Bonoli Escobar, Mariano ; Volpe, Juan Ignacio ; Mosca, Johana ; Picasso, Emilio

Respondent engagement and task complexity effects in stated choice experiments (Resumen extendido)

Ponencia presentada en:
XLI Coloquio Argentino de Estadística, 2013

XLI COLOQUIO ARGENTINO DE ESTADÍSTICA
15 A 18 DE OCTUBRE 2013
MENDOZA - ARGENTINA

RESPONDENT ENGAGEMENT AND TASK COMPLEXITY EFFECTS IN STATED CHOICE EXPERIMENTS

MARIANO BONOLI[1], JUAN IGNACIO VOLPE[2], JOHANNA MOSCA[3], EMILIO PICASSO[4]

1 *Facultad de Ingeniería, UBA,* mbonoli@fi.uba.ar
2 *Facultad de Ingeniería, UBA,* juan.volpe90@gmail.com
3 *Facultad de Ingeniería, UBA,* johamosca@gmail.com
4 *Universidad Católica Argentina,* epicasso@uca.edu.ar

ABSTRACT

Discrete Choice Modelling (DCM) provides a valuable tool for understanding the consumer behaviour. DCM can source on two kinds of empirical base. The first one are preferences revealed by individuals in their actual behaviour. The second kind are preferences stated by them in choice experiments, where the individual is exposed to hypothetical situations specifically designed to elicit decisions. The repeat nature of the choice experiment along a number of choice sets has been recognized as a feature that requires special attention, with the focus primarily on ways of accounting for the correlated structure induced by offering each respondent multiple choice sets in a sequence. Significant research has been produced about this issue after the first paper from Bradley and Daly (1994). The harsh empirical nature of the problem has made it difficult for researchers to be conclusive. They have rather built knowledge on a case by case basis, not exempt of apparent inconsistencies. The present paper brings new information about this issue, extending the present level of knowledge in three directions. A choice experiment was set up in the marketing domain, with unique characteristics vs previous research mostly focused on transportation research. Secondly, we went beyond most previous experience in terms of the number of choice tasks. This enabled us to find the respondent disengagement phenomenon. Thirdly, we address the behavioural nature of engagement along the choice experiment, finding that it is mostly due to boredom rather than due to fatigue. Two versions of DCM are employed: the simple Multinomial Logit model (MNL) and the more sophisticated Mixed Logit (MXL) model, with a suitable specification to measure the respondent disengagement effect as increased variance of utility.

KEY WORDS: discrete choice, stated choice, engagement, fatigue, boredom, consumer behaviour

2. METHODOLOGY

2.1. Data description

The data was originated in a packaged goods simulated purchase experiment with 28 products of two brands and different flavours. Consumers of the product category were recruited via street intercepts in different points of Buenos Aires, and invited to interact with a web-based instrument running the choice experiment. Assistance was available. The sample included 400 female consumers spanning all ages and socio-economic levels.

The experiment included 28 multiple purchase tasks. A subset of the products was presented in each task. The tasks number 1 to 8 had 8 alternatives to choose from, whereas tasks number 9 to 28 had 16 alternatives each. The experimental design was crafted to foster trade-off situations where the flavours offered by each brand were different, for the consumer to make a decision: to stay with preferred brand and change preferred flavour or vice versa. The prices were kept constant at the market level.

2.2. Discrete choice models

The discrete choice models are based on the theory of random utility. This theory establishes that individuals choice can be explained by the maximization of a latent variable, which we now call utility, that is a random variable.

Each individual i in the sample makes T choice tasks, among J alternatives. The structure of the random utility model is:

$$\tilde{U}_{jit} = {}^{\mathsf{T}}\beta x_{jit} + \tilde{\varepsilon}_{jit} \qquad (1)$$

Where:

U_{jit}: Utility (random) of alternative j for the individual i in the choice task t.

x_{jit}: k-dimensional vector of characteristics of the alternative j or the individual i in the choice task t performed by individual i.

β: k-dimensional vector of partial utilities (or importance) of the characteristics of alternatives or individuals.

ε_{jit}: Random component of utility, distributed as Gumbel type I (maximum).

The wavy stripe on top of variables is employed to highlight their random nature.

The explanatory variables x_{jit} can include indicator variables for alternatives. Their coefficients are then called alternative specific constants.

The MNL model has proven very useful in many applications, due to its simplicity and robustness. However it is not exempt of weaknesses. One of them is its limitation to represent individual heterogeneity. A powerful way to cope with individual heterogeneity is by means of the Mixed Multinomial Logit model (MXL). This one establishes a MNL structure for each individual in the sample:

$$P_{jit} = \frac{\exp({}^{\top}\beta_i\, x_{jit})}{\sum_{l=1}^{J} \exp({}^{\top}\beta_i\, x_{lit})} \tag{2}$$

Where the parameters vary randomly among individuals according to a pre-specified probability law.

