Modeling multiple sources of heterogeneity in multinomial logit models: Methodological and managerial issues

Chang Liu

DSE Department, CUHK

liuchang@link.cuhk.edu.hk

February 24, 2017

Overview

Introduction

This paper offers a multinomial logit model with augmented heterogeneity specification integrates three sources of unobserved differences across households:

- random variations in intrinsic brand utilities (intercepts),
- random variations in response to marketing variables,
- loyalty.

Introduction

In addition to the augmented logit model, this paper directly estimates the following three nested models:

- a model with random intercept and parameters,
- a model with random intercepts only,
- a standard model without any heterogeneity specification.

Nonnested models:

- a multinomial logit model with the traditional loyalty specification,
- a parsimonious model that measures heterogeneity on intercepts and parameters.

The models are estimated on the scanner panel data on disposable diapers.

Random Intercept and Coefficients Specification

The consumer utility for brand j(= 1, 2, 3) on purchase occasion k is:

$$U_{jk} = \alpha_j + \beta_3 \cdot PR_{jk} + \beta_4 \cdot CU_{jk} + \varepsilon_{jk}$$

where:

- PR is price,
- CU is coupon,
- α_j is brand-specific intercept term,
- price and coupon parameters are common across brands.

Random Intercept and Coefficients Specification

Consider that households are differ in their **idiosyncratic brand preferences** and their **response to marketing effort**, for a specific household n, the utility for product j on occasion k is:

$$U_{jkn} = \alpha_j + V_{jn} + (\beta_3 + V_{3n}) \cdot PR_{jk} + (\beta_4 + V_{4n}) \cdot CU_{jk} + \varepsilon_{jk}$$

where:

- V_{jn} is a random component to each of the two intercept variations in intrinsic preferences.
- V_{3n} and V_{4n} are random components for the explanation variables, they are common across alternatives.
- the paper assume that the random components follow a normal distribution across the households.

Random Intercept and Coefficients Specification

The paper treat the distributions of four random components jointly as a **four-variate normal distribution**, and **allow the price and coupon random components to covary** to check if more price-sensitive households are more likely to take advantage of coupon offering.

The variance-covariance matrix for the four-variate normal distribution is

$$\begin{bmatrix} \sigma_1^2 & & & \\ 0 & \sigma_2^2 & & \\ 0 & 0 & \sigma_3^2 & \\ 0 & 0 & \sigma_{34} & \sigma_4^2 \end{bmatrix}$$

where the mean vector of the four-variate distribution is (0, 0, 0, 0).

Brand Loyalty Specification

Traditional loyalty specification is employing a simple proportion or an exponentially weighted average of past purchases as a measure of loyalty.

The paper do this heterogeneity specification by **identifying the hardcore loyal households** and then **separating them from the switching households**.

$$L_n = (L|Z = 0)(1 - H) + (L|Z = 1)(H)$$

where:

- Z is a discrete random variable such that Z = 1 indicates a hardcore loyal household and Z = 0 indicates a switching household,
- H is the probability that a household is hardcore loyal,
- L is the likelihood before incorporating the loyalty heterogeneity.

Brand Loyalty Specification

For any household, the likelihood expression is

$$L_n = (L(1-H)+H)^D \cdot (L(1-H))^{(1-D)}$$

where

- D = 1 if the household purchases the same brand during the entire observation period, and D = 0 otherwise.

Integration of the Heterogeneity Structures

Combining the multinomial logit with four-variate normal heterogeneity with specification for hardcore loyal households and others, the resulting likelihood expression for a particular household n is

$$L_n = \frac{1}{(2\pi)^2\sqrt{1-\rho^2}\cdot\sigma_1\cdot\sigma_2\cdot\sigma_3\cdot\sigma_4}$$

$$\cdot\left[\left\{\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\prod_{k=1}^{K_n}\prod_{j=1}^{J}\left[\frac{e^{\alpha_{jk}+V_{jn}+(\beta_3+V_{3n})\mathrm{PR}_{jk}+(\beta_4+V_{4n})\mathrm{CU}_{jk}+\epsilon_{jk}}}{\sum_{j=1}^{J}e^{\alpha_{jk}+V_{jn}+(\beta_3+V_{3n})\mathrm{PR}_{jk}+(\beta_4+V_{4n})\mathrm{CU}_{jk}+\epsilon_{jk}}}\right]^{P_{jkn}}\right.\right.$$

