Do Random Coefficients and Alternative Specific Constants Improve Policy Analysis? An Empirical Investigation of Model Fit and Prediction*

H. Allen Klaiber
Department of Agricultural, Environmental and Development Economics
The Ohio State University

Roger H. von Haefen
Department of Agricultural and Resource Economics
North Carolina State University

This Draft: November 3, 2011

Abstract

In recent years, there have been numerous econometric innovations in non-market valuation applications of discrete choice models. Several particularly attractive and frequently used innovations are the inclusion of random parameters and inclusion of alternative specific constants. While these improve statistical fit, they often result in poor in sample prediction of chosen alternatives. In an effort to improve model fit and relax restrictive substitution patterns, model prediction often takes on secondary importance. In this paper, we examine the impacts of these methods on model fit and prediction, arguing that attention to poor prediction is equally important to generate credible inference. Given the apparent tradeoff between fit and prediction, we then examine a series of "second-best" modeling strategies that attempt to correct for the poor prediction we observe in several recreation data sets and find that there are a variety of ways in which to successfully overcome poor prediction.

Keywords: Discrete Choice, Recreation Demand, Revealed Preference, Stated Preference, Latent Class, Alternative Specific Constants, Random Parameters

---

*Very Early Draft – Comments welcome, but do not cite.

Do Random Coefficients and Alternative Specific Constants Improve Policy Analysis? An Empirical Investigation of Model Fit and Prediction

I. Introduction

Discrete choice logit models have become one of the most frequently used modeling frameworks for recreation demand and locational equilibrium models (Murdock, 2006; Bayer and Timmins, 2007). Within the framework, two econometric innovations that applied researchers are using with increasing regularity are random coefficients (McFadden and Train, 2000) and the inclusion of alternative specific constants (Berry, 1994). Random coefficients are an attractive mechanism for relaxing the restrictive implications of the independence of irrelevant alternatives (IIA) thus introducing more plausible substitution patterns. Including a full set of alternative specific constants allows the analyst to control for unobserved attributes that may be correlated with observed covariates.

In applications of these modeling innovations to discrete choice models, researchers have found that they generate substantial and statistically significant improvements in fit through increases in likelihood and improvement in information criteria (von Haefen and Phaneuf, 2007; Murdock, 2006). In an empirical investigation of four recreation data sets, we also find large gains in model fit. However, we find that the introduction of random parameters often results in a failure to replicate the in sample aggregate choice patterns implied by the data. This empirical regularity generates important implications for the credibility of policy analysis – why should one believe welfare measures derived from models that cannot replicate in sample aggregate choice behavior?

Our goal in this paper is to shed light on the counterintuitive empirical regularity of improved statistical fit combined with poor in sample prediction. We document this phenomenon using four recreation data sets that have been the focus of published research by previous researchers. Two of the four applications combine revealed and stated preference (RP/SP) to identify all demand parameters (Adamowicz et al., 1997; Haener et al., 2001) as previously done by von Haefen and Phaneuf (2007). The others exploit only revealed preference (RP) data (Parsons et al., 1999; von Haefen, 2003) and use a variation of a two-step estimator proposed by Berry, Levinsohn, and Pakes (2004) and used recently by Murdock (2006) in the recreation context. With all four data sets, we find the introduction of random coefficients and alternative specific constants (ASCs hereafter) substantially and significantly improves statistical fit. We also find that in sample trip predictions often (but not uniformly) deteriorate with these richer empirical specifications, particularly with the introduction of random parameters.

To develop the intuition for why these poor predictions arise in practice, we use theoretical results from Gourieroux, Monfort, and Trognon (1984) about a property of the linear exponential family of distributions as well as some Monte Carlo findings. The upshot of our discussion is that: 1) fixed coefficient logit models with a full set of ASCs will generate in sample trip predictions for each alternative that *perfectly* match the data, and 2) random coefficient models with ASCs may not predict perfectly in sample, but should be reasonably close if the analyst has correctly specified the underlying data generating process. An implication of this finding is that the poor in sample predictions that we find in our empirical applications arise because of model misspecification. This finding is a cautionary lesson to researchers as this indicates that despite adoption of recent econometric innovations, these models nevertheless may fail to account for important features of the data that are masked by focusing only on statistical fit.