2.3.Analytic strategy

We first take a simple methodological approach by using the MNL model. The model is estimated on a subset of the empirical base, formed by tasks number 9 to 16 (the earliest tasks having full complexity: 16 alternatives) for all individuals. Predicted probabilities are compared to actual choices to calculate the fraction of hits for each task, across all individuals. Two indicators are created. Top 1 indicator is calculated as the fraction of choice tasks where the top probability alternative was selected by the individual. The second indicator is Top 3, calculated as the fraction of choice tasks where the selected alternative belongs to the set of top 3 probability alternatives.

Then we employed an error component specification of the MXL model to investigate engagement more precisely. The database was partitioned into four sets: tasks number 1 to 16, 17 to 20, 21 to 24, and 25 to 28, across all individuals. An indicator variable is created for each partition: $z_{1jit} = 1$ for tasks 1 to 16, and 0 otherwise, $z_{2jit} = 1$ for tasks 17 to 20, and 0 otherwise, etc. The utility specification is as follows:

$$\tilde{U}_{jit} = {}^{\top}\beta\, x_{jit} + \tilde{\varphi}_{pi}\, z_{pjit} + \tilde{\varepsilon}_{jit} \tag{5}$$

Where $\tilde{\varphi}_{pi}$ is a random parameter with 0 mean, meant to represent the additional variance introduced in the random utility by disengagement or learning. Normal distribution is assumed for this parameter. One of the new variables has to be set as reference, and it must correspond to the group of tasks with minimum utility variance, as variance is always positive. We have chosen tasks 1 to 16 as reference.

Fatigue or boredom would be revealed by statistically significant variances for the random parameters $\tilde{\varphi}_{pi}$, increasing as the experiment makes progress.

3. RESULTS AND DISCUSSION

The MNL model result is very strong with all but one statistically significant parameters. The graphic in Figure 1 shows Top 1 and Top 3 indicators of the portion of hits. Both indicators follow parallel trends, where Top 1 lays below Top 3, as expected, being sharper. They are higher for the first 8 tasks, and they drop sharply at task number 9 to stay fairly stable from there on. This sudden drop in prediction accuracy should not be interpreted as a learning effect, oddly happening at task number 9. Task complexity provides a better explanation for the drop in prediction accuracy. Tasks number 1 to 8 have 8 alternatives to choose from, whereas tasks number 9 to 28 have 16 alternatives. The model is approximately 50% more accurate in predicting choice among 8 alternatives than among 16 alternatives.

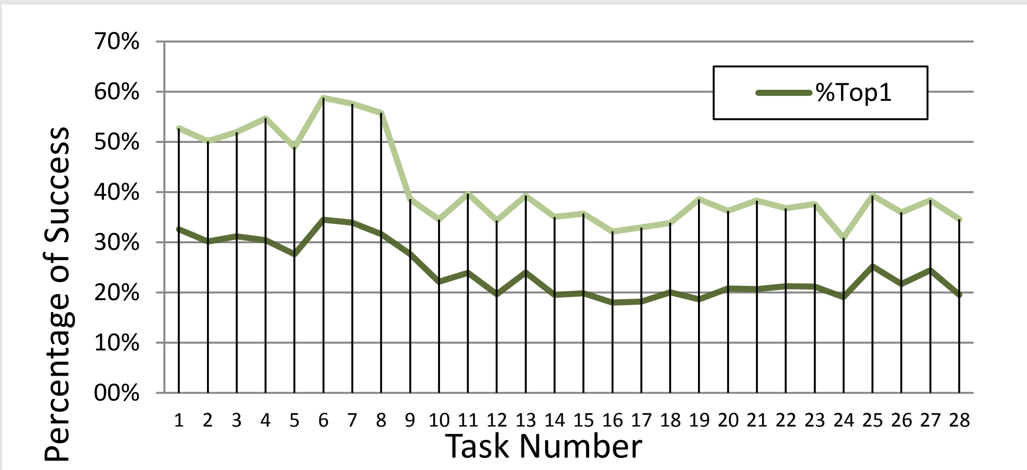

Figure 1

The results for the MXL model, specified as described in the methodology section, is shown in Table 1.

	Mean			**SD**		
Variable	Estimate	Pr(>|t|)	Signif.	Estimate	Pr(>|t|)	Signif.
Brand 1	-0.352	< 2.2e-16	***	3.000	< 2.2e-16	***
Ingredient 2	0.071	0.2600		0.097	0.0710	.
Ingredient 3	-0.821	0.0015	**	2.074	< 2.2e-16	***
Ingredient 4	-1.231	< 2.2e-16	***	1.745	< 2.2e-16	***
Ingredient 5	-1.141	< 2.2e-16	***	1.193	< 2.2e-16	***
Ingredient 6	-0.496	< 2.2e-16	***	0.795	< 2.2e-16	***
Ingredient 7	-0.384	8.16E-05	***	-0.489	0.0001	***
Ingredient 8	-0.612	6.00E-15	***	0.521	< 2.2e-16	***
Ingredient 9	-0.508	9.64E-13	***	1.126	< 2.2e-16	***
Ingredient 10	-2.733	< 2.2e-16	***	7.123	< 2.2e-16	***
Ingredient 11	-8.393	1.88E-09	***	0.082	0.9547	
Ingredient 12	0.024	0.8965		0.576	0.0751	.
Ingredient 13	-3.576	1.69E-09	***	-0.498	0.2936	
Ingredient 14	-0.449	0.0052	**	1.125	0.2544	
Ingredient 15	-0.597	1.78E-15	***	2.342	< 2.2e-16	***

