
Munich Personal RePEc Archive

Measurement of Cannibalism Effects in buying experiments using Mixed Logit Models - The Example of a new Brand of the "Fruits of Lake Constance" Association -

Profeta, Adriano

2008

Online at https://mpra.ub.uni-muenchen.de/20542/
MPRA Paper No. 20542, posted 09 Feb 2010 09:09 UTC

DISCUSSION PAPER

TUM Business School
Technische Universität München

Measurement of Cannibalism Effects in buying experiments using Mixed Logit Models

- The Example of a new Brand of the "Fruits of Lake Constance" Association -

Adriano Profeta

Discussion Paper 02-2008

Environmental Economics
and Agricultural Policy Group

Contact address: Adriano Profeta
Technische Universität München
Lehrstuhl für Volkswirtschaftslehre –
Umweltökonomie und Agrarpolitik

Alte Akademie 14
85350 Freising-Weihenstephan

phone: +49.8161.71.3593
fax: +49.8161.71.3408
email: adriano.profeta@wzw.tum.de

Discussion Papers are edited by the Environmental Economics and Agricultural Policy Group. The responsibility for the content lies solely with the author(s). Comments and critique by readers of this series are highly appreciated.

The Environmental Economics and Agricultural Policy Group is integrated in the TUM Business School.

Address: Technische Universität München
Lehrstuhl für Volkswirtschaftslehre –
Umweltökonomie und Agrarpolitik

Alte Akademie 14
85350 Freising-Weihenstephan

phone: +49.8161.71.3406
fax: +49.8161.71.3408
homepage: www.wzw.tum.de/ap/

Measurement of Cannibalism Effects in buying experiments using Mixed Logit Models

- The Example of a New Brand of the "Fruits of Lake Constance" Association -

Summary

One fundamental assumption of discrete choice regression is the assumption of independence of irrelevant alternatives (IIA). According to the IIA assumption no correlation is allowed between brands in buying experiments. As a consequence, in market simulations all remaining brands gain at the ratio of their starting market share if one brand is excluded from the simulation set. This often does not reflect the reality at the point-of-sale. Mixed-logit models offer the advantage that the IIA-assumption is completely relaxed. What is more, simulations based on mixed logit are able to model cannibalism effects. This paper applies mixed logit to buying behaviour research. A case study is presented where the introduction of a new apple brand at the German discounter "Penny" is simulated in a buying experiment.

Keywords: Mixed Logit Model, IIA-assumption, cannibalism effect, choice experiment, apples

1 Introduction

Recently, logistic regression is increasingly applied in market research (see ENNEKING ET. AL., 2007; 2003; FRODE, 2004; LOUVRIERE, 2000; PROFETA, 2006). This tool offers the advantage to analyze discrete decisions used, for example, in buying experiments (ENNEKING, 2005).

One fundamental assumption of this quantitative method is the independence of irrelevant alternatives (IIA) (MCFADDEN, 1974: 109). According to IIA no correlation is allowed between brands in buying experiments. As a consequence, in market simulations all remaining brands gain at the ratio of their primary market share if one brand is excluded from the simulation. This often does not comply with the reality at the point-of-sale.

If a certain brand is taken out of the assortment, one can assume that some of the remaining brands do profit more in market share compared to others. Consider, for example, a case where consumers can choose between two luxury brands and one discount brand in a certain product category. It is expected that if one of the luxury brands is out of stock the discount brand will profit only marginally whereas the remaining luxury brand will account for most of the released market share. In this situation the IIA assumption is violated and a simple logit model should not be applied.

Nested logit models can deal with the violation of IIA. Nevertheless, it has to be mentioned that for several nests IIA still applies (BEN-AKIVA and LERMAN, 1985). Therefore, more sophisticated methods are needed and due to improved statistical software packages (e.g.

NLOGIT 3.0, STATA 7.0, LATENT GOLD CHOICE) latent-class and mixed logit models are increasingly used in modelling of choice behaviour. Especially mixed logit models offer the advantage that the IIA-assumption can be completely relaxed and therefore correlations between alternatives can be modelled. As a consequence market simulations can integrate cannibalism effects, which is important for marketers to derive useful marketing strategies.