We conclude by exploring a number of “second best” strategies for dealing with poor in sample predictions. These range from: 1) abandoning random coefficient specifications and using fixed coefficient models with ASCs that generate perfect in sample predictions; 2) following Murdock (2006) and using the Berry (1994) contraction mapping with ASCs to force the in sample predictions to match the data perfectly; 3) using less-efficient non-panel random coefficient models that, as we demonstrate, generate more plausible in sample predictions; and 4) following von Haefen (2003) and conditioning on observed choice in the construction of welfare measures. Our results suggest that each of these strategies is effective in terms of generating plausible in sample predictions but they differ considerably in terms of their implications for statistical fit.

The paper proceeds as follows. The next section documents the performance of random coefficient and fixed coefficient logit models with a full set of ASCs using four previously used recreation data sets. Section III explores the factors that give rise to the perverse empirical findings reported in the previous section using econometric theory and a set of Monte Carlo simulations. Section IV investigates a number of “second best” empirical strategies that applied researchers may find attractive in future applications. We then conclude with some final observations and recommendations.

II. Nature of the Problem

We begin by illustrating the poor in sample prediction problem that serves as the motivation for this research. To demonstrate that this problem is not an idiosyncratic feature associated with a single data set, we consider four recreation data sets that researchers have used in previously published studies with two combining RP and SP data and two using RP data alone. As discussed in von Haefen and Phaneuf (2007), the fusion of RP and SP data is attractive in both data environments because the inclusion of a full set of ASCs confounds identification of the site attribute parameters given the relatively small number of sites in each application. For our two RP studies, we use a two-step estimator originally outlined by Berry, Levinsohn, and Pakes (1995) and applied to a recreation model by Murdoch (2006). This estimator recovers (J-1) alternative specific constants[1] and all individual varying covariates in the first stage of estimation, such as travel cost. Choice specific covariates are recovered in a second stage regression of the estimated ASCs on choice specific covariates and a constant term.

The first RP/SP data set was used by Adamowicz et al. (1997) and consists of both revealed preference (RP) and stated preference (SP) choice data for moose hunting in the Canadian province of Alberta. The RP data consists of seasonal moose hunting trips for 271 individuals to 14 wildlife management units (WMUs) throughout Alberta in 1993. The SP data consists of 16 choice experiments that were generated with a blocked orthogonal, main effects design. All eleven site attributes except travel cost in the RP and SP data are effects coded and interacted with three demographic variables. The second RP/SP data set was first used by Haener et al. (2001) and also consists of combined RP/SP data for Canadian moose hunting. This data source, however, was collected in the neighboring province of Saskatchewan in 1994. The RP data consists of seasonal moose hunting trips for 532 individuals to 11 wildlife management zones (WMZs) throughout Saskatchewan. The SP data consists of 16 choice experiments that were generated with a blocked orthogonal, main effects design. All nine

[1] Only differences in utility enter into the logit model precluding estimation of the full set of ASCs. The normalized ASC is captured by a constant in the second stage of estimation.

attributes except travel cost in the RP and SP data are effects coded and interacted with three demographic variables.

The first RP data set we consider looks at Mid-Atlantic beach visitation and was first used by Parsons et al. (1999). This data set consists of seasonal trip data to 62 ocean beaches in 1997 for 375 individuals. For each beach, we observe 14 site characteristic variables plus a constructed individual specific travel cost variable based on each recreator's home zip code. Our second RP data set focuses on recreation trips to the Susquehanna River basin for 157 nearby residents who take a combined total of 2,471 trips to one of 89 recreation sites. This data set was first used by von Haefen (2003) and includes a travel cost measure to each individual recreation site in addition to site characteristics.

