyze the statistical behavior of six competing class assignment strategies, namely: maximum prior MNL probabilities, class drawn from prior MNL probabilities, maximum posterior assignment, drawn posterior assignment, conditional individual-specific estimates, and conditional individual estimates combined with the Krinsky–Robb method to account for uncertainty. Using both a Monte Carlo study and two empirical case studies, we show that assigning individuals to classes based on maximum MNL probabilities behaves better than randomly drawn classes in market share predictions. However, randomly drawn classes have higher accuracy in predicted class shares. Finally, class assignment based on individual-level conditional estimates that account for the sampling distribution of the assignment parameters shows superior behavior for a larger number of choice occasions per individual.

Keywords

planning and analysis, transportation demand forecasting, choice models, demand estimation, forecasts/forecasting, transportation demand management, travel demand modeling

Random parameter logit models are the main empirical strategy for addressing unobserved preference heterogeneity in discrete choice analysis (*1*, *2*). Heterogeneity (or mixing) distributions are usually taken from parametric families, either continuous or discrete. In the latent class conditional logit (LCL) model, preference parameters are assumed to have a heterogeneity distribution that is discrete (*3–6*). In fact, LCL choices are governed by a conditional logit model, whereas assignment to classes is determined by a multinomial logit (MNL) specification (*2*).

Mixed logit models with parametric and continuous heterogeneity distributions—such as normally distributed parameters—provide preference estimates that can be hard to interpret or have inference problems (*7–11*). However, the discrete nature of the estimates of latent class logit models makes inference easier in relation to both interpretability and derivation of welfare measures such as willingness-to-pay (WTP) metrics. As a result, use of LCL models has proven a popular choice in empirical work. After seminal work that spread the use of LCL models in choice modeling, recent examples include a vast variety of applications, from preference location by crime offenders to valuation of endangered marine species, and recreational demand in the Alps, just to give three examples of the range of problems for which latent class logit models have been applied beyond more traditional applications (such as travel mode choice, vehicle ownership, and residential choice) (*5*, *6*, *12–21*).

Despite the growing popularity of LCL models, there is a need for better understanding of statistical inference with the model. Whereas Romero-Espinosa et al. studied statistical and asymptotic behavior of interval estimates of conditional LCL preference parameters at the individual level, further analysis is needed to characterize the

[1]Systems Engineering, Cornell University, NY
[2]Civil and Environmental Engineering, Cornell University, NY

Corresponding Author:
Wangwei Wu, ww433@cornell.edu

behavior of empirical strategies for assigning individuals to a specific class (*22*). The latent class assignment problems can relate to random draw sampling methods and Bayesian procedures of preference parameter estimation. Using different class assignment strategies may affect the accuracy of predicting latent class shares and chosen alternative shares. In addition to the popular class assignment method using prior probabilities and a maximum-value-draw strategy, there also exist other methods which need to be studied (*23–25*). In this paper we focus on these issues and analyze the statistical behavior of six competing class assignment strategies that are econometrically valid but have not been fully examined in the literature, namely: maximum prior MNL probabilities, class drawn from prior MNL probabilities, maximum posterior assignment, drawn posterior assignment, conditional individual-specific estimates, and conditional individual estimates combined with the Krinsky–Robb (KR) method (*26*). Our contribution lies in the methodical comparison of these class assignment strategies for latent class logit models.

This paper is organized as follows. The next section reviews the latent class logit model and describes the six empirical strategies for assignment to classes mentioned above. The section after that uses a Monte Carlo study to compare the empirical performance across all strategies in different regimes. Our comparison is then supplemented in the penultimate section with two case studies, one focused on response to automated electric vehicles and the other on consumer valuation of emission savings when purchasing a new vehicle. The final section concludes.

The Latent Class Logit Model

Following a general choice setting, we assume that consumer i makes discrete choices among a set of J alternatives and T choice occasions with utility maximization as the decision rule. The truncated indirect utility functions of the alternatives are characterized by individual preferences. Whereas the simplest choice models impose preference homogeneity, more flexible specifications address unobserved preference heterogeneity through the consideration of preference parameters that are random. In latent class logit models, consumers' preferences are assumed to be heterogeneous with discrete heterogeneity distributions. In fact, in LCL models consumers are assumed to belong to clusters that are modeled as latent classes. Within each class, the same preference parameters are shared as in a standard conditional logit model.

On the one hand, conditional on class q, LCL choices assume the following conditional logit choice probability:

$$P_{ijt|q} = \frac{\exp\left(\mathbf{x}'_{ijt}\boldsymbol{\beta}_q\right)}{\sum_j \exp\left(\mathbf{x}'_{ijt}\boldsymbol{\beta}_q\right)}, \tag{1}$$

where

$\mathbf{x}_{ijt}$ = the attributes of alternative j for consumer i in choice occasion t, and

$\boldsymbol{\beta}_q$ = the preference parameters of latent class q.

On the other hand, and since the underlying class of an individual is not observed, class assignment assumes the following MNL probabilities:

$$w_{iq} = \frac{\exp\ \mathbf{h}'_i\gamma_q}{\sum_{q=1}^{Q}\exp\ \mathbf{h}'_i\gamma_q};\ \ q = 1, \dot{s}, Q,\ \gamma_1 = \mathbf{0}, \tag{2}$$

where

$\mathbf{h}_i$ = socio-demographic characteristics of consumer i, and

γ_q = the parameter vector which summarizes how these characteristics are linked to a higher or lower likelihood of the consumer belonging to latent class q.

Empirical Strategies for Assignment to Classes

Using Bayes' theorem, posterior MNL probability for assignment to classes can be derived as (see Train) (*27*):

$$p(\boldsymbol{\beta} = \boldsymbol{\beta}_q|\mathbf{y}_i, \mathbf{X}_i) = \frac{\widehat{w}_{iq}\prod_{t=1}^{T}\prod_{j=1}^{J}\left[\frac{\exp\left(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q\right)}{\sum_{j=1}^{J}\exp\left(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q\right)}\right]^{y_{ijt}}}{\sum_{q=1}^{Q}\widehat{w}_{iq}\prod_{t=1}^{T}\prod_{j=1}^{J}\left[\frac{\exp\left(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q\right)}{\sum_{j=1}^{J}\exp\left(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q\right)}\right]^{y_{ijt}}}, \tag{3}$$

where

$\hat{\boldsymbol{\beta}}_q$ = the point estimate vector of preferences for latent class q, and

$\hat{w}_{iq}$ = the fitted MNL probability of consumer i belonging to that class.

Note that $\hat{w}_{iq}$ is a prior probability of class assignment that is based on the estimated parameter vectors $\hat{\gamma}_{q=1,\ldots,Q}$.

Given the prior and posterior MNL class assignment probabilities shown above, four strategies for class assignment can be derived. The first two strategies are based on prior class assignment probabilities alone, which only depend on consumer information. In contrast to prior strategies, posterior assignment strategies use the posterior MNL probabilities wherein consumers' choices are embedded. For each of the two types of assignment probabilities, we introduce two strategies for individual-level latent class assignment: *1*) assign a certain consumer a class according to maximum probability (as in Scarpa

and Thiene) (*23*); and *2*) randomly draw a class for a consumer from the estimated latent class probabilities.

Strategy 1: maximum MNL assignment probability. This strategy uses prior MNL probabilities across latent classes stated in Equation 2. Given estimated parameters and a consumer's socio-demographic information, we then compute the prior MNL probabilities $\hat{w}_{iq}$ for them and assign them the class of which the probability value is the maximum among all $q = 1, \ldots, Q$.

Strategy 2: drawn MNL assignment. The second strategy uses prior MNL probabilities through Equation 2 as well. However, this strategy makes an independent random draw for latent class assignment from the obtained prior probabilities $\hat{w}_{iq}$ for each particular consumer i, instead of taking the maximum value. Thus, consumers' assigned classes come directly from those random draws.

Strategy 3: maximum posterior assignment probability. Individuals are assigned to the class with the maximum posterior probability (Equation 3), evaluated at point estimates.

