Keep It Simple: Easy Ways To Estimate Choice Models For Single Consumers

Christine Ebling, University of Technology Sydney, christine.ebling@uts.edu.au
Bart Frischknecht, University of Technology Sydney, bart.frischknecht@uts.edu.au
Jordan Louviere, University of Technology Sydney, jordan.louviere@uts.edu.au

Abstract

We show with Monte-Carlo simulations and empirical choice data sets that we can quickly and simply refine choice model estimates for individuals based on methods such as ordinary least squares regression and weighted least squares regression to produce well-behaved in-sample and out-of-sample predictions of choices. We use well-known regression methods to estimate choice models, which should allow many more researchers to estimate choice models and be confident that they are unlikely to make serious mistakes.

Keywords: Individual-level choice models, linear probability models

Keep It Simple: Easy Ways To Estimate Choice Models For Single Consumers

Introduction

Choice models were first proposed by Thurstone (1929) for pairs of options. Models for multiple choice options are due to Luce (1959) and McFadden (1974). Except for laboratory choice experiments in psychology, it is rare to see discrete choice models estimated for single people; and after Chapman (1984), there was little work on ways to measure and model single person choices in survey applications until recently. Finn and Louviere (1993) proposed a new measurement model and associated way to elicit choices called Best-Worst Scaling that gives individual-level estimates of preferences, values, etc. Marley and Louviere (2005) prove the measurement properties of the Best-Worst approach; and Louviere, et al (2008) extended the approach to allow one to estimate choice models for single persons. They used simulated and empirical data to show that one can use several estimation methods to model individuals.

The purpose of this paper is to show that one can use simple methods to model single person choices, which extends Louviere, et al (2008) to estimation methods familiar to most academics and practitioners, such as conditional logit regression, ordinary least squares regression and weighted least squares regression. The latter two methods yield biased estimates of the choice probabilities, but we demonstrate that one can "correct" these estimates rather simply. More specifically, we show with Monte-Carlo simulations and empirical choice data sets that we can refine simple estimates to produce well-behaved in-sample and out-of-sample predictions of choices. The contribution of the paper is to describe and discuss these methods to allow many more researchers to estimate choice models and be confident that they are unlikely to make serious mistakes.

The objective of the paper needs to be seen in proper context. For approximately the last 15 years there has been the equivalent of an arms race in academic marketing and applied economics. To wit, ever more sophisticated statistical models have been proposed and applied, with the result that the barriers to entry to do frontier academic research and sophisticated applied research are now extremely high. One result has been an increasing reliance of practitioners on commercial software that "does" choice experiments and allows one to produce the associated statistical models. Considering that the closest analogue to discrete choice models are models for statistical (quantum) mechanics in physics, one might well ask what our colleagues in the physical sciences would think of many under-trained practitioners "doing" quantum mechanics using off-the-shelf commercial software. A second outcome of this arms race is a gradual decline in new academics entering the field, which also has led to a fairly significant decline in papers on choice modelling presented at conferences like EMAC, Marketing Science and ANZMAC. Thus, the field could benefit from simple, easy-to-use methods that are highly likely to produce correct answers.

Of course, correctly analysing choice data is only one part of a larger whole associated with academic or applied work in choice modeling. One must first conceptualise academic and practical problems involving consumer and other choices and understand how to design and implement choice experiments. This paper does not address the precursors to choice data analysis, which require education, training and experience; but given that academics and practitioners can acquire these skills, the approaches proposed in this paper should allow many researchers to estimate choice models for single people and predict choice outcomes in many circumstances of academic and practical interest.

Proposed Modelling Approaches

Selection of one alternative from among several is the observed dependent variable of interest in this paper. The statistical models for the analysis of such data are known as "latent dependent variable" models because the dependent variable of interest is strength of preference, but one observes only a binary indicator of that unobserved, underlying measure. Typically, the responses to choice experiments are coded one (1) for the chosen option and zero (0) for the unchosen option(s). As we later note, all latent dependent variable models have a formal identification problem, namely the standard deviation of the error component is perfectly inversely correlated with the model estimates. This poses no problem in predicting choice probabilities; however, as noted by Swait and Louviere (1993), it poses issues for comparisons of model estimates across data sources. Further, as noted by Fiebig, et al (2010), if the standard deviation of the error component is not constant across people, this can result in seriously biased model estimates and misleading inferences. The latter consideration motivates one to find ways to account for choice consistency differences across people. One way to treat differences is to estimate models for single persons making choices.

