ICT Predictions with Individual Models

TOWHIDUL ISLAM

Towhidul Islam is Associate Professor, University of Guelph, Canada

This paper proposes, describes and illustrates an alternative approach to estimate individual level models for information and communications technologies (ICT) and compares with existing established models such as mixed logit random coefficient and Hierarchical Bayes random coefficient model. In order to estimate individual level models, we have used two recent developments: (a) availability of efficient experimental designs and (b) collection of extra preference information using repeated best-worst questions. Individual level model out performed more complex models both in sample and out of sample model fit.

1 Introduction

This paper proposes, describes, illustrates and discusses a relatively simple way to collect and model individual choices for Information and Communications Technologies (ICT) by combining statistically efficient designs for discrete choice experiments (DCEs) with repeated Best-Worst (BW) questions about choice options in designed choice sets (see details in Louviere et al 2008). The proposed approach should give academics and practitioners a useful way to develop and apply the models. This paper also compares three ways to model choices that potentially can accommodate within and between individual differences among people. The approaches are: 1) a classicist mixed logit random coefficient model estimated using simulated maximum likelihood (Revelt and Train 1998), 2) a Hierarchical Bayes random coefficient model estimated using MCMC (Allenby and Rossi 1993), and 3) a new way to estimate choice models for single individuals (Louviere et al 2008). We consider the first two model types as 'top-down' approaches because one makes assumptions about distributions or errors and preferences, indirect utility functions and choice processes, which if correct, allow one to capture an aggregate distribution of preferences in a sample population. We view the third model type as a 'bottom-up' approach: it makes assumptions about indirect utility functions, choice processes and error distributions for individuals, models each person separately, and then aggregates predictions from each person's model separately.

The individual level approach is attractive on the following theoretical grounds:

1 Kenneth Arrow won the Nobel Prize (1972) for proving that one cannot aggregate individual utility functions using any measure of central tendency, such as mean utility or preference (the "impossibility theorem" 1950). Thus, choice modelers must reconcile ways of aggregating individual choices with Arrow's work, but to date few choice modelers seem to be aware of this issue. What we can say is that simply making assumptions about statistical distributions does not circumvent Arrow's result.

2 There also seems to be little recognition among choice modelers that assuming a common indirect utility specification does not imply 'common' attribute utility scales for choosers, even if the assumption is true (which is empirically unlikely). Indeed, one can show that each person or a context must have a different utility scale with different units of measurement (see eg. Lynch 1985; Louviere 1988; Louviere 2001; Louviere et al 2002; Louviere & Eagle 2006; Magidson & Vermunt 2007). Thus, choice models that do not allow for individual differences in utility scales are theoretically incorrect, however well they might fit choice data.

Individual level models capture Individual level differences – within and between subjects: preference and variance heterogeneity. There are a number of practical reasons also: (a) simple way to model individual choices and develop simulators, (b) give practitioners a choice of models.

2 Modeling Individual Choices

Louviere et al (2008) proposed a way to estimate conditional logit models for single individuals that can be described briefly as follows:

1 Obtain a full or partial ranking of options in each choice set by asking respondents two or more questions about their most and least preferred options in each set (eg. if four options per set, ask most preferred and least preferred; then ask most or least preferred of the remaining two). That is, instead of asking respondents to 'do' more choice sets, ask them more choice questions about each choice set.

 ISSN 0085-7130 ©Telenor ASA 2009

Scenario One			
	Option 1	Option 2	Option 3
Phone style[1)]	A	B	C
Brand	Sony	Nokia	Samsung
Price	$40	$140	$90
Built in camera	No camera	2 megapixel camera	5 megapixel camera
Wireless connectivity	Bluetooth and WiFi connectivity	WiFi connectivity	Bluetooth connectivity
Video capability	Video recording (up to 1 hour)	No video recording	Video recording (more than 1 hour)
Internet capability	Yes	Yes	No
Music capability	MP3 player only	MP3 player and FM radio	FM radio only
Built in memory	512 MB built in memory	2 GB built in memory	4 GB built in memory
1. Which of the three options would you be most likely to choose?	☐	☐	☐
2. Which of the three options would you be the least likely to choose?	☐	☐	☐
3. If you would choose none of the technology, check the box to the right.			☐