Ingredient 16	-0.494	2.62E-12	***	2.458	< 2.2e-16	***
z_2 (tasks 17 to 20)	-	-	-	3.603	< 2.2e-16	***
z_3 (tasks 21 to 24)	-	-	-	1.598	< 2.2e-16	***
z_4 (tasks 25 to 28)	-	-	-	15.501	1.47E-05	***
Significance codes: '***' 0.001, '**' 0.01, '*' 0.05, '.' 0.1						

Table 1

The MXL model performs very well, with all random parameters showing either significant mean or standard deviation. The variance of the indicator variables for the groups of tasks are highly significant. This means that in the three last stages of the choice experiment (tasks 17-20, 21-24 and 25-28) utility has a higher variance than in the initial stage (task 1-16). Respondent engagement shows erosion signs by tasks 17-20, as shown by a significant additional variance in utility (3.603). However, engagement recovers somewhat in tasks 21-24, with a lower additional variance (1.598), significantly different from the previous one. Finally, the additional variance of utility skyrockets to 15.501 in the last tasks: 25-28. This very interesting pattern reveals that engagement decline cannot be completely attributed to fatigue. Fatigue effect should inexorably increase the variance of utility as the experiment makes progress. However variance dropped in tasks 21-24, before going up again. A better behavioural explanation is boredom. The experiment included a brief reminder of the mission before each task. In order to keep the attention of the respondent throughout such a long experiment, a special message was placed right before task number 20, saying that the exercise was about to finish and that there were just a few tasks left. This message seems to have broken the boring climate created by 19 previous choice tasks, recovering the respondent engagement somewhat for the last part of the exercise. However, as tasks went on, and respondents found out that the "few tasks left" were not that few, they seem to became annoyed to the extent that the additional utility variance went up sharply.

4. CONCLUSIONS

We have explored the effects of task complexity and respondent engagement in choice experiments. We have selected the marketing domain. The experiment was designed with a large number of alternatives (8 to 16), a simple attribute specification (brand and ingredients only), and a large number of choice tasks (28). This enables investigating respondent engagement effects beyond the limits of most previous research.

We have found that task complexity, even if it may not impact the variance of utility, it does impact the precision of prediction. This is quite relevant for marketing applications of DCM, as results are mainly expressed in terms of choice probability. It would be interesting to model the relationship between the number of alternatives and the precision of prediction in further research. Several data points should be established for this, in the fashion of Caussade et al. (2005).

Consistently with both: Bradley and Daly (1994) and Hess et al. (2012), we have found signs of respondent engagement decline as the experiment made progress, beyond the scope of the latter. We could also identify that disengagement is mostly due to boredom rather than fatigue. This is good news, as boredom can be acted upon. A simple encouraging message can break it and recover engagement to some extent. Stimuli variation can possibly be employed to manage engagement level throughout the experiment. Great care must be taken, though, not to disappoint impatient individuals with overpromising messages, or their may disconnect and drop inconsistent data in the sample. We have identified this phenomenon in the experiment via increased utility variance, however it was not large enough to impact prediction precision within the scope of the present experiment. Further research may be designed to address the disentanglement of fatigue and boredom. Another area for further research is finding the right strategies to maintain respondent engagement throughout the experiment. Another interesting area would be measuring respondent impatience in order to offer the right number of choice tasks to each one and still collect high quality data.

REFERENCES

BECH M, KJAER T, LAURIDSEN J (2011). Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment. *Health Econ*. 20 (3), 273–286.

BHAT, C.R., CASTELAR, S., 2002. A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay area. *Transportation Research B* 36, 577–669.

BRADLEY M, DALY A (1994). Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation* 21 (2), 167–184.

BROUWER R, DEKKER T, WINDLE JRJ (2009). Choice certainty and consistency in repeated choice experiments. *Environ. Resour. Econ*. 46 (1), 93–109.

CAUSSADE S, ORTÚZAR JD, RIZZI L, HENSHER DA (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transp. Res. B* 39, 621–640.

HENSHER DA (2010). Attribute processing, heuristics, and preference construction in choice analysis. In: Hess, S., Daly, A. (Eds.), Choice Modelling: the State-of-the-Art and the State-of-Practice; Proceedings from the Inaugural International Choice Modelling Conference. Emerald, pp. 35–70 (Chapter 3).

HESS S, HENSHER DA, Daly A (2012). Not bored yet - Revisiting respondent fatigue in stated choice experiments. *Transportation Research part A*, 46: 626-644.

LOUVIERE J, HENSHER D, SWAIT J (2000): Stated choice methods: Analysis and applications. Cambridge Univ. Press, New York.

TRAIN KE (2009). Discrete choice methods with simulation. Cambridge Univ. Press.