Table 1 summarizes our initial findings on model fit and prediction. This table contains estimation results from 4 models associated with each of our empirical applications. These models include both fixed and random parameter models excluding ASCs and the same models including ASCs. For the results reported in Table 1, our two-step estimator differs from previous two-step estimators in the following way. Similar to Murdock (2006), we use maximum likelihood techniques in the first step logit estimation to recover the travel cost parameter and a full set of ASCs that subsume all site characteristics that do not vary over individuals.[2] In contrast to Murdock, our first step estimator does not employ the Berry (1994) contraction mapping algorithm, an issue we return to in a later section. Thus our first step estimator relies entirely on traditional maximum likelihood techniques, not the combination of maximum likelihood and Berry contraction mapping techniques that Murdock employs. Our second stage estimator is identical to Murdock's approach in that we regress the estimated ASCs from the first

[2] We do not include any demographic interactions in this model because preliminary testing suggested that they did not improve model fit

stage on the site characteristics and a constant term. Importantly, this approach assumes that the unobserved site attributes are uncorrelated with observed site attributes.

A comparison of log-likelihood values between fixed and random parameter specifications for all model combinations shows that inclusion of random parameters improves statistical fit significantly. Comparing across columns shows that inclusion of ASCs also results in large increases in likelihood values. Overall, the highest log-likelihood values are associated with models including both ASCs and random parameters. This general result holds regardless of application and is consistent with the general enthusiasm in the literature for incorporating these new techniques into empirical models.

To ascertain how well these models predict aggregate trip taking behavior for each alternative, we construct the following summary statistic for each model:

(1) Percentage absolute prediction error = $100 \times \sum_{i=1}^{J} s_i^S \frac{\text{abs}(s_i^S - s_i^M)}{s_i^S} = 100 \times \sum_{i=1}^{J} \text{abs}(s_i^S - s_i^M)$,

where s_i^S and s_i^M are the in sample share of trips to alternative *i* and the model's prediction of the share of trips to alternative *i*, respectively, and *J* is the number of alternatives. This prediction error statistic can be interpreted as the share weighted in sample prediction error for each alternative and thus can be used to rank order the models in terms of in sample predictions that match the observed data. Intuitively, a model that can replicate aggregate trip predictions well for each alternative would generate a low prediction error value, whereas a model with poor in sample aggregate predictions for each alternative would score a relatively high value.

Examining the columns for absolute prediction error both with and without ASCs reveals an import result that has generated much enthusiasm among researchers. For models employing fixed coefficients and ASCs, we see that the prediction errors are zero. This perfect prediction is a result of properties of the linear exponential family which we derive in the following section. The finding of perfect prediction has led to the terminology "locational equilibrium" where these models exactly replicate the observed pattern of location choice among individuals. This non-price equilibrium has recently been adopted in a recent environmental applications focusing on housing (Bayer et al., 2009; Klaiber and Phaneuf, 2010; Tra, 2010). What is troubling in our results reported above is that this pattern of perfect prediction disappears when using random parameters specifications, despite their superior model fit. In the case of the Susquehanna example, the fit deteriorates by nearly 50% with the introduction of random parameters

Finally, it is interesting to see how these differences in fit and prediction play across datasets. The best predications are consistently associated with the Mid-Atlantic Beach dataset while the worst predications are associated with the Susquehanna dataset. As the Mid-Atlantic Beach dataset contains significantly more covariates, this finding suggests that it is more fully specified and therefore able to capture the site visitation patterns more accurately. In fact, Monte Carlo estimates where the true model is known result in perfect prediction across all model specifications.[3] This result suggests that the source of the prediction error is directly associated with model misspecification in our empirical settings

To assess the robustness of these empirical findings, we repeated the exercise using seasonal models, as those are able to capture the repeated and frequent visitation patterns observed in our empirical datasets. Results from these models are presented in Table 2. These results qualitatively mirror those reported in Table 1, however several interesting quantitative differences emerge. While the prediction errors can still be quite large, approaching 50% in one case, the general magnitude of these errors is reduced. This is most apparent for the Saskatchewan RP/SP data where errors for the panel random coefficient models drop to under

[3] These simulations are available upon request.

5% for models both with and without ASCs. This is in contrast to the errors of nearly 60% and 40% associated with the corresponding trip allocation models. This finding suggests that the seasonal models better capture the underlying data generating process, and are also consistent with our Monte Carlo findings that a correctly specified model does not exhibit the poor prediction phenomenon we have characterized.