In the course of this paper the practical use of mixed logit models in consumer research is shown. A case study is presented where the introduction of a new apple brand at the German discounter "Penny" is simulated in a buying experiment.

The "Fruits of Lake Constance" Association uses the slogan "FRUITS OF LAKE CONSTANCE, BATHED IN THE SUN"[1] as a means to advertise for apples from the respective region. The brand "Fruits of Lake Constance" is relatively well-known and yields a price premium in comparison to other origins or brands. In order to increase sales it is planned to launch an additional secondary brand. It is intended to gain market shares from a main Italian competitor[2].

The challenge to the "Fruits of Lake Constance" Association is to select one of three proposals which cannibalizes market shares mainly from the foreign competitor brand and only marginally from the own premium brand "FRUITS OF LAKE CONSTANCE". In order to find out the best alternative with respect to the aim mentioned above a discrete choice buying experiment is carried out containing the brand "FRUITS OF LAKE CONSTANCE", the foreign brand as well as the three alternatives for a secondary brand which have been developed by the association.

After a brief description of the theoretical basics of the mixed logit model, the method used for data analysis and the study design is described in more detail. Subsequently, the estimation results are presented and discussed and market simulations are calculated. The simulation results on the basis of conditional logit and mixed logit are compared with each other and marketing recommendations for the introduction of a secondary brand are given.

2 Description of the analytical method: Mixed Logit Model

For the purpose of this paper we assume according to FRODE (2004: 24) that an individual's utility from any alternative can be decomposed into a) a non-stochastic, linear-in-parameters part that depends on observed data, b) a stochastic part that is normally distributed and potentially correlated and heteroscedastic, and c) another stochastic part which is independently, equally distributed over alternatives and individuals. *"Given these assumptions, the utility that individual n derives from alternative i [among J alternatives] in choice situation t is denoted by*

[1] For the sake of simplicity the appendix "Bathed in the Sun" will be left out in the following.

[2] Due to confidentiality reasons the name of the Italian competitor brand is not indicated in this study. In the following it will be named "ITALIAN DOP".

(1) $U_{nit} = \beta_i' x_{nit} + [u_{ni} + e_{nit}]$

where x_{nit} is the vector of observed non-stochastic variables including socio-economic characteristics of individual n and attributes of alternative i in choice situation t, β_i is a vector of structural parameters, u_{ni} is an error term that is normally distributed over individuals and alternatives, and e_{nit} is an extreme value error term that is independently and identically distributed over individuals, alternatives and choices by the same individual" (FRODE, 2004: 24).

It is to mention that the specified model is a panel model, where the normally distributed error term for alternative *i* is the same for all choices made by one individual. Henceforth we suppress the subscript *t*.

The density of η is denoted by f(u|Ω) where Ω are the fixed parameters of the distribution. For a given u, the IIA property holds and the conditional choice probability takes a standard multinomial logit form[3]

(2) $L_{ni}(u) = \dfrac{\exp(\beta_i' x_{ni} + u_{ni})}{\sum_{j \varepsilon J} \exp(\beta_j' x_{nj} + u_{nj})}$

The unconditional choice probability, P, in the mixed logit model is the logit formula integrated over all values of u with the density of u as weights (IBID.).

(3) $P_{ni} = \int L_{ni}(u) f(u \mid \Omega) \mathrm{d}u$

This choice probability is approximated via simulation (BROWNSTONE and TRAIN, 1999). *The estimated share of consumers choosing alternative i in a market scenario, S_i, is computed as the average of the individual participant's simulated choice probabilities in the market scenarios:*

(4) $S_i = N^{-1} \sum_{n=1}^{N} \hat{P}_{ni}$

where N is the total number of participants and $\hat{P}_{ni}$ is the simulated probability that individual n will choose alternative i" (FRODE, 2004: 25).