In particular, we investigate ordinary least squares regression models estimated directly from the 1, 0 choice indicators (known as "Linear Probability Models", or LPMs), and weighted least squares regression models estimated from full or partial rankings of the choice options in each set (hereafter termed "WLS" models) as proposed and applied by Louviere, et al (2008). The LPM approach is motivated by the work of the Nobel Laureate James Heckman and co-author Synder (1998), who proposed that LPMs provide a better way to model discrete choices when errors are asymmetric; naturally, no one knows if errors in latent utility functions are symmetric or not, so this approach is worth considering. The WLS approach is motivated by rank order explosion of choices, originally discussed by Luce and Suppes (1965) and first used in marketing by Chapman and Staelin (1982). We use rank order explosion procedure here to construct up to N-1 choice sets from a set of N rankings, or r-1 choice sets from a subset of r rankings of N options. Rank order explosion also can be seen as a way to aggregate choices; i.e., 1, 0 choices are fully disaggregated because they are associated with the most elemental level of choices, namely a particular person, a particular choice set and a set of particular choice options (alternatives). Hensher and Louviere (1984), Louviere and Woodworth (1983) and Louviere, et al (2008) (among others) show how to aggregate choices across people or choice sets by calculating and analysing choice counts or choice proportions. Non-experimental analogues of these aggregate choice counts were considered by (among others) Berkson (1944) and Theil (1969) for cases where real market choices are aggregated by choice alternatives and choice sets or other differences in choice observations. If one has such choice counts or one can transform rankings into such counts (via rank order explosion or Louviere, et al 2008), one can apply WLS regression to estimate choice model parameters. Green (1984) discusses using WLS as the first step of the maximum likelihood estimator, which is the approach used by Louviere and Woodworth (1983).

In the case of both LPMs and WLS, the model estimates are on the wrong "scale" to correctly predict the choice probabilities. That is, the parameter estimates are too small or too large and will systematically under- or over-predict the observed choice probabilities, and must be corrected. We propose and apply a simple correction that appears to work quite well under a variety of simulated and actual conditions in- and out-of-sample. Specifically, one estimates an LPM or a WLS model for a single person and uses the model to predict the observed dependent variable for each choice option in each choice set of interest. Then, one calculates the associated residual mean squared error for each person, and multiplies all LPM or WLS model estimates for an individual by a correction factor, which is 1/(mean squared error). The "corrected" parameters are used to make choice predictions for each person by using the logit formula.

We applied both LPM and WLS methods to predict observed choices in two empirical data sets representing choices of car insurance options (set 1) and choices of cross-country airline flights (set 2); and a third set of simulated choices (set 3). We also compared each approach with individual-level conditional logit models estimated from exploded choice data based on observing most and least preferred choices in each choice set. In the interests of space, we do not go into detail on exploding the choice data using the observed partial rankings, but merely note that we explode the data to be consistent with Luce and Suppes (1965) and Chapman and Staelin (1982). In the airline and car insurance data we estimate the models from choices of 150 out of 200 respondents in 12 choice sets with four options each constructed using Street and Burgess (2007) optimal design theory. We assess in-sample and out-of-sample fits using the method of sample enumeration where we average the choice share for a particular alternative across the individual choice probabilities of the in-sample or out-of-sample respondents respectively. We use observed and predicted choice alternative shares to compare models in- and out-of-sample using R-Squares, hit and fit rates (150 respondents for in-sample; the remaining 50 respondents for out-of-sample). We use an additional four choice sets to assess holdout task fits, again using R-Square and hit rates. To avoid biases that may be due to the particular respondents selected for estimation, we average our fit measures over ten different random draws of 150 individuals. We follow the same procedure for the simulated choice experiment data; we generate simulated choice data by drawing a set of parameter values from specified distributions for each simulated person. We generated 100 simulated populations, and sampled 150 people randomly ten times from each population. We average the fit measures across populations and samples.

Results

Conditional logit models estimated from the exploded data consistently perform well in-sample. When the correction factors are applied to the LPMs and the WLS model estimate, they also perform well in-sample. All models systematically tend to over-predict low choice share alternatives and under-predict high choice share alternatives for out-of-sample choice sets and samples. This is likely due to differences in error variances between estimation and hold-out data sources, such that error variances in out-of-sample sets are systematically larger than in estimation sets. The likely source of the error variance differences is model mis-specification; which in turn suggests larger mis-specification errors for the car insurance data, also implying more non-additive choice rules for insurance than flight choices.