Our final set of empirical models considers latent class models employing panel random parameters with the results are reported in Table 3. These models maintain the general pattern of poor prediction, however, it appears that the ability of these models to better capture latent heterogeneity improve prediction over all of the previous random parameters models. In particular, the best performing models appear to be seasonal models with ASCs included. While prediction errors are smaller than in our previous sets of results, they are still quite large for the Susquehanna RP dataset. In all cases, inclusion of ASCs improves model prediction over models which do not include a full set of ASCs; however, in no case do we observe perfect prediction.

In summary, the results in Table 1 through Table 3 suggest a somewhat counterintuitive result – including ASCs and especially random coefficients significantly improve overall statistical fit but do not necessarily generate in sample trip predictions that match the observed data well. For many policy oriented applications, the failure to predict in sample choice patterns accurately raises serious concerns. Richer model specifications including seasonal models and latent class models appear to improve prediction; however, they do not result in the accuracy of prediction obtained from the much simpler fixed coefficient model with a full set of alternative specific constants.

III. What explains these counterintuitive results?

In this section we use econometric theory to shed light on the counterintuitive results presented in the previous section. To motivate our main insight here, consider the log-likelihood function for a sample of *N* individuals each making separate choices from *J* alternatives:

$$(2) \qquad \ln L(\beta) = \sum_{i=1}^{N}\sum_{j=1}^{J} 1_{ij} X_{ij}\beta - \ln\left(\sum_{k=1}^{J}\exp(X_{ik}\beta)\right),$$

where 1_{ij} is an indicator function equal to 1 for individual *i's* chosen alternative and zero otherwise. The score condition associated with this log-likelihood is:

$$(3) \qquad \frac{\partial \ln L(\beta)}{\partial \beta} = \sum_{i=1}^{N}\sum_{j=1}^{J} X_{ij}\left[1_{ij} - \Pr_i(j \mid \beta)\right] = 0\,,$$

where $\Pr_i(j \mid \beta)$ is the logit probability for individual *i* choosing the *j*th alternative. If a full set of ASCs are included, then

$$(4) \qquad X_{ij} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}, \forall k\,,$$

and the score conditions associated with the ASCs can be written:

$$(5) \qquad \sum_{i=1}^{N}\left[1_{ik} - \Pr_i(k \mid \beta)\right] = 0 \quad \text{or} \quad \frac{1}{N}\sum_{i=1}^{N} 1_{ik} = \frac{1}{N}\sum_{i=1}^{N}\Pr_i(k \mid \beta), \forall k.$$

Equation 5 implies that fixed coefficient logit models with a full set of ASCs will generate in sample predictions that match the data perfectly, a result that is consistent with our empirical findings in Table 1 and well known in the discrete choice literature (see, e.g., Ben-Akiva and Lerman, 1985).

As Gourieroux, Monfort, and Trognon (1984) have shown, the logit distribution falls within the broad class of distributions known as the linear exponential family of distributions. Other notable examples include the Poisson and normal distributions. What defines this family of distributions is that they are all mean-fitting distributions, implying that with the inclusion of

ASCs, predictions from these distributions will match the data perfectly. A notable advantage of using linear exponential distributions in empirical work is that if the analyst has correctly specified the conditional expectation function of the distribution (i.e., its first moment), high order misspecification will not lead to inconsistent parameter estimates (it will, however, bias standard error estimates, but this problem can be addressed if the analyst uses robust standard errors (White, 1981) instead of traditional standard errors). Thus, if the analyst specifies the first moment correctly, consistent parameter estimates will result. This makes the fixed coefficient logit model with ASCs appealing.

It is important to note, however, that adding random coefficients to the logit distribution result in a mixture distribution that falls outside the linear exponential family.[4] Random coefficient logit models, regardless of whether ASCs are included, will not necessarily generate in sample predictions that match the data perfectly. This can be seen by looking at the score conditions for the simulated random coefficient logit model. The simulated likelihood function in this case is:

(6) $$L(\bar{\beta},\sigma)=\sum_{i=1}^{N}\sum_{j=1}^{J}1_{ij}\ln\left(\frac{1}{R}\sum_{r=1}^{R}\Pr_i(j\mid\beta_i^r)\right)=\sum_{i=1}^{N}\sum_{j=1}^{J}1_{ij}\ln\left(\frac{1}{R}\sum_{r=1}^{R}\frac{\exp(X_{ij}\beta_i^r)}{\sum_{k=1}^{J}\exp(X_{ik}\beta_i^r)}\right),$$

where $\beta_i^r=\bar{\beta}+\sigma U_i^r$, $U_i^r\sim N(0,1)$, and the score condition is:

(7) $$\frac{\partial L(\bar{\beta},\sigma)}{\partial\beta_i^r}=\prod_{i=1}^{N}\left[\frac{\prod_{j=1}^{J}\left[X_{ij}\frac{1}{R}\sum_{r=1}^{R}\Pr_i(j\mid\beta_i^r)(1-\Pr_i(j\mid\beta_i^r))\right]^{1_{ij}}}{\prod_{j=1}^{J}\left[\frac{1}{R}\sum_{r=1}^{R}\Pr_i(j\mid\beta_i^r)\right]^{1_{ij}}}\right]=0\,.$$

[4] This result also holds for latent class models.

With the inclusion of ASCs, this condition does not imply perfect in sample predictions. Thus, some degree of imperfect in sample prediction can be expected from random coefficient logit models, but the precise degree will vary across applications.

To assess how well in sample predictions from estimated logit models will match the data, we conducted an extensive Monte Carlo analysis where we know the underlying data generating process for the simulated data. Knowing the true data generating process allowed us to ascertain the in sample prediction performance of maximum likelihood estimators when model misspecification is absent. If the in sample predictions generated from these correctly specified models matches the observed data well, then we can conclude that poor in sample predictions arise due to some form of model specification, and not due to an inherent property of the estimator.

For brevity, we only summarize the main conclusions of our Monte Carlo simulation here. Across a number of specifications, we consistently found that the in sample predictions for panel and non-panel random coefficient models with and without alternative specific constants matched the simulated data nearly identically. Under none of our simulations did we find the degree of poor in sample prediction that we observed with the empirical data sets reported in Table 1. Based on these findings, we conclude that the poor predictions found in our three applications are a result of model misspecification.

The implications of the above discussion for how analysts should proceed are unclear. If the analyst estimates logit models with random coefficients and finds poor in sample predictions, the obvious ‘first best’ solution would be to continue to search for empirical specifications that fit the data well and predict well in sample. In practice, however, finding empirical specifications that satisfy these two criteria will be computationally difficult, time-consuming, and in many cases infeasible.

IV. Improving Prediction in the Presence of Misspecification

In empirical settings, obtaining the "true" model specification is unlikely, suggesting that 'second best' less demanding approaches that address the concerns of both prediction and model flexibility are warranted. Perhaps the simplest second best approach would be to estimate a fixed coefficient logit model with ASCs where the in sample aggregate predictions will match the data perfectly. One problem with this approach is that it employs models with substitution patterns that are consistent with the independence of irrelevant alternatives (IAA). These restrictive substitution patterns can be partially relaxed by using nested logit models, but the considerably more flexible substitution patterns that come with random coefficient models will not be realized.

Another second best approach involves estimating random coefficient models with ASCs using a contraction mapping (Berry, 1994) that iteratively solves for the ASC values by matching the aggregate model predictions with the data. This algorithm was first used in the industrial organization literature to estimate discrete choice models of product choice using aggregate market share data (Berry, Levinsohn, and Pakes, 1995), but Berry, Levinsohn, and Pakes (2004) apply the algorithm to a disaggregate data context. Both of these applications employed generalized method of moments estimation techniques, and it was not until Murdock (2006) that the algorithm was used within a maximum likelihood estimation framework with random coefficients.

Despite the ability to introduce heterogeneity through random parameters and achieve perfect prediction, as demonstrated by equation (7), this approach does not result in estimates consistent with maximum likelihood and creates problems for inference based on those estimates. The decision to eschew maximum likelihood estimates to achieve perfect prediction with improved model flexibility using the contraction mapping technique is similar in-spirit to the penalized likelihood estimates that Shonkwiler and Englin (2005) and von Haefen and Phaneuf (2003) have previously used.