3 Experimental Design and Concept of the Survey

The choice alternatives presented to respondents in the buying experiment were the brands "FRUITS OF LAKE CONSTANCE", "ITALIAN DOP" and three potential concepts for a secondary brand developed by the "Fruits of Lake Constance" Association. The new brands were named "MY FRUITS", "FRUIT DELIGHT" and "POP-ART". Real packaging ready for sale were used instead of cards or pictures in order to increase the reality of the experimental buying situa-

[3] URBAN, 1993.

tion. Therefore, all brands had brand-related colour designs (see Table 1). It has to be mentioned that there was no brand logo on the package of the alternative "POP-ART". This alternative had only a special colour design. The price levels for the experimental brands were 1.19 €/kg, 1.59 €/kg and 1.99 €/kg and represented real prices in the "Penny" outlets. Only brand/price-combinations the authors believed a priori to be strictly dominated by other alternatives (e.g. high priced secondary brand) were excluded from the choice set. Additionally, combinations of the foreign competitive brand and the brand "FRUITS OF LAKE CONSTANCE" with the lowest price levels were not included in the experiment because there are no such combinations in the "Penny" outlets. This resulted in a set of brand concepts which were allocated to 30 choice sets with three alternatives each using the software FEDOROV (compare Table 1).[4] Because the marketers were interested in cannibalism effects between the three new secondary brands there were no base or fixed alternatives in the choice set. Each participant was shown two different choice sets consisting of three apple boxes[5] and different price levels. The interviewees had to opt for their favoured brand taken into account a particular price. Table 2 gives an example of the choice questions.

Data was collected in outlets of the German discounter "Penny" in the cities of Neuhausen (n = 130) and Constance (n = 244) in the time from 28th June to 10th July 2005 (see Table 3). The interviews were conducted face-to-face[7] behind the cashpoint. In this way, interviewers contacted respondents who were in a shopping mood and the survey could at least partly mirror the point-of-sale context. Only those shoppers were asked to participate in the interview who at least occasionally consume apples.

[4] The software FEDOROV (FEDOROV, 1972) uses the FEDOROV-Algorithm to create optimized experimental designs.

[5] All apples were of the same type.

[7] Data were collected by four trained interviewers.

Table 1: Brand alternatives in the buying experiment

Brand	Package Design [8]	Price Levels
ITALIAN DOP		1.59 €/kg 1.99 €/kg
FRUITS OF LAKE CONSTANCE		1.59 €/kg 1.99 €/kg
MY FRUITS		1.19 €/kg 1.59 €/kg
FRUIT DE-LIGHT		1.19 €/kg 1.59 €/kg
POP-ART		1.19 €/kg 1.59 €/kg

[8] Due to confidentiality reasons the name of the foreign competitor brand (ITALIAN DOP) is garbled.

Table 2: Example of a Choice Set

		
Grade 1 1,59 €/kg	**Grade 1 1,59 €/kg**	**Grade 1 1,19 €/kg**

Question: "You see three different offers for apples from different sellers. Which of these apples would you buy at the displayed price?"

The interview started with the choice tasks in order to reduce the likelihood of bias potentially being introduced by other questions which might reveal the targets of the study. After the choice tasks, the participants made statements with respect to their general apple consumption habits and to their age and education level (see Table 3).

Table 3: Description of the sample (n=374)

Locality of interviews	- Constance (n = 244), Neuhausen (n = 130)
Gender distribution	- female = 66.31 %, male = 33.69 %
Age	- under 18 years = 2.97 % - 19 to 39 years = 20.81 % - 40 to 59 years= 18.92 % - over 59 years = 24.73 % - 12.84 % - 19.73 %
Education Level	- low = 28.15 % - middle = 32.44 % - high = 39.41 %

4. Results

Table 5 shows the estimated parameters, standard errors and p values for the mixed logit model. The corresponding results of the multinomial logit model, obtained by excluding the normally distributed error term, are included for comparison. In both models the alternative "POP-ART" serves as reference alternative which is set to zero for the purpose of estimation.

All those parameters of the mixed logit model which have a significant effect on choice have the same sign as the corresponding parameters in the multinomial logit model. The parameters are relative similar, but not identical. The most important difference between the two models is the error structure. Three out of four normally distributed residuals in the mixed logit model were heteroscedastic. This can be derived from the significant standard deviations of the non-i.i.d. residuals. The brand "MY FRUITS" (1.33) has the largest standard deviation followed by

"FRUIT DELIGHT" (1.22) and "ITALIAN DOP" (1.07).