Strategy 4: drawn posterior assignment. Similar to the second strategy, the forth strategy makes an independent random draw for the class assignment according to posterior probabilities obtained through Equation 3 for each particular consumer i. Then, individuals are assigned to the randomly picked classes.

In addition to these four strategies based on actual class assignment probabilities, consumers can also be assigned to a class by consideration of their conditional point estimate

$$\hat{\boldsymbol{\beta}}_{i|\mathbf{y}_i, \mathbf{x}_i, \hat{\boldsymbol{\theta}}}.$$

Conditional Estimates at the Individual Level

From the posterior MNL probabilities shown in the previous section, it is possible to derive conditional estimates of preference parameters at the individual level (*26*, *27*). Equation 4 for the conditional preferences is essentially an expected preference parameter vector over the posterior class assignment probabilities from Equation 3:

$$\hat{\boldsymbol{\beta}}_{i|\mathbf{y}_i, \mathbf{X}_i, \hat{\boldsymbol{\theta}}} = \mathbb{E}(\boldsymbol{\beta}_i|\mathbf{y}_i, \mathbf{X}_i, \hat{\boldsymbol{\theta}}) = \sum_{q=1}^{Q} \hat{\boldsymbol{\beta}}_q \frac{\widehat{w}_{iq} \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q)}{\sum_{j=1}^{J} \exp(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q)} \right]^{y_{ijt}}}{\sum_{q=1}^{Q} \widehat{w}_{iq} \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q)}{\sum_{j=1}^{J} \exp(\mathbf{x}'_{ijt}\hat{\boldsymbol{\beta}}_q)} \right]^{y_{ijt}}}, \quad (4)$$

where

$\hat{\boldsymbol{\theta}}$ = the point estimate of a meta-parameter vector $\boldsymbol{\theta}$ that includes parameters $\boldsymbol{\gamma}$.

In addition to the point estimates introduced above, it is also worth noting that inference on the expected preference parameters can be taken into account in the uncertainty in the determination of $\hat{\boldsymbol{\theta}}$ (*22*, *26*).Since $\boldsymbol{\theta}$ is multivariate normally distributed with mean $\boldsymbol{\theta}$ and covariance $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, the corresponding formal expression of the conditional expectation of the individual parameters is the following:

$$\mathbb{E}[\boldsymbol{\beta}_i|y_i, \boldsymbol{\theta}_i, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}] = \int_{\boldsymbol{\theta}} \mathbb{E}[\boldsymbol{\beta}_i|y_i, \boldsymbol{\theta}] \mathcal{N}(\boldsymbol{\theta}|\boldsymbol{\theta}_i, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) d\boldsymbol{\theta}. \quad (5)$$

To circumvent the heavy computations associated with evaluation of the multi-dimensional integral, a sampling method can be set to calculate the conditional estimates that accounts for the sampling distribution of the meta-parameter $\boldsymbol{\theta}$. This approximation is close to the KR method, which was introduced by Greene et al. for a random parameter logit model with continuous heterogeneity distributions (*26*, *28*). In practice, the computation steps of the empirical counterpart of the expectation in Equation 5 optimize toward $\hat{\mathbb{E}}[\boldsymbol{\beta}_i|y, \boldsymbol{\theta}_i, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}] = \frac{1}{R}\sum_{r=1}^{R} \mathbb{E}[\boldsymbol{\beta}_i|y_i, X_i\boldsymbol{\theta}^r$
where

R = the number of simulation repetitions, where $\boldsymbol{\theta}$ is sampled R times and the average of those conditional estimates at the individual level is taken.

In sum, the possibility of working with conditional estimates at the individual level provides another empirical strategy of assignment to classes, where a given individual is assigned to the class with population point estimates that are closest to the conditional point estimates. Thus, it is possible to implement the following:

Strategy 5: conditional individual-specific estimates. Individual i is assigned to the class with population parameters that are closest to $\mathbb{E}(\boldsymbol{\beta}_i|\mathbf{y}_i, \mathbf{X}_i, \hat{\boldsymbol{\theta}})$, which uses Equation 4.

Strategy 6: conditional individual-specific estimates with KR implementation. Individual i is assigned to the class with population parameters that are closest to $\frac{1}{R}\sum_{r=1}^{R} \mathbb{E}[\boldsymbol{\beta}_i|y_i, X_i\boldsymbol{\theta}^r$, which further assumes uncertainty in meta-parameters $\boldsymbol{\theta}$ and takes simulated integral in Equation 5 by sampling meta-parameter values.

Monte Carlo Study

Simulation Plan

To compare the performance of the several competing latent class assignment strategies described in the previous section, we conducted a Monte Carlo study. Similar to the simulation done by Sarrias and Daziano, we also assumed multiple sets of scenarios with three alternatives

($J = 3$) for hypothetical individuals without individual-level socio-demographic data (*22*). Whereas Sarrias and Daziano focused their study on interval estimation, we change the focus to choice predictions based on class assignment (*22*).

Formally, the true latent utility of alternative j for individual i in choice occasion t is implemented as:

$$U_{ijt} = \boldsymbol{\beta}_{1i}x_{1ijt} + \boldsymbol{\beta}_{2i}x_{2ijt} + \epsilon_{ijt}, \tag{6}$$

where

$\mathbb{1}_{ijt}$ = a Type-1 extreme value distributed preference shock,

x_1 = an independent and standard normally distributed attribute, and

x_2 = an attribute which is assumed to be a dummy variable created from an indicator function $\mathbb{1}(u<0.5)$ of a uniformly distributed random variable $0<u<1$.

For the preference parameters, we assumed the following discrete unobserved heterogeneity distributions ($Q = 3$):

$$\boldsymbol{\beta}_{1i} = \begin{cases} 2 & \text{with probability } 0.25 \\ 0 & \text{with probability } 0.5 \\ 2 & \text{with probability } 0.25 \end{cases} \tag{7}$$

$$\boldsymbol{\beta}_{2i} = \begin{cases} 0.5 & \text{with probability } 0.25 \\ 1 & \text{with probability } 0.5 \\ 1.5 & \text{with probability } 0.25 \end{cases} \tag{8}$$

Simulated databases were constructed for a baseline size $N = 1{,}000$. For each individual, five scenarios with differing numbers of choice occasions were created, namely $T \in \{1, 5, 10, 20, 50\}$. For each scenario, 300 independently sampled databases were implemented.

Given this simulation plan, we were able to compare statistical results across different latent class assignment strategies mainly from the perspective of preference space estimates, assigned latent class shares, and predicted choice shares. Specifically, we focused on the following statistics:

$$\text{Mean} = \hat{\boldsymbol{\beta}}_S = \frac{1}{S}\sum_{s=1}^{S} \frac{1}{N}\sum_{i=1}^{N} \hat{\boldsymbol{\beta}}_{is}$$

$$\text{(Absolute)Bias} = \frac{1}{S}\sum_{s=1}^{S} \frac{1}{N}\sum_{i=1}^{N} |\hat{\boldsymbol{\beta}}_{is} - \boldsymbol{\beta}_i|$$

$$\text{Absolute percentage bias} = \frac{1}{S}\sum_{s=1}^{S} \frac{1}{N}\sum_{i=1}^{N} \frac{|\hat{\boldsymbol{\beta}}_{is} - \boldsymbol{\beta}_i|}{\boldsymbol{\beta}_i}$$

$$\text{SE} = \frac{1}{S}\sum_{s=1}^{S} \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (\hat{\boldsymbol{\beta}}_{is} - \hat{\boldsymbol{\beta}}_S)^2},$$

where

$\boldsymbol{\beta}_i$ = the true parameters for each individual i, and

$\hat{\boldsymbol{\beta}}_{is}$ = individual-level parameters obtained through the different latent class assignment strategies under analysis.

Besides this series of statistics, we also considered the empirical coverage probability (COV). COVs tell the proportion of simulated samples for which the estimated 95% interval includes the true individual-level parameter. A correct interval inference should produce a 95% coverage, with lower or higher figures indicating respective narrow or wide estimated intervals on average.