The idea behind maximum penalized likelihood estimation is that one maximizes the likelihood subject to a function that penalizes the likelihood for some undesirable behavior, in this case poor prediction. Random coefficient logit models with ASCs that are estimated within the maximum likelihood framework using the Berry contraction mapping are observationally equivalent to estimating random coefficient logit models with ASCs within the maximum penalized likelihood framework using an infinitely weighted penalty function for poor in sample predictions. A limitation with this approach is that the asymptotic properties of maximum penalized likelihood estimators are not well understood, but it does directly address the poor in sample prediction problem.

A third approach for dealing with poor in sample predictions involves estimating non-panel random coefficient models with ASCs within the maximum likelihood framework. This approach sacrifices the efficiency gains (which may be substantial) from introducing correlations across an individual's multiple trips for improved (but not perfect) in sample predictions. Moreover, it makes estimation more computationally intensive. Our final second-best strategy to improve prediction is the incorporation of observed choice into the construction of welfare measures as suggested by von Haefen (2003). The idea of incorporating observed choice into welfare measurement construction is attractive because it simulates the unobserved determinants of choice in a way that implies perfect prediction for every observation and then uses the model's implied structure of substitution to ascertain how behavior and welfare change with changes in price, quality, and income. The approach can be used with any set of model estimates, but it requires a somewhat more computationally intensive algorithm for calculating welfare estimates (see von Haefen (2003) for details) and does not directly address the underlying source of model misspecification.

Table 4 provides estimation results for each of our empirical applications using these 4 approaches with ASCs included in the model specification. As a baseline for comparison, the first row for each application presents the fixed coefficient model with ASCs that consistently exhibits the worst fit among all models. The second, third, and fourth rows for each application present non-panel random parameters, maximum penalized likelihood estimates similar to those used by Murdoch (2006), and conditional estimates following von Haefen (2003). As can be seen, each of these is associated with large improvements in prediction error with errors typically less than 2% compared to well over 30% as reported in Table 1.

The results for the conditional models are identical to the panel random coefficient estimates including ASCs that are reported in table 1, where prediction is improved through conditioning on observed choice only for welfare measures. This model does not correct for poor-prediction in terms of the estimated parameters (which are not shown). The slightly less than perfect prediction associated with maximum penalized likelihood arises from the empirical estimates having converged to numerical stopping criteria prior to reaching perfect prediction. Comparing the conditional and penalized likelihoods one can see that the likelihood values for penalized models are inferior to the maximum likelihood values, reflecting the results shown in section III. The non-panel random parameters models significantly improve prediction, but give up the majority of the gains achieved through panel random parameters.

V. Conclusion

Our goal in this research has been threefold: 1) to document the somewhat counterintuitive in sample prediction problems that arise with random coefficient logit models that include ASCs; 2) to explore the sources of these problems using economic theory; and 3) to suggest and evaluate alternative, second best, strategies for dealing with the poor in sample predictions that researchers might find attractive in future empirical work. Across four data sets, we document that the addition of ASCs and especially panel random coefficients generates significant improvements in statistical fit but do not uniformly improve model prediction. We argue that failure to predict in sample choice patterns accurately raises validity concerns for policy analysis.

Building on economic theory, we show that the fixed coefficient logit model falls within the larger family of linear exponential distributions, and thus the inclusion of a full set of ASCs will generate in sample trip predictions for each site that match the data perfectly. The introduction of random coefficients, however, results in a mixture distribution that falls outside the linear exponential family and thus will not imply perfect in sample predictions. Results from an extensive Monte Carlo analysis suggest that the poor in sample predictions observed in our four applications are likely due to some form of misspecification.

To account for these model shortcomings, the analyst may find attractive one of the second best strategies that we empirically evaluate for addressing poor in sample predictions. Our empirical results suggest that all of these strategies are effective in controlling for poor in sample predictions, but the use of non-panel random coefficients significantly degrades model fit. Our investigation of alternative discrete choice models found that while both seasonal and latent class models improved model fit relative to non-panel random coefficients, they nevertheless failed to fully capture the in-sample predictions of site choice. The use of either a penalized likelihood estimator or the Berry contraction mapping would generate improved prediction, but at the cost of obtaining maximum likelihood estimates with known asymptotic properties.

