11

METHODOLOGICAL DEVELOPMENTS IN ACTIVITY-TRAVEL BEHAVIOR ANALYSIS

David Brownstone
University of California, Irvine, USA

1. INTRODUCTION

Ten years ago I wrote a resource paper on "Discrete Choice Modeling for Transportation" for the 9th IATBR conference on the Gold Coast of Australia (Brownstone, 2001). While the scope of the resource paper for this conference is broader, it is still interesting to review the claims and recommendations made ten years ago to see how they have been received in our profession. The purpose of this paper is to stimulate discussion about methodological developments in travel behavior analysis. I will therefore concentrate on topics that I find relevant instead of trying to survey the large literature in this area.

Although there have been many advances in methodology over the last decade, I will argue that we are not paying enough attention to numerical and inference problems with maximum simulated likelihood estimation. We should also concentrate more efforts on getting valid inferences for forecasts and willingness-to-pay measures (WTP), and these inferences should account for measurement error and model uncertainty. All of these are best handled using Bayesian models and estimation techniques.

A fairly close parody of what most empirical researchers (including frequently me) do is to fit the most sophisticated and complicated model we can subject to the constraint that key parameters have the expected sign and have t-statistics at least two in absolute magnitude. As we work with larger and more accurate data, this methodology will lead to larger models with more parameters. This approach is reinforced since this type of work is relatively easy to publish in high-prestige methodological journals. Unfortunately almost all of the models we used are parameterized so that quantities we frequently are interested in (e.g., WTP, elasticities, scenario forecasts) are complex nonlinear functions of many parameters and data. There is no guarantee that more complex models will yield more accurate estimates of these interesting quantities. For example, Greene and Hensher (2010) find no statistically significant differences between multinomial logit model elasticity estimates and those derived from a variety of increasingly complex models of mode choice in a Sydney data set. Brownstone and Small (2005) similarly find no differences in WTP estimates from a variety of high-occupancy toll lane choice models.

We should therefore concentrate on models that produce accurate forecasts of interesting quantities. Ideally the quality of these forecasts should be checked using independent data, but in practice this usually means a “holdout sample” of the data which is set aside and not used for model estimation (see e.g., Brownstone and Fang, 2009). Since there is usually uncertainty about the correct model specification, these forecasts should be done for a range of reasonable models. Bayesian model averaging (described in the third section of this paper) provides an excellent methodology for incorporating model specification uncertainty into confidence regions for model forecasts.

2. FLEXIBLE DISCRETE CHOICE MODELS

Ten years ago I said that mixed logit models were the wave of the future, but that there needed to be improvements in software and computation to fully realize their promise. During the next decade there have been considerable improvements in software. ALOGIT, BIOGEME, NLOGIT, and even STATA all have mixed logit code, and most of these use improved “low-dispersion” to speed up calculations. Train’s 2003 (now available in the second edition (Train, 2009)) book also helped popularize mixed logit, but Hensher and Greene (2003) cautioned that there are many pitfalls not fully appreciated by researchers applying mixed logit. Hensher and Greene point out that there are many choices to be made when specifying mixed logit models, and those seemingly innocuous choices about distributions of random components can lead to very unrealistic distributions of derived quantities like willingness to pay measures. Hensher and Greene also point out that our typical data do not clearly identify the complex heterogeneity patterns commonly specified by mixed logit users.

The most general parameterization of mixed logit (from Greene and Hensher, 2010) is given by:

$$
\begin{aligned}
P_{i,j,t} &= \frac{exp\left(V_{itj}\right)}{\sum_{j=1}^{J} exp\left(V_{itj}\right)} \\
V_{itj} &= \boldsymbol{\beta}'_i \boldsymbol{x_{itj}} \\
\boldsymbol{\beta}_i &= \sigma_i[\boldsymbol{\beta} + \Delta \boldsymbol{z_i}] + [\gamma + \sigma_i(1-\gamma)]\boldsymbol{\Gamma}\boldsymbol{v_i} \\
\sigma_i &= exp[\bar{\sigma} + \boldsymbol{\delta}'\boldsymbol{h_i} + \tau w_i]
\end{aligned}
\qquad (1)
$$

Note that the “hyperparameters” to be estimated are: $\boldsymbol{\beta}, \Delta, \gamma, \boldsymbol{\Gamma}, \bar{\sigma}, \boldsymbol{\delta}, \tau$. γ controls the relative importance of scale heterogeneity, and $\boldsymbol{\delta}=\tau=0$ implies the standard random parameters form of the mixed logit model. If in addition $\boldsymbol{\Delta}=\boldsymbol{\Gamma}=0$, then the model collapses to the conditional multinomial logit model. Of course some restrictions are required to identify the general structure given above.