On the other hand, to evaluate different strategies' behavior on market shares and latent class shares, the following additional statistics were calculated:

$$\text{RMSE} = \frac{1}{S}\sum_{s=1}^{S} \sqrt{\frac{1}{N}\sum_{i=1}^{N} \frac{1}{T}\sum_{t=1}^{T} (1 - P_{ijt}(\hat{\boldsymbol{\beta}}_{is}|y_{it}, X_{it}))^2}$$

$$\text{MAE} = \frac{1}{S}\sum_{s=1}^{S} \frac{1}{N}\sum_{i=1}^{N} \frac{1}{T}\sum_{t=1}^{T} (1 - P_{ijt}(\hat{\boldsymbol{\beta}}_{is}|y_{it}, X_{it}))$$

where

y_{it} = the true choice made by individual i in choice occasion t.

In addition, consideration is given to average percent correctly predicted (PCP). In fact, just as with assignment probabilities, we present PCP results obtained through two approaches, namely: (1) choosing an alternative with a maximum choice probability, and (2) randomly picking an alternative according to the estimated choice probabilities. In the tables, these two statistics are denoted by PCP(max) and PCP(drawn), respectively.

Finally, Brier scores were calculated to evaluate prediction performance (*29*, *30*):

$$\text{BrierScore} = \frac{1}{N \times T \times J}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J} (\mathbf{1}\{y_{i,t} = j\} - \hat{p}_{i,t,j})^2 \tag{9}$$

where

$\mathbf{1}\{y_{i,t} = j\}$ = the true choices made by individual i on occasion t, and

$\hat{p}_{i,t,j}$ = the estimated choice probability, correspondingly. Lower Brier scores are an indication of better predictive accuracy.

Results

Tables 1 and 2 summarize parameter recovery results obtained through the six competing class assignment strategies. Table 1 displays aggregate Monte Carlo results for the estimation of the preference parameter $\boldsymbol{\beta}_1$. To understand this aggregate analysis, consider that point estimates of $\boldsymbol{\beta}_1$ have been contrasted with the true expected value of the parameter which is equal to 0.

Table 1. Latent Class Conditional Logit (LCL) Model—Three Classes: Aggregate Parameter Recovery (β_1)

	Mean		Bias		CP	
	Mean	SE	Mean	SE	Mean	SE
$T = 1$						
Maximum MNL assignment probability	0.0678	0.5246	1.2627	0.9344	0.9533	0.0890
Drawn MNL assignment	−0.2142	5.9109	3.3305	5.9542	0.9509	0.1563
Maximum posterior assignment probability	−0.1092	2.6107	1.5107	2.1677	0.9505	0.1552
Drawn posterior assignment	−0.1716	5.9121	2.9994	5.7863	0.9507	0.1566
Conditional individual-specific estimates	−0.1966	3.4742	2.5171	3.4366	0.9500	0.1217
Conditional individual-specific estimates + KR	−0.0626	3.1302	2.4203	3.1569	1.0000	0.0000
$T = 5$						
Maximum MNL assignment probability	0.0034	0.0379	1.0190	0.9821	0.9533	0.0890
Drawn MNL assignment	0.0042	1.4143	1.5209	1.2956	0.9509	0.1350
Maximum posterior assignment probability	0.0078	1.4326	0.3522	0.6665	0.9500	0.1364
Drawn posterior assignment	0.0026	1.4159	0.4695	0.7720	0.9512	0.1344
Conditional individual-specific estimates	0.0038	1.2664	0.4201	0.4901	0.9500	0.1746
Conditional individual-specific estimates + KR	0.0031	1.2696	0.4162	0.4900	1.0000	0.0000
$T = 10$						
Maximum MNL assignment probability	−0.0016	0.0178	1.0089	0.9917	0.9533	0.0890
Drawn MNL assignment	0.0018	1.4203	1.5129	1.3116	0.9505	0.1329
Maximum posterior assignment probability	0.0035	1.4258	0.1288	0.4094	0.9504	0.1332
Drawn posterior assignment	0.0009	1.4186	0.1735	0.4954	0.9508	0.1324
Conditional individual-specific estimates	−0.0003	1.3697	0.1592	0.3365	0.9500	0.1533
Conditional individual-specific estimates + KR	−0.0005	1.3665	0.1607	0.3369	1.0000	0.0000
$T = 20$						
Maximum MNL assignment probability	0.0269	0.0503	1.0182	0.9533	0.9867	0.0890
Drawn MNL assignment	−0.0161	1.4014	1.4977	0.9508	0.9506	0.1312
Maximum posterior assignment probability	−0.0131	1.4000	0.0769	0.2119	0.9508	0.1306
Drawn posterior assignment	−0.0139	1.3990	0.0833	0.2449	0.9506	0.1315
Conditional individual-specific estimates	−0.0142	1.3920	0.0811	0.1932	0.9500	0.1376
Conditional individual-specific estimates + KR	−0.0018	1.4076	0.0504	0.1484	0.9999	0.0002
$T = 50$						
Maximum MNL assignment probability	0.0079	0.0192	1.0083	0.9938	0.9533	0.0890
Drawn MNL assignment	−0.0035	1.4055	1.5019	1.3123	0.9507	0.1327
Maximum posterior assignment probability	−0.0039	1.4058	0.0303	0.0449	0.9500	0.1339
Drawn posterior assignment	−0.0039	1.4058	0.0303	0.0449	0.9500	0.1339
Conditional individual-specific estimates	−0.0039	1.4058	0.0303	0.0443	0.9500	0.1340
Conditional individual-specific estimates + KR	−0.0046	1.3665	0.0145	0.0142	1.0000	0.0000

Note: SE = standard error; CP = empirical coverage; KR = Krinsky–Robb method; MNL = multinomial logit.

Recall that the law of large numbers ensures that an unbiased estimator has its estimates converge to the true value as sample size goes infinitely large. Thus, we should observe that bias values shrink as the choice situation number T goes large.

In the case of analyzing choice predictions, each individual has been assigned to a class; the point estimates of the respective class can be used to evaluate the conditional logit choice probabilities of Equation 1 which can be then plugged into the expressions of root mean square error (RMSE) and maximum absolute error (MAE) as well as exploited for making an actual predicted alternative either using the maximum probability rule or the drawn alternative method. As shown in Table 1, and as expected, we can observe that assignment strategies related to posterior probabilities and conditional estimates generate lower bias when T goes up. Nevertheless, the first two assignment strategies, which are associated with prior assignment probabilities, do not show lower bias values when T is large. This observation results from the setting of our Monte Carlo study plan. In our simulated databases, we only randomly simulated alternative attribute levels but not the individual-level socio-demographic data. This kind of setting causes prior probabilities to be unable to effectively identify each individual's latent class.

Comparing bias values across class assignment strategies, strategy 3 displays the lowest bias when $T = 5, 10$. This result implies that posterior choice probabilities behave better in the cases of limited individual-level choice information (i.e., T is at a moderate value). As choice situations T becomes larger ($T = 20, 50$), the

conditional individual-specific estimates with the KR method estimates produce the lowest bias values on the aggregate estimations of parameter $\boldsymbol{\beta}_1$.

Given a certain kind of latent class assignment probabilities, say, for example, prior probabilities, the maximum probability rule methods (strategies 1 and 3) bring lower bias values compared with those of the random draw method (strategies 2 and 4). This phenomenon can be explained through a linear programming problem in a 1 norm space. Given a set of class probabilities $\boldsymbol{p}_{iq}$ of an individual i of class q, a predictor's job is to solve the following linear programming problem: $\max_{||a||_1 \leq 1} : \sum_{q=1}^{Q} a_q p_{iq} \text{Bias}_q$,

where

$\boldsymbol{a} = (a_1, \ldots, a_Q)$ = the probabilities of latent classes by a predictor.

The feasible area of a, $||a||_1 \leq 1$ shows that these prediction probabilities are required to follow a one-sum rule. To this extent, an optimal solution to this linear programming formulation implies that the maximum probability assignment strategy suggests the best classification for individuals when a predictor uses an unbiased Maximum Likelihood Estimation (MLE) estimator.