[1)] Phone style is disguised for confidentiality

Table 1 A Sample Survey Scenario for the Study

2 Use the ranking obtained from 1 above to estimate choice models via the method of rank order explosion (Luce and Suppes 1965; Chapman and Staelin 1982); or use the ranking to assign weights to each rank order position based on the expected choice counts that should be observed if a person chose consistently with their ranking in each possible choice set (there are 2^J subsets of J choice options in each set). For example, four options per set (J = 4) gives 16 possible choice sets (one is empty). Let the options be A, B, C, D, respectively ranked, 1, 2, 3, 4 (1 = best, 4 = worst), the expected choices (weights) are 8, 4, 2, 1 for these ranks. Louviere et al (2008) show how to estimate conditional logit models based on these weights, and show the model estimates are unbiased.

We use this weighted conditional logit approach to estimate models for individuals. We compare in- and out-of-sample predictions of this approach to predictions from mixed logit (MIXL) and Hierarchical Bayes (HB) models for mobile phones hand set data sets described below. Table 1 lists attributes and levels of mobile phone sets for discrete choice experiments (DCEs) used in the model comparison. We constructed optimal Street and Burgess (2007) main effects only designs for each DCE. Data (ie. sample of 461 individuals) were collected in Guelph, Canada.

3 Results

As previously noted, we estimated MIXL and HB models (Allenby & Rossi 1993). The probability of choosing option i in MIXL can be written as:

$$P_i = \exp[(\beta_k + \omega_k)X_{ki}] \,/\, \Sigma_{j \in C} \exp[(\beta_k + \omega_k)X_{kj}],$$

where β_k is a vector of random effects of the k^{th} attribute including alternative-specific intercepts with associated disturbance terms, ω_k, corresponding to the design matrix of covariates, X_{ki} and X_{kj}. Like many researchers, we assume that the random effects are normally distributed. The MIXL model is not closed form, so we estimate it with simulated maximum likelihood (Revelt & Train 1998) using Ken Train's GAUSS code. We used Bayes rule to estimate the posterior distribution for each person following Train (2003, Chap 11) to estimate utilities for each person. We also compared the individual level model estimates with HB estimates from a similar model specification. We used the HB information on the aggregate taste distribution together with the individuals' choices to derive conditional estimates of each individual's parameters. For both MIXL and HB we use the individual-level estimates to predict the choices of each person and aggregate these choices via sample enumeration.