$$\cdot e^{-(V_{1n}^2/2\sigma_1^2)-(V_{2n}^2/2\sigma_2^2)-(1/2(1-\rho^2))[(V_{3n}^2/\sigma_3^2)-2\cdot\rho\cdot(V_{3n}\cdot V_{4n}/\sigma_3\cdot\sigma_4)+(V_{4n}^2/\sigma_n^2)]}$$

$$\left.\left.\cdot\, dV_{1n}\cdot dV_{2n}\cdot dV_{3n}\cdot dV_{4n}\cdot(1-H)\right]\right\}^{D}$$

$$\cdot\left\{\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\prod_{k=1}^{K_n}\prod_{j=1}^{J}\left[\frac{e^{\alpha_{jk}+V_{jn}+(\beta_3+V_{3n})\mathrm{PR}_{jk}+(\beta_4+V_{4n})\mathrm{CU}_{jk}+\epsilon_{jk}}}{\sum_{j=1}^{J}e^{\alpha_{jk}+V_{jn}+(\beta_3+V_{3n})\mathrm{PR}_{jk}+(\beta_4+V_{4n})\mathrm{CU}_{jk}+\epsilon_{jk}}}\right]^{P_{jkn}}\right.\right.$$

$$\cdot e^{-(V_{1n}^2/2\sigma_1^2)-(V_{2n}^2/2\sigma_2^2)-(1/2(1-\rho^2))[(V_{3n}^2/\sigma_3^2)-2\cdot\rho\cdot(V_{3n}\cdot V_{4n}/\sigma_3\cdot\sigma_4)+(V_{4n}^2/\sigma_4^2)]}$$

$$\left.\left.\cdot\, dV_{1n}\cdot dV_{2n}\cdot dV_{3n}\cdot dV_{4n}\cdot(1-H)\right]+H\right\}^{(1-D)}\Bigg],$$

Integration of the Heterogeneity Structures

(continued from previous page)

where

- $P_{jkn} = 1$ if brand j is purchased on occasion k by household n and $P_{jkn} = 0$ otherwise,
- p measures the correlation between the random price and coupon parameter distributions across households.

The likelihood expression across N households is

$$L = \prod_{n=1}^{N} L_n.$$

Comparison: Nested Models

A number of simpler heterogeneity specifications are nested within the model. The nested models the paper estimates are listed below.

- **Random intercepts and random parameters model.**
  By ignoring the heterogeneity due to distinction of the hardcore loyal households from others. ($\mathrm{H} = 0$)
- **Random intercepts model.**
  By restricting the parameters of the explanatory variables to remain fixed across households. ($V_{3n} = V_{4n} = 0$, $\rho = 0$)
- **Multinomial logit model with no heterogeneity.**
  By setting all the random components and the hardcore loyalty probability to zero. ($V_{jn} = 0$)

Comparison: Nonnested Models

To facilitate comparison with the augmented model, we propose to estimate two additional nonnested models.

- **Multinomial logit model with traditional loyalty specification.**
- **Parsimonious heterogeneity model.** A common additive component would capture the common heterogeneity in the parameters but would not measure any common heterogeneity in intercepts.

$$U_{jkn} = (\alpha_j + \beta_3 \cdot PR_{jk} + \beta_4 \cdot CU_{jk}) \cdot V_n + \varepsilon_{jk}.$$

The multiplicative specification imposes the **strong restriction** that the random deviation from the mean value is common to the intercepts and the parameters of the model. Only need to estimate the variance of only one component.

Data

The paper uses the UPC scanner data on disposable diapers collected by the A. C. Nielsen Company containing purchase histories of households over a 52 week period.

- The estimation and holdout samples consist of households that buy at least one of the three national brands(90% of the total market) of disposable diapers.
- The full sample consists of 152 households with a total of 2,675 purchases.
- Nearly half of all the purchases are made with a coupon. The most common value for the manufacturer's coupon is \$1.00 and does not vary by diaper size. Hence, coupon can be treated as a **dummy variable**.