In Table 2, we illustrate detailed parameter recovery based on the databases with $T = 50$. In this case, comparisons are made with respect to each of the three possible values of the parameter. Similar to the situation in Table 1, strategies with prior probabilities exhibit worse performance than that of other strategies, with all other strategies having corresponding MEV metrics close to the true values.

In Table 3, where we analyze correct disaggregate choice predictions by class assignment strategy, we observe that summary statistics including RMSE, MAE, and PCP do not become better when T goes up. All posterior assignment strategies (strategies 3 and 4) perform equivalently well in choice prediction. However, for $T = 1$, exploiting conditional estimates at the individual level makes superior choice inference, especially when using the maximum probability rule in the calculation of PCP. Combined with the KR method for accounting for the sampling distribution of the class assignment parameters, for $T = 1$, PCP(max) achieves a value over 81%. In fact, PCP(max) values are larger than PCP(drawn) across all assignment strategies. This phenomenon can be explained by the same linear programming in the 1-norm space argument used before.

In addition, Table 4 reports that the aggregate class shares of all strategies closely approximate the true class shares even when $T = 5$. The exception is maximum prior assignment probability, which is expected, as Class 2 having the greatest share, the result is an almost deterministic assignment to Class 2. Another reason behind this is also similar to the case in Table 1 where the maximum MNL assignment probability assignment strategy predicts $\boldsymbol{\beta}_i$ worse. The lack of individual background information causes inefficiency in class assignment through maximum prior probabilities.

As a whole, we can see that, even though a larger T results in more accurate estimated parameters, good aggregate class shares can be achieved even with a small T. Table 4 also shows that the maximum MNL (prior) assignment probability (strategy 1) predicts aggregate class shares significantly worse than other strategies.

Looking at disaggregate correct class assignment, Table 5 reports the proportions of correct class assignments for all assignment strategies. Again, using prior probabilities (strategies 1 and 2) performs significantly worse than other strategies because of the lack of individual socio-demographic information in the Monte Carlo study setting. These observations match the analysis from Tables 1 and 2, as elaborated previously. For instance, while increasing T in general improves correct posterior assignment, even with a relatively low number of choice occasions ($T = 5$) correct assignment to classes is over 85% and can achieve 95% for $T = 10$ (which is a common number found in practice for choice experiments).

Finally, Table 6 reports Brier scores of individual-level choice predictions made by the different class assignment strategies. These score values reflect the accuracy of choice predictions, and a lower value implies more accurate predictions. These figures again confirm that strategies embedded with posterior probabilities have rather accurate choice predictions. However, the posterior probability strategies do not necessarily generate lower Brier scores when T goes large.

Empirical Case Studies

Data

To supplement the Monte Carlo study, the six class assignment strategies were applied to two empirical datasets using actual choice experiments. Both case studies relate to purchase preferences toward low-emission vehicles. Whereas the first case study, in addition to electrification, focuses on automated features, the second case study uses data that were collected to analyze economic valuation of carbon abatement. There are two main differences with the simulation setting. First, true parameters and true classes are, of course, unknown. Second, class assignment is informed by socio-demographic characteristics of the consumers.

Case Study 1: Automated Electric Vehicles

We first use microdata from a choice experiment that was designed to analyze early-market response to vehicle automation (*31*, *32*). The choice experiment was designed

Table 2. Latent Class Conditional Logit (LCL) Model—Three Classes: Detailed Parameter Recovery (T = 50, N = 1,000)

	True	MEV	MAB	APB	FSSE
Maximum MNL assignment probability					
Class 1: β_1	−2.0	0.0063	2.0063	1.0031	2.0063
Class 2: β_1	0.0	0.0064	0.0063	Inf	0.0063
Class 3: β_1	2.0	0.0064	1.9936	0.9968	1.9936
Class 1: β_2	−0.5	1.0182	1.5182	1.0031	1.5182
Class 2: β_2	1.0	1.0182	0.0182	0.0182	0.0182
Class 3: β_2	1.5	1.0182	0.4817	0.3211	0.4817
Drawn MNL assignment					
Class 1: β_1	−2.0	−0.0990	1.9009	0.9504	2.3668
Class 2: β_1	0.0	0.1665	0.9713	Inf	1.3889
Class 3: β_1	2.0	0.0679	1.9429	0.9714	2.3986
Class 1: β_2	−0.5	0.7009	1.2009	2.4019	1.4305
Class 2: β_2	1.0	0.8468	0.4621	0.4621	0.7312
Class 3: β_2	1.5	0.7834	0.7263	0.4842	1.0406
Maximum posterior assignment probability					
Class 1: β_1	−2.0	−1.9534	0.0465	0.0233	0.0465
Class 2: β_1	0.0	0.0063	0.0063	Inf	0.0063
Class 3: β_1	2.0	2.0203	0.0203	0.0101	0.0203
Class 1: β_2	−0.5	−0.4933	0.0066	0.0133	0.0066
Class 2: β_2	1.0	1.0182	0.0182	0.0182	0.0182
Class 3: β_2	1.5	1.5182	0.0182	0.0121	0.0182
Drawn posterior assignment					
Class 1: β_1	−2.0	−1.9534	0.0465	0.0232	0.0465
Class 2: β_1	0.0	0.0064	0.0064	Inf	0.0064
Class 3: β_1	2.0	2.0203	0.0203	0.0101	0.0203
Class 1: β_2	−0.5	−0.4933	0.0066	0.0132	0.0066
Class 2: β_2	1.0	1.0182	0.01826	0.0182	0.0182
Class 3: β_2	1.5	1.5182	0.01822	0.0121	0.0182
Conditional individual-specific estimates					
Class 1: β_1	−2.0	−1.9534	0.0465	0.0232	0.0465
Class 2: β_1	0.0	0.0064	0.0063	Inf	0.0064
Class 3: β_1	2.0	2.0203	0.0203	0.0102	0.0203
Class 1: β_2	−0.5	−0.4933	0.0066	0.0133	0.0066
Class 2: β_2	1.0	1.0182	0.01826	0.0182	0.0182
Class 3: β_2	1.5	1.5182	0.01822	0.0121	0.0182
Conditional individual-specific estimates—KR					
Class 1: β_1	−2.0	−1.9915	0.0084	0.0042	0.0084
Class 2: β_1	0.0	−0.0013	0.0016	Inf	0.0029
Class 3: β_1	2.0	2.0358	0.0358	0.0179	0.0358
Class 1: β_2	−0.5	−0.4551	0.0448	0.0896	0.0448
Class 2: β_2	1.0	0.9950	0.0049	0.0049	0.0049
Class 3: β_2	1.5	1.4785	0.0214	0.0143	0.0214

Note: APB = absolute percentage bias; FSSE = finite sample standard error; Inf = Infinity; KR = Krinsky–Robb method; MAB = mean absolute bias; MEV = mean estimated value; MNL = multinomial logit.

around three levels of automation of private light duty vehicles, namely: no automation, partial automation, and full automation. Automation was allowed for low-emission powertrains (hybrid electric, plug-in hybrid, and full battery electric). Details about both the design of the experiment and the data are provided by Daziano et al. (*31*). The conditional indirect utility for individual i choosing alternative j was specified in WTP space as:

$$U_{ij} = x'_{ij}\boldsymbol{\omega}_i \quad \alpha_i \text{price}_{ij} \quad \gamma_i \text{PVFC}_{ij} + \epsilon_{ij}$$

where

x_{ij} = vehicle design attributes,
PVFC = expected present value of fuel costs, and
ϵ_{ij} = i.i.d. distributed Type 1 extreme value as an error term.
The parameters $<\boldsymbol{\omega}_i, \alpha_i, \gamma_i>$ are assumed to be random with a discrete heterogeneity distribution.