The correlation error terms is reported in the lower part of table 5. The error terms of the secondary brands "MY FRUIT", "FRUIT DELIGHT" and the "ITALIAN DOP" are positively correlated, indicating that they are close substitutes. This result can not be derived from the non-stochastic part of the model. An estimation carried out with "FRUIT DELIGHT" as reference alternative showed that the brand "POP-ART" is positively correlated with "FRUITS OF LAKE CONSTANCE" and negatively correlated with the "ITALIAN DOP". Significant heteroscedastic and correlated error terms prove that the IIA property does not hold and that the error structure in the multinomial logit model is not adequate.

In order to rank the five alternatives, we modelled the same price for all alternatives and insert this price into the alternative specific utility function. By subtracting the utility of "POP-ART" from the utilities of the other alternatives we estimate the relative utility compared to the utility of "POP-ART". This results in zero utility for "POP-ART" and positive utilities for the other alternatives. The ranking results according to the average utilities can be specified as follows: "FRUITS OF LAKE CONSTANCE"(0.78), "ITALIAN DOP"(0.75), "FRUIT DELIGHT"(0.69), "MY FRUITS"(0.42) and "POP-ART"(0.00). This is equal to the parameter estimates because no individual specific data entered the estimation.

Market simulations were calculated for identifying cannibalism effects predicted by the estimated mixed logit and multinomial logit model. The individual participant's probability of choosing each of the alternatives under various market scenarios was estimated and summed up to obtain the simulated share of each alternative in the market (compare equation 5).

All in all three market scenarios were constructed. In these scenarios the price for "FRUITS OF LAKE CONSTANCE" and "ITALIAN DOP" was fixed at 1.99 €/kg whereas the other alternatives were set to 1.19 €/kg. In scenario 1 to 3 one of the new secondary brands was added to the present brands "FRUITS OF LAKE CONSTANCE" and "ITALIAN DOP". In this way these scenarios simulated the introduction of a new secondary brand. The estimated parameters of the mixed logit model and the corresponding multinomial logit model, allowed the calculation of market shares for the above specified scenarios. The results are presented in Table 6.

Table 5: Estimated parameters for the mixed logit and the multinomial logit model

	Multinomial Logit				*Mixed Logit*		
	Para-meter	Std. err	p value		Para-meter	Std. err	p value
LAKE CONSTANCE	0.67	0.15	0.00		0.78	0.17	0.00
FRUIT DELIGHT	0.59	0.16	0.00		0.69	0.20	0.00
MY FRUITS	0.35	0.16	0.03		0.42	0.20	0.03
ITALIAN DOP	0.68	0.18	0.00		0.75	0.21	0.00
Price	-0.91	0.21	0.00		-1.02	0.27	0.00
					St. Dev. of non-i.i.d. residuals		
				LAKE CON.	0.00	0.00	0.17
				FRUIT DEL.	1.22	0.00	0.00
				MY FRUITS	1.33	0.46	0.00
				I. DOP	1.07	1.81	0.05

Summary statistics		*Correlation matrix for non-i.i.d. residuals*			
Number of observations: 748		LAKE C.	FRUIT D.	MY F.	I. DOP
Number of participants: 374	LAKE C.	1.00	-0.71	-0.64	-0.84
Log likelihood random choice:	FRUIT D.	-0.71	1.00	0.99	0.86
Log likelih. of multinom. logit at converg.: -812.90168	MY F.	-0.64	0.99	1.00	0.87
Log likelihood of mixed logit at converg.: -775.4520	I. DOP	-0.84	0.86	0.87	1.00

Estimated with Nlogit 3.0 as panel data. Halton draws. Replications for simulated probability = 1000.

As mentioned above, the relative size of the parameters is not identical in the two models and the mixed logit model has heteroscedastic and correlated error terms. However, heteroscedastic and correlated error terms have negative effects on the market simulations, as will be shown subsequently.

In the case where taste parameters are small relative to the error variance, the stochastic part of the utility will dominate the non-stochastic part. Furthermore, the relative variance of the utility distributions also affects the simulated market shares. An alternative with a relatively low mean utility and a high variance can take on a large share of the choices as a result of the thick tails of the utility distribution. Additionally, the utility correlation between alternatives has an impact on the simulated market shares (FRODE, 2004: 34).