Estimation

- **Estimate the augmented model and nested models**: Gauss Hermite quadrature.
- **Estimate the logit model with traditional loyalty**: The paper determines the exponential smoothing constant by searching for the value that maximizes the likelihood.
- **Disaggregate measure of the fit of a model**: Hauser's (1978) U^2 index. The measure is defined as

$$1 - (\ln L / \ln N)$$

where:

  - $\ln L$ is the model log-likelihood value,
  - $\ln N$ is the log-likelihood value of the null model.

The paper consider

  - U_0^2: the null model of equal probability of buying any of the three brands on each occasion.
  - U_1^2: a logit model with intercepts alone measuring the uncertainty explained by the share of brands.

Estimates of Augmented and Nested Models

Variables	Augmented Model		Random Intercept and Parameter Model		Random Intercept Model	
	Parameter	t-Ratio	Parameter	t-Ratio	Parameter	t-Ratio
Intercept A	Normalized		Normalized		Normalized	
Intercept B	−1.0775	−3.86	−1.8290	−6.29	−1.0003	−3.91
Intercept C	1.7057	6.21	2.0758	4.11	1.2497	8.26
Price	−0.4957	−3.07	−0.4048	−5.91	−0.2647	−7.54
Coupon	1.6052	6.99	1.4963	3.49	1.0594	3.38
S.D. (Intercept B)	1.7757	8.11	2.3802	2.02	1.8492	2.93
S.D. (Intercept C)	2.5221	9.96	3.1694	2.15	2.4076	4.34
S.D. (Price)	1.0566	5.02	1.2795	9.13	—	—
S.D. (Coupon)	1.2848	6.85	1.4814	3.63	—	—
Correlation Between Price and Coupon Distribution	−0.5611	−3.06	−0.4997	−5.02	—	—
Loyalty (Stayer) Probability	0.1864	4.81	—	—	—	—
Log Likelihood	−733.23		−765.35		−819.76	
Chi-Squared Statistic			64.24		108.82	
Degree of Freedom			1		3	
Time (minutes)	418		396		52	
U_0^2 (Equally Likely)	0.5559		0.5365		0.5035	
U_I^2 (Intercept)	0.4872		0.4648		0.4267	

Note: S.D. is the standard deviation equal to the square root of the variance of the random coefficients.

Results: Comparison of the Augmented Model with the Nested Models

TABLE 2

Performance on the Holdout Sample: Augmented and Nested Models

	Holdout Sample		Augmented Model	Random Intercept and Parameter Model	Random Intercept Model
	Brand	Share			
Market Share	A	0.2231	0.2167	0.2323	0.2367
	B	0.2066	0.2002	0.1740	0.2161
	C	0.5703	0.5831	0.5937	0.5472
Log Likelihood For Estimates in Table 1			−412.87	−425.99	−464.69
U_0^2 (Equally Likely)			0.4374	0.4195	0.3668
U_I^2 (Intercept only)			0.3698	0.3497	0.2907

Results: Comparison of the Augmented Model with the Nonnested Models

TABLE 3

Estimates of Nonnested Models

Variables	Multinomial Logit with Loyalty		Parsimonius Heterogeneity Model	
	Parameter	*t*-Ratio	Parameter	*t*-Ratio
Intercept A	Normalized		Normalized	
Intercept B	−0.3027	−2.79	0.3352	5.74
Intercept C	0.1860	2.01	0.7615	7.79
Price	−0.1814	−2.81	−0.0510	−2.40
Coupon	0.9191	8.52	0.0765	2.41
Loyalty	3.0928	28.14	—	—
Common Heterogeneity Component	—	—	4.88	6.04
Log Likelihood	−811.76		−934.18	
Time (minutes)	12		24	
U_0^2 (Equally Likely)	0.5076		0.4342	
U_I^2 (Intercept)	0.4323		0.3467	

Results: Comparison of the Augmented Model with the Nonnested Models

TABLE 4

Predictions on the Holdout Sample: Nonnested Models

	Holdout Sample		Multinomial Logit with Loyalty	Parisimonius Heterogeneity Model
	Brand	Share		
Market Share	A	0.2231	0.2321	0.3105
	B	0.2066	0.2174	0.1730
	C	0.5703	0.5515	0.5165
Log Likelihood For Estimates in Table 2			−449.22	−509.79
U_0^2 (Equally Likely)			0.3879	0.3054
U_I^2 (Intercept only)			0.3143	0.2218

Managerial Implications

- **Households' response to price and promotion.** The paper find that price-sensitive households respond to coupon more than other households.