One fundamental problem is that mixed logit models are typically estimated by maximum simulated likelihood techniques, and the noise introduced by the simulations can mask identification problems and allow seemingly sensible estimates to be calculated from theoretically unidentified models and data. Chiou and Walker (2007) demonstrate this problem with small examples and some real data applications. They conclude that it is critical to use far more than 200 simulation draws (using modified Halton sequences, or 2000 pseudo-random draws) to avoid this problem. Of course their results only apply to the models they tested, and it is likely that thousands (or even tens of thousands) of draws may be necessary for more realistic large-scale models. Increasing the number of simulation draws by an order of magnitude will of course increase computation time by a similar amount. These results have been confirmed in a recent Monte Carlo study by Hess and Train (2010), and they provide strong justification for the Bayesian methods discussed in the next section of this paper.

Simulation noise also impacts inference in mixed logit and more general maximum simulated likelihood estimators. While it has long been recognized (see Revelt and Train, 1998) that it is necessary to use the "sandwich estimator" instead of the simpler information matrix of the simulated likelihood function, the difficulty of computing the hessian of the simulated likelihood function has not been as widely appreciated. Bhat, *et al*, (2010a) show that there are numerical problems computing the hessian matrix using numerical derivatives, and the analytic second derivatives can be very tedious to derive for many models.

The mixed logit model has been applied to many areas of travel behavior analysis, and a key theme to these applications is the complexity of heterogeneity in many settings. A number of papers have explored the idea that some heterogeneity is due to scale variation across respondents. This type of heterogeneity manifests itself in the absolute scale of the utility parameters (σ_i) varying across respondents, and this rules out the standard specification where each parameter varies independently across respondents. Greene and Hensher (2010) show how to include this in a mixed logit model, but they do not find any strong evidence of scale heterogeneity in the one application they examined.

Scale heterogeneity has also been an important issue when jointly modeling SP and RP data. Yáñez, *et al*, (2010) show how to handle the resulting heteroskedasticity in both mixed logit and GEV models. Their paper highlights the problem of forecasting from these models when the scales and/or correlations across alternatives depends on SP data. Since we typically want to forecast the demand for new alternatives in the real world, we would therefore like to use the RP scales. In many settings this is not possible, and this raises some important issues in using these models for forecasting. Forecasting from random parameter models like mixed logit is tricky and has not been carefully studied. I will discuss this area more later in this paper.

Modern SP studies typically include multiple choice tasks for each respondent. Revelt and Train (1998) suggested a computationally simple approach for capturing the resulting intra-respondent correlation across choice tasks. Revelt and Train specified that the utility function parameters varied across respondents but were constant across choices made by the same respondent. This specification allows for the estimation of individual-specific utility functions when there are sufficient numbers of choice tasks for each respondent. Hess and Hensher (2010) have used these individual-specific choice parameters to infer which attributes were ignored in a complex SP experiment by groups of respondents. It is very easy to use standard mixed logit code to estimate this specification, since it only requires holding the parameter draws in the simulation fixed for each choice made by the same respondent.

More recent work (see Hess and Train, 2011) has found evidence of more complex patterns of heterogeneity within and across choices made by the same respondent. These findings considerably complicate the task of specifying and estimating accurate mixed logit models in settings with repeated choices. These situations arise naturally with SP experiments, but as Yáñez, *et al*, (2011) point out they can be very important in analyzing panel data with repeated RP measurements. As Hensher and Greene (2003) clearly point out, specification choices about how to model intra and inter-respondent heterogeneity impact all outcomes from mixed logit models – including willingness-to-pay (WTP) measures.