The results of the simulations show that "FRUITS OF LAKE CONSTANCE" loses market shares in all three scenarios of the mixed logit model compared to the multinomial logit model. In a similar study carried out by FRODE (2004) and based on a mixed model the strongest alternative was positively correlated with all other alternatives and gained market shares. How can these results be explained? Alternatives with positively correlated utilities compete for the same consumers. In the extreme case where the utilities of several alternatives have identical

mean and variance and are perfectly correlated, the joint share for these alternatives is independent of how many of these alternatives are offered in the market. This is known as cannibalisation of market shares between close substitutes. If, in a choice set of four alternatives, two brands are positively correlated and one of these correlated alternatives is excluded from a simulation than the other gains most of the released market share. The reverse is true if the two mentioned brands are negatively correlated.

What happens if a new brand enters a set and if this brand is negatively correlated to one of the former brands (this situation occurred in the present study)? The outcome depends on the "brands strength" which can be derived from the estimated alternative specific parameters. If the brand in the starting choice set is weaker than the new one this weak brand profits from a negative correlation because it loses less market share in comparison to the uncorrelated situation. The reverse holds if the new brand is weaker.

The simulation results in table 6 support these considerations. As already mentioned "FRUITS OF LAKE CONSTANCE," loses market shares in all three scenarios of the mixed logit model compared to the multinomial logit whereas the secondary brands and the "ITALIAN DOP" gain market shares.

In this study the alternative "MY FRUITS" is negatively correlated to the brand "FRUITS OF LAKE CONSTANCE" but in comparison to the other brands this correlation is rather weak. These results have the following consequences. "MY FRUITS" is negatively correlated to "FRUITS OF LAKE CONSTANCE" and thus profits in terms of market share. To the contrary "FRUITS OF LAKE CONSTANCE" loses according to the negative correlation. Nevertheless it has to be mentioned that this loss is still higher for the brand "FRUITS OF LAKE CONSTANCE" if the secondary brand "FRUIT DELIGHT" is introduced. Thus, "FRUITS OF LAKE CONSTANCE" should rather be marketed together with "MY FRUIT" than in tandem with the other secondary brand "FRUIT DELIGHT".

"POP-ART" is positively correlated (0.36) to "FRUITS OF LAKE CONSTANCE" and negatively correlated to the "ITALIAN DOP". This explains the small loss for "FRUITS OF LAKE CONSTANCE" in scenario three of the mixed model.

Whereas "MY FRUITS" shows disadvantages as a secondary brand this brand has the highest increase in market share (+1.51 %) when the mixed and the multinomial logit models are compared. This is due to the thick tails of the utility distribution. Because "MY FRUIT" is positively correlated to "ITALIAN DOP" it gains most of the market shares from this brand. Therefore "ITALIAN DOP" has the lowest increase in market share in scenario two of the mixed logit model.

In order to determine the effect of the best secondary brand we assumed that each year 10.000 t of apples are sold in "Penny"-outlets. The turnover of one alternative can thus be calculated

by multiplying the absolute market share and the corresponding price of the alternative. In order to determine the total share of the "Fruits of Lake Constance" Association the revenues of the brand "FRUITS OF LAKE CONSTANCE" and the secondary brand have to be added.
According to our calculation based on the mixed model "My FRUITS" (10.523 Mio. €) is the alternative that maximizes the turnover of the association.

Table 6: Simulated market shares

Multin. logit model	Scenario 1	Scenario 2	Scenario 3
LAKE CONSTANCE	25.47 %	28.51 %	32.58 %
ITALIAN DOP	25.78 %	28.84 %	32.96 %
FRUIT DELIGHT	**48.75 %**	0.00	0.00
MY FRUITS	0.00	42.65 %	0.00
POP-ART	0.00	0.00	34.46 %
Turnover in Mio. €	**10.870 (1)**	10.748 (2)	10.583 (3)
Mixed logit model	Scenario 1	Scenario 2	Scenario 3
LAKE CONSTANCE	21.91 % (-3.56%)	26.47 % (-2.04 %)	31.52 % (-1.06 %)
ITALIAN DOP	28.63 % (+2.85 %)	29.38 % (+0.54 %)	33.66 % (+0.70 %)
FRUIT DELIGHT	49.47 % (+0.72%)	0.00	0.00
MY FRUITS	0.00	**44.16 % (+1.51 %)**	0.00
POP-ART	0.00	0.00	34.82 % (+0.36)
Turnover in Mio. €	10.246 (3)	**10.523 (1)**	10.416 (2)

Prices in the simulation: "FRUITS OF LAKE CONSTANCE" = 1.99 €/kg, "ITALIAN DOP" = 1.99 €/kg Secondary Brands = 1.19 €/kg. Market Volume = 10,000 t.