  The augmented model suggests considerably higher price elasticities.

TABLE 6

Augmented Model: Impact of Explanatory Variables

	Price Elasticities		
	A	B	C
A	−3.48	1.04	1.00
B	1.66	−3.89	2.28
C	0.96	0.54	−1.66

Managerial Implications

TABLE 5

Comparison of Price Elasticities Across Models

	Brand	Price Elasticities		
		A	B	C
Random Intercept and Parameter Model	A	−2.68	0.66	0.65
	B	1.72	−3.71	2.02
	C	0.74	0.28	−1.17
Random Intercept Model	A	−0.88	0.30	0.31
	B	0.58	−1.06	0.47
	C	0.40	0.18	−0.59
Multinominal Logit with Loyalty	A	−2.02	0.46	1.56
	B	0.59	−2.17	1.56
	C	0.59	0.46	−1.02

Managerial Implications

Since coupon is a dummy variable, the total change in share due to coupon is presented below.

	Coupon Effect		
	A	B	C
A	0.0809	−0.0336	−0.0607
B	−0.0307	0.0880	−0.0547
C	−0.0502	−0.0544	0.1254

Note: Coupon is a dummy variable. The values reflect the total change in share of the row brand due to coupon offering of column brand.

The augmented model offers strong evidence for influence of coupon on choice.

Managerial Implications

- **Variations across households and directions for segmentation.** Households substantially differ in their response to price changes as the variance is high.
  A brand manager may generate the following hypotheses from focus groups as to why a household is relatively price insensitive:
  - higher income,
  - parents show greater concern for the first child,
  - light user of disposable diapers (e.g. only during travel),
  - do not worry about the brand of diaper when the children become older.
- **Brand loyalty.** The paper finds that high-income households are more likely to remain hardcore loyal. Identifying the hardcore loyal segment enables the brand manager to develop marketing strategies directed to the segment to encourage trial of the firm's brand first.

Research Issues

- **Computational issues.** As the number of choices and explanatory variables increases, computation becomes burdensome.
- **Comparison with the multinomial probit and computational issues.** An important limitation of the logit model in contrast to the probit model is the well-known independence from the irrelevant alternatives (IIA) problem. However, the multinomial probit model would involve a six-level numerical integration compared to four in the augmented logit model. Hence the probit model is more difficult to compute.

Conclusion

The main methodological and managerial implications of this model:

- The three measures of heterogeneity across households are all strongly supported.
- The average impact of the marketing efforts on brand choice is greater than the level implied by simpler models with limited heterogeneity specifications.
- Incorporation of important sources of heterogeneity reduces the potential bias in estimation of parameters.
- While different sources of heterogeneity exist, partial or parsimonious control of such sources may not be adequate.
- Guidelines on market segmentation.
- Answer several substantive issues on the interrelationships between a household's reactions to marketing inputs.

Future Extensions

- A more disaggregate specification of heterogeneity is possible. A logical extension is to allow the random component to differ by parameter and brand.
- Relax the definition of a seemingly loyal household. From a practical point of view, researchers may classify a household that purchases one brand on most of the occasions (for example, 98% of the total purchases) as a loyal household.
- Incorporate a more detailed loyalty specification in the random parameters model.

Strengths and weakness

Strengths:

- Compare the augmented model with several nested models to test the significance of each source of heterogeneity.
- Use nonnested models to facilitate comparison with the augmented model.
- Use holdout samples and validation samples to test performance of each model.
- Provide many managerial implications.

Weakness:

- Lack in details of nonparametric estimation.
- Lack of testing this model in other dataset or other brands.
- Computational issues.

Possible future extensions

- Consider multiple sources of heterogeneity in both logit and probit models.
- Different approaches in customer segmentation.
- Calculate and analyze the elasticities within each customer segment.
- If the number of brands becomes larger, it is also possible to specify different groups of households and add a random component to each of the group intercept.
- Take state dependence into account.

Thank you!