Scarpa, *et al*, (2008) show that it is possible to reparameterize the utility function to isolate WTP in a single parameter instead of the usual ratio of time and cost parameters. This reparameterization allows them to control the distribution of WTP and also separate out the impacts of scale heterogeneity. However the utility function is now a nonlinear function of the parameters and this makes it very difficult to estimate using maximum simulated likelihood techniques. However they were able to estimate the model using hierarchical Bayes methods which will be described in the next section of the paper.

2.1 Other Models

There has been a lot of new work on extensions to the classic discrete choice model. Bhat (2005; 2008) and Fang (2008) have both modeled household choice of vehicles and utilization. This situation is formally a joint discrete/continuous choice model with the added problem that a household can hold more than one vehicle. Since the standard discrete choice model assumes that each respondent chooses one and only one alternative, this requires "exploding" the choice set to be all possible combinations of vehicles a household can hold. Bhat's MCDEV model provides an elegant solution based on directly solving the household's constrained optimization problem. The simplest version of this model retains the IIA assumption across vehicles, but Bhat has also specified mixed versions of the model to relax this constraint.

Bhat's model is derived under the assumption that the total miles driven by a household is fixed, and this can be unrealistic for evaluating policies that are aimed at reducing vehicle travel. Fang takes a more reduced form approach and specifies a system of ordered probit and tobit equations with unconstrained error correlations across all equations. The ordered probits correspond to the household choice of vehicle types, and the tobits correspond to miles driven by the chosen vehicles. Fang estimates her model by Bayesian techniques, and demonstrates that the model can predict reductions in vehicle holdings and vehicle miles traveled in response to changes in residential density.

One difficulty in modeling joint discrete/continuous choices is that it has been difficult to specify tractable joint distributions for the discrete and continuous parts of the model. Recent work with copulas have greatly alleviated this problem by allowing modelers to “stitch together” convenient marginal distributions for the discrete and continuous parts into a flexible joint distribution. Spissu, *et al*, (2009) provide a good introduction to this promising methodology.

When copulas are applied to modeling spatial dependence (e.g., detailed location choice) then the likelihood function rapidly becomes intractable. Bhat and coauthors have recently adopted approximate Composite Maximum Marginal Likelihood methods to deal with this problem (Bhat, *et al*, 2010a; b; Bhat, 2011; Bhat and Sidharthan, 2011). The key idea here is to only consider pairs (or triplets) of discrete alternatives and maximize the marginalized likelihood across these choices. If enough different pairs are included then all parameters of even very high dimensional problems can be identified. Although this method is not as efficient as maximum likelihood, it can be applied in situations where even maximum simulated likelihood is infeasible. Furthermore, Bhat and Sidharthan (2011) show simulation evidence that approximate composite maximum likelihood estimators perform at least as well (and many times better) than maximum simulated likelihood for inference in complex discrete choice models.

3. BAYESIAN MODELS

Train’s (2009, Chapter 12) book popularized the idea of using Bayesian numerical techniques to estimate mixed logit model. He pointed out that in large samples Bayes posterior mean estimates and their associated highest posterior density confidence regions will equal the classical maximum likelihood estimators. For example, Scarpa, *et al*, (2008) used Bayesian techniques to estimate their choice models in WTP space since maximum simulated likelihood estimation didn’t work well for their non-linear utility formulation. Similarly Bento, *et al*, (2009) used Bayesian estimation techniques for their joint model of household vehicle choice and utilization. Sillano and Ortúzar (2005) compared Bayesian and maximum simulated likelihood for a mixed logit model. They concentrated on WTP measures, and found Bayesian methods generally superior. Wang and Kockelman (2009) use Bayesian methods to estimate a complex dynamic ordered probit model, and they report that these methods are much easier and more accurate than maximum simulated likelihood. Alvarez-Daziano and Bolduc (2011) use Bayesian techniques to estimate a hybrid choice model of “green vehicle” choice, and they report that maximum simulated likelihood is infeasible for this model. Kim, *et al*, (2003) use Bayesian methods to estimate a high-dimensional multinomial probit model. Washington, *et al*, (2009) is a particularly interesting paper since it uses Bayesian multiple imputation techniques to deal with imputation of non-chosen attribute values in RP surveys – a potentially serious problem that is frequently ignored.

Since Train’s (2009) Chapter 12 does an excellent job describing general Bayesian methodology and its application to discrete choice modeling, I will limit this section to describing Bayesian methodology that is not mentioned in Train (2009).