5. Conclusions and Discussion

The mixed logit market simulations illustrate that two of the potential secondary brands compete with "ITALIAN DOP" for the same consumers. The more restrictive multinomial logit model does not detect the substitution patterns inherent in the choice data and predicts proportional changes in market shares. According to the mixed model results and the described market situation "MY FRUITS" can be recommended to the "Fruits of Lake Constance" Association as a secondary brand which optimizes the cumulated turnover of two brands.

In the past five years a multitude of multinomial logit analyses has been carried out in order to examine the preferences for (branded) food products. The authors heavily recommend to re-analyze the respective data sets using a mixed logit approach. In particular in the food sector it can be assumed that the IIA assumption does not hold very often. This could be verified by

means of a meta-analysis. If this is in fact the case many marketing recommendations relying on simple MNL-models might be flawed. As shown in this study the brands with the best alternative specific parameter in the mixed model is worst when simulated based on a mixed model.

An interesting field for future research relying on the mixed logit model is the integration of more situational factors as experimental variables. For example in the case where a well-known and strong brand is deleted from a retailer's assortment some consumers probably would not buy a cheaper alternatives or another secondary brand but would rather go to the retailer where the missing brand is still available. Integrating the variable "retailer" may help to analyze assortments of different retailer according to cannibalism effects.

Literature

BEN-AKIVA, M. and S.R. LERMAN (1985): Discrete Choice Analysis: Theory and Application to Travel Demand. Cambrigde, MA: MIT Press.

BROWNSTONE, D. and K. TRAIN (1999): Forecasting new product penetration with flexible substitution patterns. In: Journal of Econometrics 89: 109-129.

ENNEKING, U. (2003): Die Analyse von Lebensmittelpräferenzen mit Hilfe von Discrete-Choice-Modellen am Beispiel ökologisch produzierter Wurstwaren. In: Agrarwirtschaft 52 (5): 254-267.

ENNEKING, U. (2005): Präferenzen für Lebensmittelqualität aus mikroökonomischer und konsumentenpsychologischer Sicht, bisher unveröffentlichte Habilitationsschrift, Fakultät für Agrarwissenschaften der Universität Göttingen.

ENNEKING, U., C. NEUMANN ANDJJ, S. HENNEBERG (2007): How important intrinsic and extrinsic product attributes affect purchase decision. In: Food Quality and Preference (Volume 18): 133-138.

FEDOROV, V.V. (1972): Theory of Optimal Experiments. Translated and edited by STUDDEN, W.J. and E.M. KLIMKO. Academic Press, New York.

FRODE, A. (2004): Stated Preferences for imported and hormone-treated beef: application of a mixed logit model. In: European Journal of Agricultural Economics 31 (1): 19-37.

GREENE, W.H. (2002): Nlogit Version 3.0 Reference Guide. New York. Econometric Software, Inc.

LOUVIERE, J., D. HENSHER and J. SWAIT (2000): Stated Choice Methods. Analysis and Application. Cambridge: Cambridge University Press.

MCFADDEN, D. (1974): Conditional Logit Analysis of Qualitative Choice Behavior. In: ZA REMBKA, R. (ed.): Frontiers in Econometrics. New York: Academic Press: 105-142.

PROFETA, A. (2006): Der Einfluss geschützter Herkunftsangaben auf das Konsumenten-

vehalten bei Lebensmitteln - Eine Discrete-Choice-Analyse am Beispiel Bier und Rindfleisch. Studien zum Konsumentenverhalten, Bd. 7, Hamburg.

PROFETA, A. and U. ENNEKING (2004): Interactions of Protected Geographical Indications (PGI) with brand names: The case of "Bavarian Beer" and "Munich Beer". Unpublished Article. Technische Universität München, Weihenstephan.

URBAN, D. (1993): Logit-Analyse. Statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen. Gustav Fischer Verlag, Stuttgart.