In this dataset, there are 1,260 individuals (N = 1,260) that responded to a choice experiment with four alternatives, namely: a hybrid electric vehicle (HEV), a plug-in hybrid electric vehicle (PHEV), a battery electric vehicle (BEV), and a gasoline vehicle (GAS). In the statistical

Table 3. Latent Class Conditional Logit (LCL) Model—Three Classes: Prediction Metrics

	RMSE		MAE			
Class assignment strategy	Mean	SE	Mean	SE	PCP (max)	PCP (drawn)
$T = 1$						
Maximum MNL assignment probability	0.6636	0.0197	0.6343	0.0331	0.4095	0.3675
Drawn MNL assignment	0.6894	0.0094	0.6320	0.05820	0.3998	0.3569
Maximum posterior assignment probability	0.4811	0.0586	0.4254	0.0893	0.7255	0.5134
Drawn posterior assignment	0.5318	0.0237	0.4608	0.0782	0.62746	0.5173
Conditional individual-specific estimates	0.4325	0.0401	0.3678	0.0908	0.7906	0.5173
Conditional individual-specific estimates + KR	0.4058	0.0636	0.3334	0.1214	0.8142	0.5254
$T = 5$						
Maximum MNL assignment probability	0.6482	0.0019	0.6344	0.0142	0.4204	0.3675
Drawn MNL assignment	0.6874	0.0048	0.6353	0.0522	0.3992	0.3569
Maximum posterior assignment probability	0.5331	0.0038	0.4682	0.0651	0.6215	0.5134
Drawn posterior assignment	0.5434	0.0040	0.4792	0.0644	0.6070	0.5173
Conditional individual-specific estimates	0.5228	0.0036	0.4642	0.0589	0.6492	0.5173
Conditional individual-specific estimates + KR	0.5233	0.0036	0.4652	0.0583	0.64938	0.5254
$T = 10$						
Maximum MNL assignment probability	0.6479	0.0013	0.6342	0.0139	0.4219	0.3675
Drawn MNL assignment	0.6869	0.0046	0.6353	0.0518	0.3997	0.3569
Maximum posterior assignment probability	0.5415	0.0024	0.4772	0.0644	0.6081	0.5134
Drawn posterior assignment	0.5436	0.0025	0.4795	0.0641	0.6052	0.5172
Conditional individual-specific estimates	0.5372	0.0024	0.4749	0.0624	0.6231	0.5173
Conditional individual-specific estimates + KR	0.5377	0.0022	0.4757	0.0621	0.6231	0.5254
$T = 20$						
Maximum MNL assignment probability	0.6482	0.0017	0.6342	0.0141	0.4212	0.3676
Drawn MNL assignment	0.6866	0.0047	0.6356	0.0512	0.3997	0.3569
Maximum posterior assignment probability	0.5450	0.0065	0.4818	0.0640	0.6035	0.5134
Drawn posterior assignment	0.5451	0.0065	0.4820	0.0639	0.6033	0.51728
Conditional individual-specific estimates	0.5443	0.0065	0.4814	0.0637	0.6066	0.5173
Conditional individual-specific estimates + KR	0.5436	0.0046	0.4803	0.0637	0.6073	0.5254
$T = 50$						
Maximum MNL assignment probability	0.6480	0.0011	0.6343	0.0138	0.4217	0.3675
Drawn MNL assignment	0.6867	0.0040	0.6352	0.0516	0.4001	0.3569
Maximum posterior assignment probability	0.5445	0.0043	0.4809	0.0638	0.6036	0.5134
Drawn posterior assignment	0.5445	0.0043	0.4809	0.0638	0.6036	0.5172
Conditional individual-specific estimates	0.5445	0.0043	0.4809	0.0638	0.6036	0.5173
Conditional individual-specific estimates + KR	0.5440	0.0011	0.4802	0.0638	0.6041	0.5254

Note: KR = Krinsky–Robb method; MAE = mean absolute error; max = maximum; MNL = multinomial logit; PCP = percent correctly predicted; RMSE = root mean square error; SE = standard error.

analysis, we assume three latent classes ($Q = 3$) following the model selection strategy conducted in the article by Daziano et al., which was based on both Bayesian information criterion (BIC) and Akaike information criterion (AIC) (*31*).

Case Study 2: Emission Valuation in Vehicle Purchases

The second choice experiment was designed to analyze the impact of environmental information framings on the maximum WTP for CO_2 abatement (see Daziano et al.) (*33*). This experiment considered a binary choice between two unlabeled vehicles. The experimental attributes were: purchase cost, fuel costs per year, and CO_2 emissions (in pounds). The following indirect utility in preference space was adopted:

$$U_{ij} = \boldsymbol{\beta}_{i1}\text{price}_{ij} + \boldsymbol{\beta}_{i2}\text{PVFC}_{ij} + \boldsymbol{\beta}_{i3}\text{PVFE}_{ij} + \epsilon_{ij}$$

where

PVFE = present value of the expected future emissions, and

PVFC = present value of expected future fuel costs.

The parameter vector $\boldsymbol{\beta}_i$ is random with a discrete heterogeneity distribution.

In this dataset, there are 1,580 individuals with two alternative vehicles. We assume two latent classes according to the original probabilistic model selection study conducted in the paper by Daziano et al. (*33*).

Table 4. Latent Class Conditional Logit (LCL) Model—Three Classes: Class Shares

Class assignment strategy	Aggregate class shares			
	Class 1	Class 2	Class 3	χ^2 p-value
True class shares	0.25	0.50	0.25	na
$T = 1$				
Maximum MNL assignment probability	0.1267	0.6533	0.2200	0.0000
Drawn MNL assignment	0.2597	0.4416	0.2985	0.0002
Maximum posterior assignment probability	0.2391	0.4770	0.2837	0.0206
Drawn posterior assignment	0.2591	0.4414	0.2993	0.0003
Conditional individual-specific estimates	0.2392	0.4891	0.2716	0.0206
Conditional individual-specific estimates + KR	0.2028	0.5228	0.2743	0.0015
$T = 5$				
Maximum MNL assignment probability	0.0000	1.0000	0.0000	0.0000
Drawn MNL assignment	0.2518	0.4981	0.2500	0.9830
Maximum posterior assignment probability	0.2573	0.4853	0.2572	0.7550
Drawn posterior assignment	0.2528	0.4971	0.2500	1.0000
Conditional individual-specific estimates	0.2547	0.4904	0.2547	0.7550
Conditional individual-specific estimates + KR	0.2543	0.4888	0.2568	0.7727
$T = 10$				
Maximum MNL assignment probability	0.0000	1.0000	0.0000	0.0000
Drawn MNL assignment	0.2504	0.4984	0.2510	0.9928
Maximum posterior assignment probability	0.2519	0.4946	0.2534	0.9482
Drawn posterior assignment	0.2500	0.4997	0.2502	1.0000
Conditional individual-specific estimates	0.2516	0.4950	0.2533	0.9482
Conditional individual-specific estimates + KR	0.2526	0.4946	0.25287	0.9482
$T = 20$				
Maximum MNL assignment probability	0.0033	0.9533	0.0433	0.0000
Drawn MNL assignment	0.2513	0.5051	0.2435	0.7507
Maximum posterior assignment probability	0.2509	0.5051	0.2439	0.7820
Drawn posterior assignment	0.2506	0.5055	0.2437	0.7494
Conditional individual-specific estimates	0.2503	0.4949	0.2546	0.7820
Conditional individual-specific estimates + KR	0.2532	0.4884	0.2583	0.7402
$T = 50$				
Maximum MNL assignment probability	0.0067	0.9900	0.0033	0.0000
Drawn MNL assignment	0.2526	0.5029	0.2445	0.8831
Maximum posterior assignment probability	0.2524	0.5033	0.2441	0.9423
Drawn posterior assignment	0.2524	0.5033	0.2441	0.9423
Conditional individual-specific estimates	0.2499	0.5016	0.2484	0.9423
Conditional individual-specific estimates + KR	0.2525	0.4974	0.2500	0.9830

Note: KR = Krinsky–Robb method; MNL = multinomial logit; na = not applicable.