The key difference between Bayesian and classical statistics is that Bayesians treat parameters as random variables. Bayesians are therefore led to summarize their prior knowledge about parameters θ by a *prior* distribution, π(θ). The sampling distribution, or likelihood function, is given by $f(x|\theta)$. After observing some data, the information about θ is given by the *posterior* distribution:

$$p(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d(\theta)} \qquad (2)$$

Note that all inference is based on this posterior distribution. In many circumstances (under quadratic loss) the optimal Bayes estimator is the mean of the posterior distribution, and Bayesian confidence bands are typically given by the smallest region of the posterior distribution with the specified coverage probability. Bayesian confidence regions are interpreted as fixed regions containing the random parameter θ with the specified coverage probability (called Highest Posterior Density regions). This is very different from the classical confidence region, which is a region with random endpoints that contain the true value θ with the specified probability over independent repeated realizations of the data. Classical inference therefore depends on the distribution of unobserved realizations of the data, whereas Bayesian inference conditions on the observed data. Bayesian inference is also exact and does not rely on asymptotic approximations.

A key problem with applied econometric work is model uncertainty; we are never certain that we have specified the correct model and almost all empirical work just reports inference conditional on some model being correct. Leamer (1983) clearly shows that this practice (sometimes also called pretest, data mining, or data snooping bias) can lead to very misleading inferences. There is no solution to this problem within the classical statistical paradigm - particularly in the case where competing models are not nested in each other. For example, many applications of nested logit models have correlation structures (commonly called tree structures) that cannot be nested in a larger nested logit model.

The Bayesian approach to this problem does not require the choice of a correct model. Inference can be carried out unconditional on model choice. Suppose there are *M* competing models indexed by *m* with likelihood $f_m(x|\theta)$ and prior density $P_m(\theta)$. Let π_m be the prior probability that model *m* is correct. The marginal data density for model *m* by:

$$f_m(x) = \int f_m(x|\theta)p_m(\theta)d\theta \qquad (3)$$

Note that the marginal density can be hard to simulate in complex models, but new algorithms by Jeliazkov and Chib (2001), Chib and Jeliazkov (2005) have greatly reduced the computational burden. The posterior probability that model *m* is correct is given by:

$$\overline{\pi_m} = \frac{\pi_m f_m(x)}{\sum_{j=1}^{M} \pi_j f_j(x)} \tag{4}$$

These posterior probabilities might suggest that there is an obvious correct model, but in any case the unconditional posterior distribution for θ is then given by:

$$\bar{p}(\theta|x) = \sum_{j=1}^{M} \overline{\pi_j} p_j(\theta|x) \tag{5}$$

In the common case where there is uncertainty about the correct model, then averaging over models as in equation (3.4) will almost always yield better results than arbitrary choice of a "correct" model.

One important advantage of Bayesian analysis of flexible discrete choice models can be clarified by considering the all too frequent case where at least some of the covariance parameters are poorly identified. This means that the likelihood function will be almost flat along the dimensions corresponding to these parameters, and classical methods will therefore have problems converging to the optimum. As long as proper prior distributions are used Bayesian sampling methods should converge, but it is very likely that the resulting posterior distribution for the poorly identified parameters will have the same shape as the prior distribution for these parameters. However, it is still possible to carry out informative Bayesian inference on other parameters of interest. More generally, comparison of the information conveyed in the prior and likelihood distributions is an excellent way to quantify the relative importance of these two inputs into Bayesian inference.

4. INFERENCE AND FORECASTING FOR COMPLEX CHOICE MODELS

While we all fit increasingly complex models of travel behavior, we are still mostly interested in derived quantities like WTP measures, elasticities, or scenario forecasts. Bayesians base all forecasts off the posterior distribution and inference for these forecasts are given by highest posterior density regions calculated from the posterior distributions. The situation is not so clear for classical statisticians as illustrated by Sillano and Ortuzar's (2005) finding that the implied distribution of WTP measures from a panel data mixed logit model are very different depending on whether the computations are conditioned on the individual-specific utility parameter estimates. Bayesians would implicitly condition on these individual-specific parameters since they condition all inference on the observed choices.