Finally, it is worth stepping back and directly addressing the fundamental question that motivated this research: do random coefficients and alternative specific constants improve analysis? With regard to random coefficients, we believe that the richer substitution patterns implied by random coefficients are quite attractive, but the poor in sample predictions that often result from these models (especially panel random coefficient versions) need to be addressed in some way. If not, policy implications and inference is suspect. With regard to alternative specific constants, we believe that their ability to control for unobserved attributes makes them extremely attractive. One limitation with their inclusion, however, is that one needs either an RP data set with many objects of choice (sites in recreation models, or neighborhoods in locational equilibrium models) or additional SP data to identify the part worths of the different site attributes. When these data are available, we believe that ASCs are an attractive modeling innovation and should be routinely included in discrete choice applications.

References

Adamowicz, W., J. Swait, P. Boxall, J. Louviere, and M. Williams, "Perceptions vs. objective measures of environmental quality in combined revealed and stated preference models of environmental valuation," *Journal of Environmental Economics and Management* 32 (1997): 65-84.

Bayer, P. and C. Timmins, "Estimating equilibrium models of sorting across locations," *Economic Journal* 117(2007): 353-374.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management,* 58(1).

Ben-Akiva, M., and S. Lerman. *Discrete Choice Analysis*. MIT Press: Cambridge, 1985.

Berry, S. "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics,* 25(2): 242-262.

Berry, S., J. Levinsohn, and A. Pakes, "Automobile prices in equilibrium," *Econometrica* 63 (1995): 841-890.

Berry, S., J. Levinsohn, and A. Pakes, "Differentiated products demand systems from a combination of micro and macro data: the new car market," *Journal of Political Economy* 112 (2004): 68-105.

Gourieroux, C., A. Monfort, and A. Trognon, "Pseudo-maximum likelihood methods: application to Poisson models," *Econometrica* 52(1984), 701-720.

Haener, M., P. Boxall, and W. Adamowicz, "Modeling recreation site choice: Do hypothetical choices reflect actual behavior?" *American Journal of Agricultural Economics* 83 (2001): 629-642.

Klaiber, H. Allen and Daniel J. Phaneuf. 2010. “Valuing Open Space in a Residential Sorting Model of the Twin Cities.” *Journal of Environmental Economics and Management*, 60(2): 57-77.

McFadden, D. and K. Train, “Mixed MNL of discrete response,” *Journal if Applied Econometrics,* 15(2000): 447-470.

Murdock, J., “Handling unobserved site characteristics in random utility models of recreation demand,” *Journal of Environmental Economics and Management* 51 (2006): 1-25.

Parsons, G., D.M. Massey, and T. Tomasi, “Familiar and favorite sites in a random utility model of beach recreation,” *Marine Resource Economics,* 14(1999): 299-315.

Shonkwiler, J.S., and J. Englin, “Welfare losses due to livestock grazing on public lands: a count data systemwide treatment,” *American Journal of Agricultural Economics* 87(2005): 302-313.

Timmins, C. and J. Murdock, “A revealed preference approach to the measurement of congestion in travel cost models,” *Journal of Environmental Economics and Management* forthcoming.

Train, K., *Discrete Choice with Simulation.* Cambridge University Press: Cambridge, 2003.

Train, K., “Recreation demand models with taste variation over people,” *Land Economics,* 74:2(1998); 230-239.

Tra, Constant I., 2010. “A Discrete Choice Equilibrium Approach to Valuing Large Environmental Changes.” *Journal of Public Economics*, 94 (1-2): 183-196.

von Haefen, R., “Incorporating observed choice into the construction of welfare measures from random utility models,” *Journal of Environmental Economics and Management* 45:2(2003): 145-165.

von Haefen, R., D. Phaneuf, "Identifying demand parameters in the presence of unobservables: a combined revealed and stated preference approach," *Journal of Environmental Economics and Management* forthcoming.