Table 5. Latent Class Conditional Logit (LCL) Model—Three Classes: Individual-Level Class Assignments

T	Max prior	Drawn prior	Max posterior	Drawn posterior	Conditional est.	Conditional est. + KR
PCP of class assignments						
1	0.4175	0.3616	0.5091	0.4687	0.5091	0.4683
5	0.5000	0.3717	0.8617	0.8018	0.8617	0.8618
10	0.5000	0.3720	0.9553	0.9331	0.9553	0.9554
20	0.4850	0.3766	0.9242	0.9210	0.9242	0.9721
50	0.4975	0.3777	0.9875	0.9875	0.9875	0.9725

Note: est. = estimate; KR = Krinsky–Robb method; max = maximum; PCP = percent correctly predicted.

Empirical Analysis

For both case studies, before applying each class assignment strategy, we trained a latent class logit model with a meta-parameter $Q = 3$ (i.e., three latent classes). The numerical results in Table 7 (choice predictions, case study 1) and Table 9 (choice predictions, case study 2)

Table 6. Latent Class Conditional Logit Model (LCL)—Three Classes: Brier Scores

T	Max prior	Drawn prior	Max posterior	Drawn posterior	Conditional est.	Conditional est. + KR
Brier scores						
1	0.2367	0.2573	0.1949	0.2067	0.1851	0.1827
5	0.2178	0.2464	0.1989	0.2015	0.1962	0.1962
10	0.2176	0.2465	0.2009	0.2014	0.1998	0.1997
20	0.2180	0.2455	0.2017	0.2017	0.2015	0.2013
50	0.2178	0.2460	0.2016	0.2016	0.2016	0.2014

Note: est. = estimate; KR = Krinsky–Robb method; Max = maximum.

Table 7. Case Study 1: Prediction Metrics—Vehicle Automation Discrete Choice Experiment

Class assignment strategy	PCP (drawn)	RMSE	MAE	PCP (max)
Fitted values (full sample)				
Maximum MNL assignment probability	0.3743	0.7083	0.6345	0.3756
Drawn MNL assignment	0.3297	0.7362	0.6632	0.3404
Maximum posterior assignment probability	0.5586	0.5351	0.4446	0.6476
Drawn posterior assignment	0.5910	0.5373	0.4468	0.6432
Conditional individual-specific estimates	0.3392	0.6847	0.6694	0.4448
Conditional individual-specific estimates + KR	0.3128	0.7515	0.6881	0.3160
20% testing (80% training)				
Maximum MNL assignment probability	0.3648	0.7207	0.6507	0.3683
Drawn MNL assignment	0.3389	0.7358	0.6683	0.3465
Maximum posterior assignment probability	0.5503	0.5321	0.4478	0.6524
Drawn posterior assignment	0.5548	0.5354	0.4503	0.6478
Conditional individual-specific estimates	0.3191	0.6875	0.6742	0.4375
Conditional individual-specific estimates + KR	0.3063	0.7523	0.6895	0.3102
LOOCV				
Maximum MNL assignment probability	0.3652	0.7089	0.6346	0.3762
Drawn MNL assignment	0.3309	0.7401	0.6691	0.3318
Maximum posterior assignment probability	0.5555	0.5350	0.4443	0.6479
Drawn posterior assignment	0.5535	0.5368	0.4457	0.6454
Conditional individual-specific estimates	0.3311	0.6845	0.6689	0.4483
Conditional individual-specific estimates + KR	0.3046	0.7556	0.6919	0.3036
k-fold cross validation ($k = 5$)				
Maximum MNL assignment probability	0.3509	0.7240	0.6530	0.3526
Drawn MNL assignment	0.3110	0.7514	0.6833	0.3122
Maximum posterior assignment probability	0.5504	0.5370	0.4466	0.6428
Drawn posterior assignment	0.5890	0.5387	0.4473	0.6397
Conditional individual-specific estimates	0.3172	0.6927	0.6768	0.4237
Conditional individual-specific estimates + KR	0.3048	0.7540	0.6912	0.3179
Repeated k-fold cross validation ($k = 5$)				
Maximum MNL assignment probability	0.3450	0.7259	0.6556	0.3512
Drawn MNL assignment	0.3244	0.7449	0.6759	0.3238
Maximum posterior assignment probability	0.5527	0.5367	0.4460	0.6451
Drawn posterior assignment	0.5541	0.5386	0.4476	0.6424
Conditional individual-specific estimates	0.3255	0.6912	0.6754	0.4275
Conditional individual-specific estimates + KR	0.3065	0.7554	0.6928	0.3118

Note: KR = Krinsky–Robb method; LOOCV = leave-one-out cross-validation; MAE = mean absolute error; max = maximum; MNL = multinomial logit; PCP = percent correctly predicted; RMSE = root mean square error.

summarize critical statistics for the evaluation of the class assignment strategies in relation to their ability to reproduce actual choices. Because correct out-of-sample predictions cannot be known in empirical data, five validation scenarios were considered, namely: fitted values for the whole sample (a predicted choice is built for each pseudoindividual in the sample); 80% of the sample used for estimation and the remaining 20% was used for

Table 8. Case Study 1: Market Shares—Vehicle Automation Discrete Choice Experiment. The Four Alternatives are: Hybrid Electric Vehicle (HEV), Plug-in Hybrid Electric Vehicle (PHEV), Battery Electric Vehicle (BEV), and Gasoline Vehicle (GAS)

	Aggregate shares				
Class assignment strategy	GAS	HEV	PHEV	BEV	p-value
Actual shares (full sample)	0.3692	0.1228	0.3100	0.1977	na
LCL fitted shares (full sample)					
Maximum MNL assignment probability	0.3463	0.1317	0.3103	0.2115	1.5448e-06
Drawn MNL assignment	0.3904	0.1203	0.2978	0.1913	2.5648e-04
Maximum posterior assignment probability	0.3676	0.1238	0.3093	0.1991	0.9673
Drawn posterior assignment	0.3703	0.1220	0.3109	0.1967	0.9819
Conditional individual-specific estimates	0.3831	0.1232	0.2973	0.1962	0.0156
Conditional individual-specific estimates + KR	0.3699	0.1207	0.3131	0.1962	0.0156
LCL fitted shares 20% testing					
Actual shares (testing sample)	0.3490	0.1229	0.3384	0.1895	na
Maximum MNL assignment probability	0.3539	0.1246	0.3154	0.2059	0.1103
Drawn MNL assignment	0.3558	0.1281	0.3090	0.2069	0.0317
Maximum posterior assignment probability	0.3490	0.1283	0.3150	0.2074	0.0720
Drawn posterior assignment	0.3510	0.1247	0.3205	0.2037	0.2601
Conditional individual-specific estimates	0.3780	0.1289	0.2898	0.2032	0.0001
Conditional individual-specific estimates + KR	0.3710	0.1211	0.3122	0.1956	0.0156

Note: KR = Krinsky–Robb method; LCL = latent class conditional logit model; MNL = multinomial logit; na = not applicable.

out-of-sample; leave-one-out cross-validation (LOOCV) testing, where one pseudoindividual is repeatedly left out for estimation and is reserved for testing; k-fold cross validation with $k = 5$, where five groups are created and reserved as hold-out observations for testing; and repeated 5-fold cross validation, where the previous cross validation method is repeated. These cross validation methods are common practice in the machine learning community, but are not traditionally used in choice modeling. In general, and confirming the observations of the Monte Carlo study, posterior probability strategies perform better than those based on prior probabilities, especially for the first case study (which involves three alternatives as opposed to the binary nature of the second case study). Beyond PCP, posterior probabilities also bring more stable predictions, as is revealed in both RMSE and MAE. On the other hand, working with maximum probability assignment (strategies 1 and 3) has slightly higher correct prediction rates compared against drawn classes. Holding other environmental information, it is natural to see that a choice set with larger cardinality leads to a lower likelihood of a certain alternative being chosen. There are no remarkable difference across cross validation methodologies, with all class assignment strategies having similar metric values in testing and training. This latter observation implies that predictions made through all assignment strategies are generally consistent without over-fitting.