Hess and Daly (2009) have produced a nice Excel tool for calculating standard errors for derived quantities like WTP measures. They claim that their methods are superior to a bootstrap procedure where the standard error of the WTP measure is simulated by draws from the asymptotic covariance matrix of the underlying parameters. The correct Bayesian approach would be to simulate draws of the WTP measure from the posterior distribution of the parameters and calculate highest posterior density confidence regions from these simulations. This Bayesian procedure is asymptotically equivalent to the "percentile" method of constructing bootstrap confidence regions, and repeating the Monte Carlo exercises in Hess and Daly (2009) shows that these bootstrap percentile intervals perform similarly to their intervals. I should also note that Krinsky and Robb (1986) provide another accurate method for calculating confidence intervals for WTP measures.

Ten years ago I argued that we needed to be much more careful about giving confidence regions for all of our derived measures and forecasts. The methods for computing these confidence regions are more widely available, but still not widely applied. I suspect that the reason for this is that these intervals are probably embarrassingly wide. If we calculate and publish more of these intervals then we will be forced to face the tradeoff between simpler (possibly biased) models and wide confidence regions.

5. CONCLUSION

Over the last ten years there have been many more applications of mixed logit, and there have also been a number of papers pointing out problems with maximum simulated likelihood estimation of these models. As these problems have become more evident there has been an explosion (albeit from a very small base!) of papers using Bayesian techniques to estimate these models. Bhat and his co-authors have also developed very promising non-simulation estimators using copulas and composite marginal likelihood estimation.

However, almost all of the recent Bayesian applications have failed to take full advantage of the Bayesian approach to quantify problems caused by model uncertainty. If this were done then I am quite certain that many of the more complex new models would not turn out to be much of an improvement over simpler approaches. There have also been very few applications of multiple imputations (see Steimetz and Brownstone, 2005 and Washington, *et al*, 2009) even though there are many potential applications in transportation. Hopefully these shortcomings will be addressed by the members of IATBR during the next ten years!

ACKNOWLEDGEMENT

Chandra Bhat, David Hensher, Stephane Hess, and Juan de Dios Ortúzar provided many useful suggestions and references. Of course, none of these people are responsible for any of the errors or omissions in this paper. Financial support was provided by the University of California Transportation Center.

REFERENCES

Alvarez-Daziano, R. and Bolduc, D., 2011. Incorporating pro-environmental preferences toward green automobile technologies through a bayesian hybrid choice model. *Transportmetrica* , March, 1-33, iFirst.

Bento, A.M., Goulder, L.M., Jacobsen, M.R. and von Haefen, R.H., 2009. Distributional and efficiency impacts of increased US gasoline taxes. *American Economic Review*, 99(3), pp.667-699.

Bhat, C.R., 2005. A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B*, 39(8), pp.679-707.

Bhat, C.R., 2008. The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B*, 42(3), pp.274-303.

Bhat, C.R., 2011. The maximum approximate composite marginal likelihood (MACML) estimation of multinomial probit-based unordered response choice models. *Transportation Research Part B*, 45(7), pp.923-939.

Bhat, C.R., Varin, C. and Ferdous, N., 2010a. A comparison of the maximum simulated likelihood and composite marginal likelihood estimation approaches in the context of the multivariate ordered response model. In: W.H. Greene and R.C. Hill, eds. *Advances in Econometrics, Vol. 26: Maximum Simulated Likelihood Methods and Applications*. Bingley: Emerald Group Publishing Limited, pp.65-106.

Bhat, C.R., Sener, I.N. and Eluru, N., 2010b. A flexible spatially dependent discrete choice model: formulation and application to teenagers' weekday recreational activity participation. *Transportation Research Part B*, 44(8-9), pp.903-921.

Bhat, C.R. and Sidharthan, R., 2011. A simulation evaluation of the maximum approximate composite marginal likelihood (MACML) estimator for mixed multinomial probit models. *Transportation Research Part B*, 45(7), pp.940-953.

Brownstone, D., 2001. Discrete Choice Modeling for Transportation. In: D.A. Hensher, ed. *Travel Behaviour Research: The Leading Edge*. Amsterdam: Pergamon, pp.97-124.

Brownstone, D. and Small, K.A., 2005. Valuing time and reliability: assessing the evidence from road pricing demonstrations. *Transportation Research Part A*, 39(4), pp.279-29.