The results are summarised in Table 1, which reports model performances for four cases: 1) in-sample, same people, same choice sets; 2) out of sample, same people, different choice sets; 3) out of sample, different people, same choice sets; and 4) out of sample, different people, different choice sets. Shown are the hit rates, i.e. the average number the alternatives assigned the highest choice probabilities in a given choice set that were the same as the alternatives chosen, the fit rates, i.e. the mean squared deviation between predicted choice probabilities and individual observed 0/1 choices, the R^2 for the line y=x (i.e., aggregate observed shares equal to aggregate predicted shares), as well as the R^2 and the slope of the regression of the observed aggregate choice shares on the predicted aggregate choice probabilities. Whereas hit and the fit rates measure a models' ability to predict individual person choices, R^2 values and regression slopes reflect a models' ability to produce good aggregate predictions.

Table 1: Results of Model Estimation Tests

Model Performance	Exploded Cond Logit			OLS not corrected			OLS x 1/(msq error)			WLS not corrected			WLS x 1/(msq error)		
	Set1	Set2	Set3	Set1	Set2	Set3	Set1	Set2	Set3	Set1	Set2	Set3	Set1	Set2	Set3
In Sample (Same respondents, same choice sets)															
Hit rate	0.92	0.93	0.56	0.91	0.92	0.66	0.91	0.91	0.66	0.93	0.93	0.65	0.92	0.93	0.65
Fit index	0.03	0.02	0.14	0.13	0.12	0.16	0.03	0.03	0.11	0.10	0.10	0.14	0.03	0.03	0.13
R^2, perfect	1.00	1.00	0.90	0.40	0.44	0.39	0.98	0.99	0.90	0.61	0.60	0.65	0.99	0.99	0.86
R^2, best fit	1.00	1.00	0.92	0.90	0.94	0.90	0.98	0.99	0.90	0.91	0.91	0.90	0.99	0.99	0.91
Slope, best	1.02	1.02	0.88	3.92	3.67	4.04	1.04	1.01	1.02	2.38	2.42	2.12	1.06	1.04	0.83
Out of Sample within respondents (same respondents, different choice sets)															
Hit rate	0.63	0.60	0.33	0.61	0.59	0.31	0.61	0.62	0.31	0.64	0.62	0.32	0.64	0.64	0.32
Fit index	0.17	0.18	0.23	0.16	0.15	0.19	0.16	0.15	0.23	0.14	0.14	0.19	0.14	0.14	0.27
R^2, perfect	0.93	0.91	0.78	0.30	0.37	0.36	0.87	0.97	0.84	0.47	0.56	0.64	0.89	0.96	0.70
R^2, best fit	0.97	0.92	0.85	0.84	0.85	0.85	0.91	0.97	0.85	0.86	0.83	0.87	0.95	0.98	0.85
Slope, best	1.24	1.10	0.81	5.02	3.95	4.18	1.25	1.01	1.05	3.04	2.35	2.07	1.32	1.14	0.72
Out of Sample across respondents (different respondents, same choice sets)															
Hit rate	0.55	0.61	0.40	0.57	0.61	0.35	0.56	0.60	0.35	0.56	0.61	0.35	0.56	0.60	0.35
Fit index	0.15	0.13	0.18	0.17	0.16	0.18	0.14	0.13	0.18	0.16	0.15	0.18	0.14	0.13	0.18
R^2, perfect	0.92	0.95	0.71	0.38	0.45	0.29	0.90	0.93	0.63	0.58	0.60	0.49	0.92	0.94	0.58
R^2, best fit	0.93	0.96	0.75	0.84	0.92	0.65	0.91	0.95	0.64	0.86	0.89	0.66	0.93	0.95	0.65
Slope, best	0.97	1.01	0.87	3.89	3.53	3.87	0.99	0.94	0.97	2.38	2.32	2.05	1.03	0.98	0.79
Out of Sample across respondents across choice sets(different respondents, different choice sets)															
Hit rate	0.68	0.66	0.37	0.71	0.68	0.38	0.71	0.68	0.38	0.71	0.68	0.38	0.71	0.68	0.38
Fit index	0.12	0.13	0.19	0.17	0.16	0.18	0.12	0.12	0.18	0.15	0.15	0.18	0.12	0.13	0.18
R^2, perfect	0.93	0.89	0.61	0.29	0.36	0.30	0.84	0.94	0.68	0.46	0.53	0.53	0.86	0.91	0.51
R^2, best fit	0.95	0.90	0.70	0.81	0.81	0.70	0.87	0.95	0.71	0.83	0.79	0.72	0.91	0.94	0.70
Slope, best	1.16	1.13	0.81	5.05	4.00	4.22	1.23	1.04	1.06	3.04	2.38	2.08	1.31	1.16	0.71
R^2 perfect = estimated thru origin; R^2 best = estimated linear fit; Slope best = slope of estimated linear fit															