When comparing aggregate shares in Tables 8 and 10, there is no evident conclusion about which strategy has generally better market share predictions. Whereas predictions from Drawn MNL assignment have a higher p-value for a χ^2 test of fit in the second case study compared with that of Maximum MNL probability assignment, the first case study shows inverse results. Nonetheless, and matching results from the Monte Carlo study, the use of conditional estimates at the individual level implemented together with the KR method is characterized by a more consistent performance in the two case studies.

Conclusions

In this paper, we have discussed and applied six different class assignment strategies for latent class logit models. Whereas maximum prior and posterior class assignment have been applied in some previous studies, we argue that individuals can also be assigned to a class by randomly drawing a class from a multinomial distribution with probabilities given by MNL probabilities, either prior or posterior. We have also argued and implemented class assignment exploiting individual-level parameter estimates that come from the expected posterior means that are conditional to the sequence of choices made by the individual. Appendix A presents pseudocode of the implementation of the six class assignment strategies under study. By conducting a Monte Carlo study, we have analyzed the behavior of the identified class assignment strategies focusing on preference parameter recovery, choice predictions, and class share inference. In addition, we used two empirical case studies to supplement the results of the simulation study.

Table 9. Case Study 2: Prediction Metrics—Emission Valuation Discrete Choice Experiment

Class assignment strategy	PCP (drawn)	RMSE	MAE	PCP (max)
Fitted values (full sample)				
Maximum MNL assignment probability	0.6338	0.4578	0.3654021	0.6872
Drawn MNL assignment	0.6076	0.5020	0.3918	0.6342
Maximum posterior assignment probability	0.6313	0.4658	0.3674	0.6805
Drawn posterior assignment	0.6313	0.4658	0.3675	0.6804
Conditional individual-specific estimates	0.6035	0.4429	0.3903	0.7026
Conditional individual-specific estimates + KR	0.6102	0.4969	0.3895	0.6421
20% testing (80% training)				
Maximum MNL assignment probability	0.6413	0.4667	0.3633	0.6943
Drawn MNL assignment	0.5994	0.5076	0.3992	0.6284
Maximum posterior assignment probability	0.6094	0.4830	0.3902	0.6537
Drawn posterior assignment	0.6247	0.4824	0.3880	0.6534
Conditional individual-specific estimates	0.6091	0.4502	0.3966	0.6898
Conditional individual-specific estimates + KR	0.6159	0.4935	0.3866	0.6491
LOOCV				
Maximum MNL assignment probability	0.6349	0.4579	0.3652	0.6887
Drawn MNL assignment	0.6095	0.5011	0.3903	0.6362
Maximum posterior assignment probability	0.6235	0.4751	0.3763	0.6639
Drawn posterior assignment	0.6226	0.4766	0.3773	0.6617
Conditional individual-specific estimates	0.6095	0.4429	0.3903	0.7026
Conditional individual-specific estimates + KR	0.6104	0.4950	0.3881	0.6442
k-fold cross validation ($k = 5$)				
Maximum MNL assignment probability	0.6455	0.4591	0.3602	0.6918
Drawn MNL assignment	0.6036	0.5016	0.3922	0.6309
Maximum posterior assignment probability	0.6113	0.4859	0.3889	0.6384
Drawn posterior assignment	0.6045	0.4885	0.3914	0.6332
Conditional individual-specific estimates	0.6139	0.4436	0.3910	0.7017
Conditional individual-specific estimates + KR	0.6065	0.5044	0.3953	0.6309
Repeated k-fold cross validation ($k = 5$)				
Maximum MNL assignment probability	0.6385	0.4585	0.3607	0.6939
Drawn MNL assignment	0.6063	0.5036	0.3935	0.6305
Maximum posterior assignment probability	0.6168	0.4806	0.3842	0.6512
Drawn posterior assignment	0.6167	0.4812	0.3846	0.6502
Conditional individual-specific estimates	0.6097	0.4437	0.3908	0.7018
Conditional individual-specific estimates + KR	0.6101	0.4966	0.3897	0.6400

Note: KR = Krinsky–Robb method; LOOCV = leave-one-out cross-validation; MAE = mean absolute error; max = maximum; MNL = multinomial logit; PCP = percent correctly predicted; RMSE = root mean square error.

The results of the Monte Carlo study have the following implications. Given a moderate number of choice occasions by a consumer (i.e., $T = 5, 10$), the maximum posterior strategy (i.e., strategy 3, assigning a consumer to the class having the highest posterior probability) performs best at parameter recovery. On the one hand, class assignment according to maximum probabilities (strategies 1 and 3) can be seen as optimal strategies in expectation. On the other hand, posterior class assignment probabilities take advantage of the information contained in the sequence of choices made by the individual so that the posterior evaluation is more accurate in reproducing the correct class. However, in the case of a larger number of choice occasions (i.e., $T = 20, 50$), class assignment based on individual-level conditional estimates that account for the sampling distribution of the assignment parameters (through a KR method type of procedure) shows superior behavior, with a very good percentage of correct predictions of the actual classes. This can be explained by the maximum posterior strategy having over-fitting effects because posterior probabilities optimize toward choice likelihood. In contrast, the KR procedure can reproduce extreme cases where an individual makes low-likelihood choices according to their true latent class. In addition, the capacity to collect socio-demographic information will help all strategies do better jobs, especially for strategies 1 and 2 using prior probabilities.

The results of the two empirical case studies, when actual classes are not known, suggest that drawn posterior assignment (strategy 4) performs best from the perspective of aggregate shares. However, maximum probability assignment performs better at predicting individual choices.

Table 10. Case Study 2: Market Shares—Emission Valuation Discrete Choice Experiment

	Aggregate shares		
Class assignment strategy	Alt1	Alt2	p-value
Actual shares (full sample)	0.4676	0.5324	na
LCL fitted shares (full sample)			
Maximum MNL assignment probability	0.5342	0.4657	0.0000
Drawn MNL assignment	0.4664	0.5335	0.7351
Maximum posterior assignment probability	0.5522	0.4477	0.0000
Drawn posterior assignment	0.5515	0.4484	0.0000
Conditional individual-specific estimates	0.4657	0.5342	0.5878
Conditional individual-specific estimates + KR	0.4664	0.5336	0.8081
LCL Fitted shares 20% testing			
Actual shares (testing sample)	0.4628	0.5371	na
Maximum MNL assignment probability	0.5057	0.4942	0.0000
Drawn MNL assignment	0.4627	0.5372	0.9902
Maximum posterior assignment probability	0.5508	0.4491	0.0000
Drawn posterior assignment	0.5421	0.4578	0.0000
Conditional individual-specific estimates	0.4586	0.5413	0.6076
Conditional individual-specific estimates + KR	0.4687	0.5313	0.2484

Note: KR = Krinsky–Robb method; LCL = latent class conditional logit model; MNL = multinomial logit; na = not applicable.

In sum, just plugging in point estimates in the MNL probabilities of class assignment—which is equivalent to prior class assignment and is commonly used in practice—should be avoided. For analyzing expected class shares, individual-level conditional estimates implemented with the KR method should be preferred, although posterior assignment performs almost as well.

Finally, as future avenue of research we would like to explore how the different class assignment strategies behave in the context of the novel latent class logit specification with consumer-surplus feedback from the class-specific conditional logit models to the class assignment MNL model (*34–36*).

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: W. Wu, R. Daziano; data collection: W. Wu, R. Daziano; analysis and interpretation of results: W. Wu, R. Daziano; draft manuscript preparation: W. Wu, R. Daziano. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Technical developments were supported by the National Science Foundation Grant SES-2031841.

ORCID iDs

Wangwei Wu https://orcid.org/0000-0003-4252-8316
Ricardo Daziano https://orcid.org/0000-0002-5613-429X

Supplemental Material

Supplemental material for this article is available online.