Table 1. Baseline Performance using ASCs and Panel Random Coefficients for Trip Allocation Models

		Alternative Specific Constants							
		NO				YES			
Empirical Example	Specification	Log-Likelihood	Abs Pred Err	# Observations	# Parameters	Log-Likelihood	Abs Pred Err	# Observations	# Parameters
Mid-Atlantic Beach RP	Fixed	-13,160	13.3%	375	15	-12,982	0.0%	375	44
	Random	-11,014	21.7%	375	29	-10,821	12.7%	375	72
Alberta RP/SP	Fixed	-5,655	29.9%	271	50	-5,377	0.0%	271	63
	Random	-4,823	33.9%	271	62	-4,537	19.6%	271	75
Susquehanna RP	Fixed	-5,303	61.5%	157	6	-4,385	0.0%	157	89
	Random	-4,381	86.4%	157	11	-3,604	48.9%	157	94
Saskatchewan RP/SP	Fixed	-7,655	26.3%	532	42	-7,482	0.0%	532	52
	Random	-6,673	59.7%	532	52	-6,567	35.2%	532	62

Table 2. Baseline Performance using ASCs and Panel Random Coefficients for Seasonal Models

		Alternative Specific Constants							
		NO				YES			
Empirical Example	Specification	Log-Likelihood	Abs Pred Err	# Observations	# Parameters	Log-Likelihood	Abs Pred Err	# Observations	# Parameters
Mid-Atlantic Beach RP	Fixed	-27,759	1.9%	540	22	-27,564	0.0%	540	51
	Random	-22,592	19.3%	540	37	-22,545	14.5%	540	81
Alberta RP/SP	Fixed	-8,890	3.2%	271	54	-8,683	0.0%	271	67
	Random	-7,792	10.5%	271	67	-7,549	15.4%	271	80
Susquehanna RP	Fixed	-11,599	9.3%	157	10	-10,743	0.0%	157	93
	Random	-9,392	40.6%	157	16	-8,526	49.0%	157	105
Saskatchewan RP/SP	Fixed	-10,565	1.1%	532	46	-10,172	0.0%	532	56
	Random	-9,258	4.8%	532	57	-9,194	3.0%	532	67

Table 3. Latent Class Panel Random Parameter Models

		Alternative Specific Constants									
		NO					YES				
Empirical Example	Specification	Log-Likelihood	Abs Pred Err	# Observations	# Parameters	Latent Classes	Log-Likelihood	Abs Pred Err	# Observations	# Parameters	Latent Classes
Mid-Atlantic Beach RP	Trip	-11,786	13.3%	375	99	4	-12,312	6.6%	375	64	2
	Seasonal	-24,443	1.9%	540	66	3	-24,733	1.6%	540	85	3
Alberta RP/SP	Trip	-4,865	31.0%	271	113	7	-4,678	23.0%	271	110	6
	Seasonal	-8,041	4.4%	271	106	6	-8,096	1.3%	271	68	3
Susquehanna RP	Trip	-4,370	66.6%	157	42	5	-3,561	33.3%	157	125	5
	Seasonal	-10,036	10.3%	157	27	3	-9,276	12.9%	157	126	5
Saskatchewan RP/SP	Trip	-6,662	47.4%	532	113	8	-6,627	29.5%	532	95	6
	Seasonal	-9,475	2.7%	532	64	4	-9,504	0.8%	532	59	3

Table 4. Potential Solutions to Poor Prediction (all with ASCs included)

Empirical Example	Specification	Log-Likelihood	Abs Pred Err	# Observations	# Parameters
Mid-Atlantic Beach RP	Fixed	-12,982	0.0%	375	44
	Non-Panel	-12,838	2.2%	375	72
	Penalty	-10,911	1.0%	375	72
	Conditional	-10,821	0.0%	375	72
Alberta RP/SP	Fixed	-5,377	0.0%	271	63
	Non-Panel	-5,368	0.1%	271	75
	Penalty	-4,572	0.8%	271	75
	Conditional	-4,537	0.0%	271	75
Susquehanna RP	Fixed	-4,385	0.0%	157	89
	Non-Panel	-4,384	0.9%	157	94
	Penalty	-3,747	3.6%	157	94
	Conditional	-3,604	0.0%	157	94
Saskatchewan RP/SP	Fixed	-7,482	0.0%	532	52
	Non-Panel	-7,467	5.1%	532	62
	Penalty	-6,598	1.9%	532	62
	Conditional	-6,567	0.0%	532	62