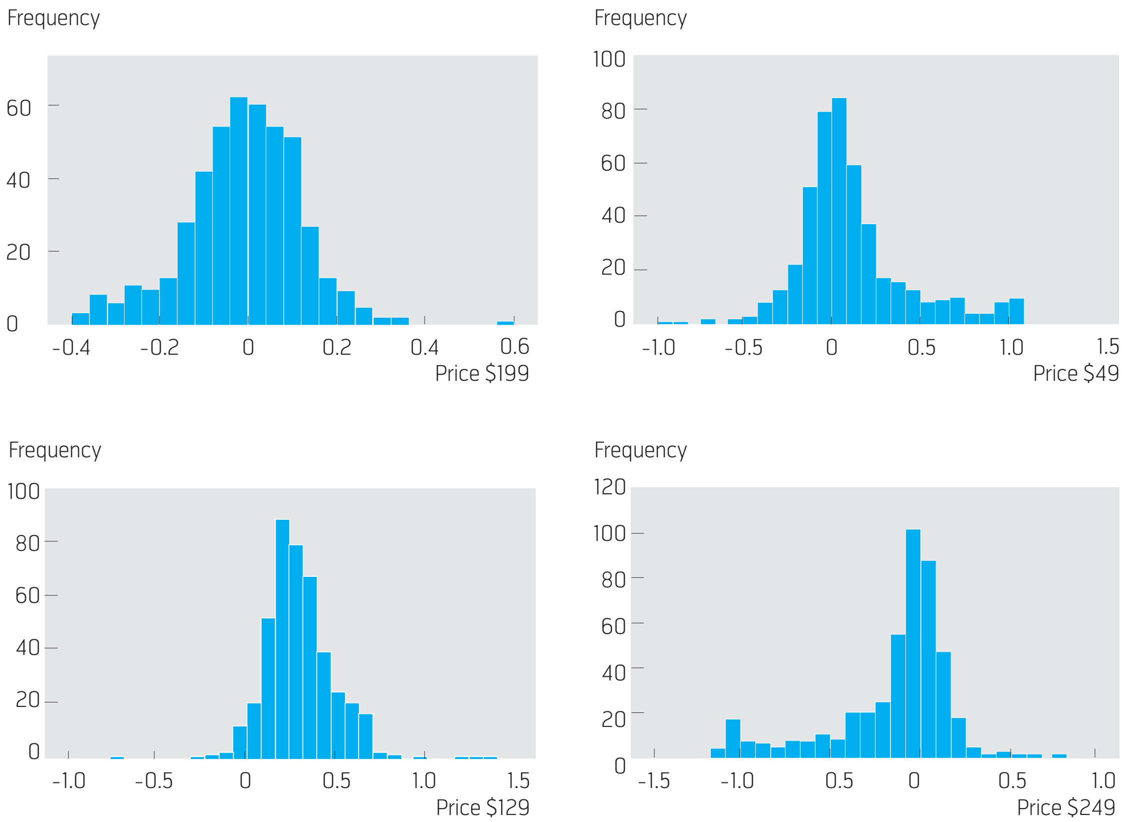

Figure 1 Mobile Phones: Empirical Distributions from Individual Models

Parameters	Estimates	Std Errors	t-ratios	prob
Price $49	0.2963	0.0263	11.27	0.0000
Price $99	0.0705	0.0208	3.39	0.0007
Price $129	-0.0166	0.0231	-0.72	0.4740
Price $199	-0.3502	0.0350	-10.02	0.0000
No camera	-0.4483	0.0249	-18.00	0.0000
2 megapixel camera	0.0408	0.0230	1.78	0.0760
3 megapixel camera	0.1141	0.0228	5.00	0.0000
5 megapixel camera	0.2935	0.0289	10.14	0.0000
No Bluetooth or WiFi connectivity	-0.1752	0.0218	-8.02	0.0000
WiFi connectivity	-0.0283	0.0222	-1.27	0.2034
Bluetooth connectivity	0.0759	0.0193	3.93	0.0001
Bluetooth and WiFi connectivity	0.1276	0.0240	5.33	0.0000
No video recording	-0.1335	0.0219	-6.09	0.0000
Video recording (up to 15 minutes)	0.0115	0.0221	0.52	0.6037
Video recording (up to 1 hour)	0.0635	0.0211	3.01	0.0026
Video recording (more than 1 hour)	0.0585	0.0226	2.59	0.0097
Internet Access	0.1491	0.0185	8.08	0.0000
No music capability	-0.2322	0.0218	-10.63	0.0000
MP3 Music Player only	0.1070	0.0192	5.57	0.0000
FM Radio only	-0.0852	0.0208	-4.10	0.0000
MP3 Music Player and FM Radio	0.2104	0.0255	8.24	0.0000
64 MB built-in memory	-0.1143	0.0219	-5.21	0.0000
512 MB built-in memory	0.0232	0.0202	1.15	0.2499
2 GB built-in memory	-0.0110	0.0220	-0.50	0.6180
4 GB built-in memory	0.1021	0.0236	4.32	0.0000

Table 2 Sample results of an individual

		R^2: Within Sample		R^2: Cross Sample
MIXL HB Indv. model	Sample 1	0.8337 0.8354 0.8807	Sample 1 estimates used to predict Sample 2	0.8287 0.8293 0.8678
MIXL HB Indv. model	Sample 2	0.8247 0.8249 0.8958	Sample 2 estimates used to predict Sample 1	0.7988 0.7958 0.8669

Table 3 Comparison of in- and out-of-sample predictive validity (R-square)

Previous MIXL and HB comparisons (eg. Huber & Train 2001), suggest that both models give nearly equivalent conditional estimates; hence, familiarity, personal preference and estimation ease typically dictate the particular approach that researchers use. Our experience is that aggregate estimates of means and standard deviations of assumed preference distributions agree closely, but individual-level estimates can differ, particularly if less than full design information is used in estimation (typical of many applications), which is why we chose to compare them in this way.

The sample distribution of price coefficients in Figure 1 clearly shows the extent of non-normality of empirical distributions of price preference.

Estimates (excluding brand coefficients for confidentiality) of an individual are shown in Table 2.