References

1. McFadden, D., and K. Train. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, Vol. 15, No. 5, 2000, pp. 447–470.
2. McFadden, D. The Measurement of Urban Travel Demand. *Journal of Public Economics*, Vol. 3, No. 4, 1974, pp. 303–328.
3. Kamakura, W. A., and G. J. Russell. A Probabilistic Choice Model for Market Segmentation and Elasticity Structure. *Journal of Marketing Research*, Vol. 26, 1989, pp. 379–390.
4. DeSarbo, W., V. Ramaswamy, and S. Cohen. Market Segmentation with Choice-Based Conjoint Analysis. *Marketing Letters*, Vol. 6, 1995, pp. 137–147.
5. Greene, W. H., and D. A. Hensher. A Latent Class Model for Discrete Choice Analysis: Contrasts With Mixed Logit. *Transportation Research Part B: Methodological*, Vol. 37, No. 8, 2003, pp. 681–698.
6. Bhat, C. R. An Endogenous Segmentation Mode Choice Model with an Application to Intercity Travel. *Transportation Science*, Vol. 31, No. 1, 1997, pp. 34–48.
7. Wedel, M., W. Kamakura, N. Arora, A. Bemmaor, J. Chiang, T. Elrod, R. Johnson, P. Lenk, S. Neslin, and C. S. Poulsen. Discrete and Continuous Representations of Unobserved Heterogeneity in Choice Modeling. *Marketing Letters*, Vol. 10, No. 3, 1999, pp. 219–232.

8. Hensher, D. A., and W. H. Greene. The Mixed Logit Model: The State of Practice. *Transportation*, Vol. 30, No. 2, 2003, pp. 133–176.
9. Cherchi, E., and J. Polak. Assessing User Benefits with Discrete Choice Models: Implications of Specification Errors Under Random Taste Heterogeneity. *Transportation Research Record: Journal of the Transportation Research Board*, 2005. 1926: 61–69.
10. Fosgerau, M., and S. Hess. A Comparison of Methods for Representing Random Taste Heterogeneity in Discrete Choice Models. *European Transport-Trasporti Europei*, Vol. 42, 2009, pp. 1–25.
11. Dong, X., and F. S. Koppelman. Comparison of Continuous and Discrete Representations of Unobserved Heterogeneity in Logit Models. *Journal of Marketing Analytics*, Vol. 2, No. 1, 2014, pp. 43–58.
12. Frith, M. J. Modelling Taste Heterogeneity Regarding Offence Location Choices. *Journal of Choice Modelling*, Vol. 33, 2019, p. 100187.
13. Lew, D. K. Place of Residence and Cost Attribute Non-Attendance in a Stated Preference Choice Experiment Involving a Marine Endangered Species. *Marine Resource Economics*, Vol. 34, No. 3, 2019, pp. 225–245.
14. Notaro, S., G. Grilli, and A. Paletto. The Role of Emotions on Tourists' Willingness to Pay for the Alpine Landscape: A Latent Class Approach. *Landscape Research*, Vol. 44, No. 6, 2019, pp. 743–756.
15. El Zarwi, F., A. Vij, and J. L. Walker. A Discrete Choice Framework for Modeling and Forecasting the Adoption and Diffusion of New Transportation Services. *Transportation Research Part C: Emerging Technologies*, Vol. 79, 2017, pp. 207–223.
16. Hurtubia, R., M. H. Nguyen, A. Glerum, and M. Bierlaire. Integrating Psychometric Indicators in Latent Class Choice Models. *Transportation Research Part A: Policy and Practice*, Vol. 64, 2014, pp. 135–146.
17. Shen, J. Latent Class Model or Mixed Logit Model? A Comparison by Transport Mode Choice Data. *Applied Economics*, Vol. 41, No. 22, 2009, pp. 2915–2924.
18. Ferguson, M., M. Mohamed, C. D. Higgins, E. Abotalebi, and P. Kanaroglou. How Open are Canadian Households to Electric Vehicles? A National Latent Class Choice Analysis with Willingness-to-Pay and Metropolitan Characterization. *Transportation Research Part D: Transport and Environment*, Vol. 58, No. December 2017, 2018, pp. 208–224.
19. Ho, K. A., M. Acar, A. Puig, G. Hutas, and S. Fifer. What do Australian Patients with Inflammatory Arthritis Value in Treatment? A Discrete Choice Experiment. *Clinical Rheumatology*, Vol. 39, No. 4, 2020, pp. 1077–1089.
20. Rozier, M. D., A. A. Ghaferi, A. Rose, N. J. Simon, N. Birkmeyer, and L. A. Prosser. Patient Preferences for Bariatric Surgery: Findings from a Survey Using Discrete Choice Experiment Methodology. *JAMA Surgery*, Vol. 154, No. 1, 2019, pp. 1–10.
21. Romero-Espinosa, D., M. Sarrias, and R. Daziano. Are Preferences for City Attributes Heterogeneous? An Assessment Using a Discrete Choice Experiment. *Papers in Regional Science*, Vol. 100, No. 1, 2021, pp. 251–272.
22. Sarrias, M., and R. A. Daziano. Individual-Specific Point and Interval Conditional Estimates of Latent Class Logit Parameters. *Journal of Choice Modelling*, Vol. 27, 2018, pp. 50–61.
23. Scarpa, R., and M. Thiene. Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preferences. *Land Economics*, Vol. 81, No. 3, 2005, pp. 426–444.
24. Nunez Velasco, J. P., H. Farah, B. van Arem, and M. P. Hagenzieker. Studying Pedestrians' Crossing Behavior When Interacting with Automated Vehicles Using Virtual Reality. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 66, 2019, pp. 1–14.
25. Elliott, M. R., Z. Zhao, B. Mukherjee, A. Kanaya, and B. L. Needham. Methods to Account for Uncertainty in Latent Class Assignments when Using Latent Classes as Predictors in Regression Models, With Application to Acculturation Strategy Measures. *Epidemiology (Cambridge, Mass.)*, Vol. 31, No. 2, 2020, p. 194.
26. Greene, W., M. N. Harris, and C. Spencer. *Estimating the Standard Errors of Individual-Specific Parameters in Random Parameters Models*. Bankwest Curtin Economics Centre, Curtin University, Bentley WA, 2013.
27. Train, K. *Discrete Choice Methods with Simulation -Introduction*. Cambridge University Press, Cambridge, UK, 2009, pp. 1–8.
28. Krinsky, I., and A. L. Robb. On Approximating the Statistical Properties of Elasticities. *The Review of Economics and Statistics*, Vol. 68, 1986, pp. 715–719.
29. Brier, G. W. Verification of Forecasts Expressed in Terms of Probability. *Monthly Weather Review*, Vol. 78, No. 1, 1950, pp. 1–3.
30. Gneiting, T., and A. E. Raftery. Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, Vol. 102, No. 477, 2007, pp. 359–378.
31. Daziano, R. A., M. Sarrias, and B. Leard. Are Consumers Willing to Pay to Let Cars Drive for Them? Analyzing Response to Autonomous Vehicles. *Transportation Research Part C: Emerging Technologies*, Vol. 78, 2017, pp. 150–164.
32. Leard, B. Consumer Inattention and the Demand for Vehicle Fuel Cost Savings. *Journal of Choice Modelling*, Vol. 29, 2018, pp. 1–16.
33. Daziano, R. A., E. Waygood, Z. Patterson, and M. B Kohlova.. Increasing the Influence of CO2 Emissions Information on Car Purchase. *Journal of Cleaner Production*, Vol. 164, 2017, pp. 861–871.
34. Vij, A., and J. L. Walker. Preference Endogeneity in Discrete Choice Models. *Transportation Research Part B: Methodological*, Vol. 64, 2014, pp. 90–105.
35. Vij, A., S. Gorripaty, and J. L. Walker. From Trend Spotting to Trend'splaining: Understanding Modal Preference Shifts in the San Francisco Bay Area. *Transportation Research Part A: Policy and Practice*, Vol. 95, 2017, pp. 238–258.
36. Hossain, S., M. Hasnine, and K. N. Habib. A Latent Class Joint Mode and Departure Time Choice Model for the Greater Toronto and Hamilton Area. *Transportation*, Vol. 48, No. 3, 2021, pp. 1217–1239.