Brownstone, D. and Fang, H.A., 2009. A vehicle ownership and utilization choice model with endogenous residential density. In: *88th Annual Meeting of the Transportation Research Board*. Washington, DC, USA, 11-15 January 2009.

Chib, S. and Jeliazkov, I., 2005. Accept-reject Metropolis-Hastings sampling and marginal mikelihood estimation. *Statistica Neerlandica*, 59(1), pp.30-44.

Chiou, L. and Walker, J.L., 2007. Masking identification of discrete choice models under simulation methods. *Journal of Econometrics*, 141(2), pp.683-703.

Fang, A., 2008. A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B*, 42(9), pp.736-758.

Greene, W.H. and Hensher, D.A., 2010. Does scale heterogeneity across individuals matter? an empirical assessment of alternative logit models. *Transportation*, 37(3), pp.413-428.

Hensher, D.A. and Greene,W.H., 2003. The mixed logit model: the state of practice. *Transportation*, 30(2), pp.133-176.

Hess, S. and Daly, A., 2009. Calculating errors for measures derived from choice modeling estimates. In: *88th Annual Meeting of the Transportation Research Board*. Washington, DC, USA, 11-15 January 2009.

Hess, S. and Hensher, D.A., 2010. Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. *Transportation Research Part B*, 44(6), pp.781-790.

Hess, S. and Train, K.E., 2010. Approximation issues in simulation-based estimation of random coefficient models. In: *89th Annual Meeting of the Transportation Research Board*. Washington, DC, USA, 10-14 January 2010.

Hess, S. and Train, K.E., 2011. Recovery of inter- and intra-personal heterogeneity using mixed logit models. *Transportation Research Part B*, 45(7), pp.973-990.

Jeliazkov, I. and Chib, S., 2001. Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association*, 96(453), pp.270-281.

Kim, Y., Kim, T.-Y. and Heo, E., 2003. Bayesian estimation of multinomial probit models of work trip choice. *Transportation*, 30(3), pp.351-365.

Krinsky, I. and Robb, A.L., 1986. On approximating the statistical properties of elasticities. *Review of Economics and Statistics*, 68(4), pp.715-719.

Leamer, E.E., 1983. Let's take the con out of econometrics. *American Economic Review*, 73(1), pp.31-43.

Revelt, D. and Train, K.E., 1998. Mixed logit with repeated choices. *The Review of Economics and Statistics*, 80(4), pp.647-657.

Scarpa, R., Thiene, M. and Train, K.E., 2008. Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps. *American Journal of Agricultural Economics*, 90(4), pp.994-1010.

Sillano, M. and Ortúzar, J.de D., 2005. Willingness-to-pay estimation with mixed logit models: some new evidence. *Environment and Planning A*, 37(3), pp.525-550.

Spissu, E., Pinjari, A.R., Pendyala, R.M. and Bhat, C.R., 2009. A copula-based joint multinomial discrete-continuous model of vehicle type choice and miles of travel. *Transportation*, 36(4), pp.403-422.

Steimetz, S. and Brownstone, D., 2005. Estimating commuters' "value of time" with noisy data: a multiple imputation approach. *Transportation Research Part B*, 39(10), pp.865-889.

Train, K.E., 2009. *Discrete Choice Methods with Simulation*. 2nd ed. Cambridge: Cambridge University Press.

Wang, X. and Kockelman, K.M., 2009. Bayesian inference for ordered response data with a dynamic spatial-ordered probit model. *Journal of Regional Science*, 49(5), pp.877-913.

Washington, S., Ravulaparthy, S., Rose, J., Hensher, D.A., Pendyala, R.M. and Konduri, K.C., 2009. Bayesian imputation of non-chosen attribute values in revealed preference surveys. In: *88th Annual Meeting of the Transportation Research Board*. Washington, DC, USA, 11-15 January 2009.

Yáñez, M.F., Cherchi, E. and Ortúzar, J.de D., 2010. Defining inter-alternative error structures for joint RP-SP modelling: some new evidence. *Transportation Research Record*, 2175, pp.65-73.

Yáñez, M.F., Cherchi, E., Heydecker, B.G. and Ortúzar, J.de D., 2011. On the treatment of the repeated observations in panel data: efficiency of mixed logit parameter estimates. *Networks and Spatial Economics*, 11(3), pp.393-418.