Within and Cross-Sample Model Performance

We evaluated the fit of the models in- and out-of-sample using different model fit criteria (R^2, MSE & χ^2) using data from two different samples. All fit criteria agreed, so we only report R^2 values in Table 3. As most choice modelers are unaccustomed to R^2 values, we note that many DCEs assign the same choice sets to multiple choosers, allowing calculation of choice proportions for each option in each choice set. Probabilistic discrete choice models predict choice probabilities, so it is natural to ask how well observed and predicted choice proportions agree.

Table 3 results suggest that the individual-level model approach dominates currently popular and widely used MIXL and HB.

4 Discussion and Conclusion

We compared three approaches to estimating choice models that can produce individual-level parameters. We found that the bottom-up approach of Louviere et al (2008) provided consistently superior in-sample fits. It also dominated out-of-sample fit comparisons. This approach is easy to implement and apply, and can be used in conjunction with many other discrete choice model approaches. It does not require assumptions about preference distributions, and hence is preferred on those grounds to approaches that do require such assumptions. Because one can estimate the unobserved variability associated with each person, one can 'adjust' the models to take this into account. Another advantage is that it can identify lexicographic and similar strategies, as demonstrated by Louviere et al (2008). It also avoids two theoretically problematic issues noted earlier.

Naturally, there are limitations, such as assumptions about choice processes and indirect utility functions. We also do not yet know the extent to which this can be relaxed and made more flexible. There also probably are upper limits to possible numbers of attributes and choice sets, although to date we have had success with problems as large as 32 choice sets, with five options per set and 13 attributes. Our results suggest that researchers should give this approach at least as much attention as popular, but more complex top-down models.

5 References

Allenby, G M, Rossi, P E. 1993. A Bayesian approach to estimating household parameters. *Journal of Marketing Research*, 30 (2), 171-182.

Arrow, K J. 1950. A Difficulty in the Concept of Social Welfare. *Journal of Political Economy*, 58 (4), 328–346.

Chapman, R G, Staelin, R. 1982. Exploiting rank order choice set data within the stochastic utility model. *Journal of Marketing Research*, 19, 288-301.

Huber, J, Train, K. 2001. On the similarity of Bayesian estimates of individual mean partworths. *Marketing Letters*, 12 (3), 259-269.

Louviere, J J. 1988. *Analyzing Decision Making: Metric Conjoint Analysis*. Newbury Park, CA, Sage Publications, Inc. (Sage University Papers Series No. 67)

Louviere, J J. 2001. What if consumer experiments impact variances as well as means: Response variability as a behavioral phenomenon. *Journal of Consumer Research*, 28, 506-511.

Louviere, J J, Eagle, T. 2006. Confound It! That pesky little scale constant messes up our convenient assumptions. *Proceedings, 2006 Sawtooth Software Conference*, 211-228. Sequem, WA, USA, Sawtooth Software, Inc.

Louviere, J J et al. 2002. Dissecting the random component of utility. *Marketing Letters*, 13 (3), 177-193.

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. *The Journal of Choice Modeling*, 1 (1), 128-163.

Feinberg, F, Salisbury, L. 2008. Alleviating the constant stochastic variance assumption in decision research: Theory, measurement and experimental test. *Marketing Science*, forthcoming.

Revelt, D, Train, K. 1998. Mixed logit with repeated choices: Household choices of appliance efficiency level. *Review of Economic Statistics*, 80, 647-57.

Luce, D, Suppes, P. 1965. Preference, utility, and subjective probability. In: Bush, R R, Galanter, E (eds.). *Handbook of Mathematical Psychology*. New York, John Wiley and Sons Inc., 249-410.

Lynch, J G. 1985. Uniqueness issues in the decompositional modelling of multiattribute overall evaluations. *Journal of Marketing Research*, 22, 1-19.

Magidson, J, Vermunt, J K. 2007. *Removing the scale factor confound in multinomial logit choice models to obtain better estimates of preference*. Proceedings, Sawtooth Software Conference, Santa Rosa, CA October.

Revelt, D, Train, K. 1998. Mixed logit with Repeated Choices of Appliance Efficiency Levels. *Review of Economics and Statistics*, LXXX, 4, 647-657.

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

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

For a presentation of Dr. Towhidul Islam please turn to page 82.

ISSN 0085-7130 ©Telenor ASA 2009