Theoretically, model performance should systematically decline from cases 1 to 4. Let us now consider the various performance measures: 1) Hit rates – these are uniformly high for case 1, indicating one correctly predicts the chosen options a very high proportion of the time in-sample; hit rates decline for cases 2 and 3, but surprisingly rise in case 4; 2) Fit rates for both corrected LPM and corrected WLS models are closer to 0, and thus better, than exploded conditional logit models. Considering hit-rates and fit-rates, corrected LPM and WLS models perform as well for individual persons as exploded conditional logit models. 3) R^2 values are uniformly high for conditional logits using exploded choices; they are consistently high for corrected LPM and WLS models, and are consistently lower for non-corrected models; 4) Slopes are close to one for conditional logit on exploded choices and corrected LPM and WLS models, but differ widely for non-corrected models. All models tend to over-predict low choice share alternatives and under predict high choice share alternatives out-of-sample with different choice sets.

Discussion and Conclusions

We proposed and investigated two simple ways to model individual choices, namely linear probability models (LPMs) and Weighted Least Squares regression based on Louviere, et al (2008). We tested the performance of the two approaches in- and out-of-sample. Our results suggest that LPMs and WLS models corrected for error variance difference between individuals perform well in both real and simulated data once the parameter estimates are weighted by the inverse of the mean squared error. Because of their simplicity and ease of use, we suggest that both approaches are likely to prove attractive to researchers who want to better understand and model consumer choices, but find complex statistical models or black box commercial software challenging. In turn, this should allow both academics and practitioners to do a reasonable job of predicting choices without making large mistakes.

References

Berkson, J. 1944. Application of the Logistic Function to Bio-Assay. Journal of the American Statistical Association 39, 357-365.

Chapman, R.G., Staelin, R. 1982. Exploiting Rank Order Choice Set Data Within the Stochastic Utility Model. Journal of Marketing Research 19, 288-301.

Chapman, R.G. 1984. An Approach to Estimating Logit Models of a Single Decision Maker's Choice Behavior. Advances in Consumer Research 11, 656-661.

Fiebig, D.G., Keane, M.P., Louviere, J., Wasi, N. 2010. The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. Marketing Science 29, 393-421

Finn, A., Louviere, J.J. 1992. Determining the Appropriate Response to Evidence of Public Concern: The Case of Food Safety. Journal of Public Policy and Marketing 11 (1) 12-25.

Green, P.J. 1984. Iteratively Reweighted Least Squares for Maximum Likelihood Estimation, and some Robust and Resistant Alternatives. Journal of the Royal Statistical Society, Series B (Methodological) 46(2), 149-192.

Heckman, J.J., Snyder, J.M. 1997. Linear Probability Models of the Demand for Attributes With an Empirical Application to Estimating the Preferences of Legislators. Rand Journal of Economics 28, 142-189.

Hensher, D.A., Louviere, J.J. 1983. Identifying Individual Preferences for International Air Travel: An Application of Functional Measurement Theory. Journal of Transport Economics and Policy 17(3), 225-245.

Louviere, J.J., Woodworth, G. 1983. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. Journal of Marketing Research 20, 350-367.

Louviere, J.J., Street, D., Burgess, L., Wasi; N., Islam, T., Marley, A.A.J. 2008. Modeling the Choices of Individual Decision-Makers by Combining Efficient Choice Experiment Designs with Extra Preference Information. Journal of Choice Modelling 1(1), 128-163.

Luce, R.D. 1959. Individual Choice Behavior: A Theoretical Analysis. New York: Wiley.

Luce, D., Suppes, P. 1965. Preference, Utility, and Subjective Probability. In Bush, R.R., Galanter, E. (eds.), Handbook of Mathematical Psychology, 249-410. John Wiley and Sons Inc., New York.

Marley A.A.J., Louviere, J.J. 2005.Some Probabilistic Models of Best, Worst, and Best-Worst choices. Journal of Mathematical Psychology 49, 464-480.

McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (ed.), Frontiers In Econometrics, Academic Press: New York, 105-142.

Street, D., Burgess, L. 2007. The Construction of Optimal Stated Choice Experiments: Theory and Methods. New York: Wiley.

Swait, J.D., Louviere, J.J. 1993. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. Journal of Marketing Research 30, 305-314.

Theil, H. 1969. A Multinomial Extension of the Linear Logit Model. International Economic Review 10, 251-259.

Thurstone, L.L. 1927. A Law of Comparative Judgement. Psychological Review 34, 273-286.