Discrete choice models, which one performs better?

R. Karina Gallardo and Jaebong Chang
Gallardo is an assistant scientist and extension specialist in the School of Economic Sciences, Tree Fruit Research and Extension Center, Washington State University, U.S., and Chang is a researcher at the Korea Rural Economic Institute, Seoul, Korea.
Contact author: R. Karina Gallardo, School of Economic Sciences, Tree Fruit Research and Extension Center, karina_gallardo@wsu.edu, (509) 663-8181 x. 261.

Poster prepared for presentation at the Agricultural & Applied Economics Association's 2010 AAEA, CAES & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010.

Discrete choice models, which one performs better?

Karina Gallardo and Jaebong Chang

Washington State University and Korea Rural Economic Institute

Selected poster 11543

Poster prepared for presentation at the Agricultural & Applied Economics Association's 2010

AAEA, CAES & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010

Korean Rural Economic Institute

Introduction

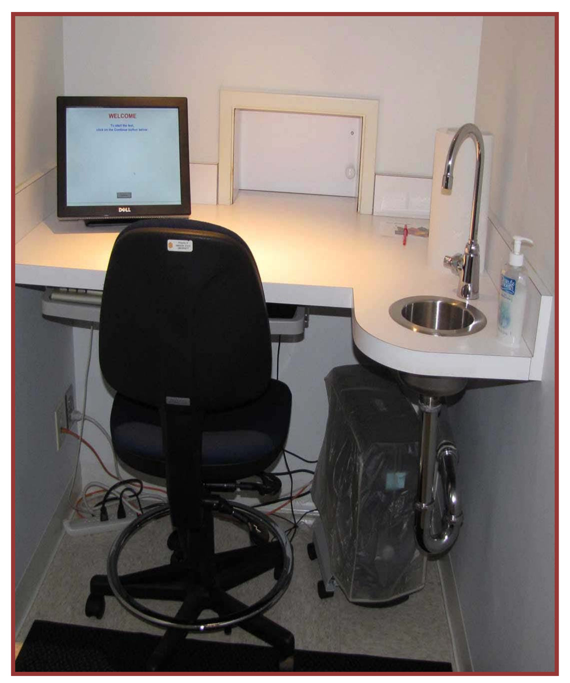
- With the developments of alternatives to the standard multinomial logit model (MNL), an increasing number of studies are focused testing improvement in predictability between competing discrete choice models and the standard MNL.
- Competing discrete choice models mostly generalize preference heterogeneity. One increasingly popular is the random parameter logit (RPL) model. This model relaxes the IID and IIA conditions leading to a flexible specification and behavioral richness. RPL's open form solution requires simulations to evaluate the likelihood function not ensuring a globally optimal estimate set.
- Another extension of the MNL is the error component multinomial logit (ECMNL) model. This specification is more straightforward than RPL and includes an additional error term to the utility function to capture unobserved individual specific random effects.
- The relative performance of discrete choice econometric models has been investigated based on in-sample statistics and out-of-sample criteria. However, it is well known that as more complexity is added to a model, the better the model will fit the data in-sample, while the contrary tends to be true out-of-sample. This suggests the need to incorporate both in-and out-of-sample criteria to compare the reliability and validity of advanced discrete choice models.

Objective

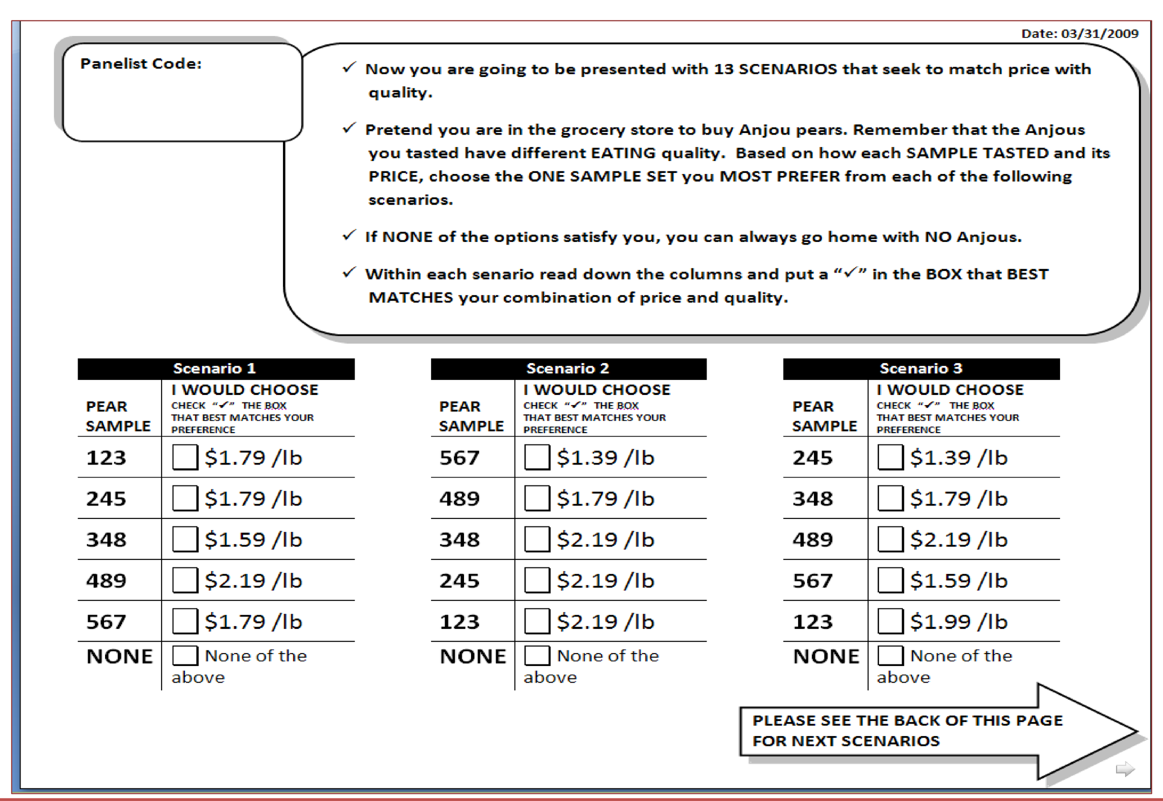
- To compare the performance of three discrete choice models - the MNL, the RPL, and the ECMNL, measured in terms of WTP valuations, market share estimates and the prediction success index within sample. Moreover, this study compares the models' ability to predict holdout sample choices.

Data

- We utilized response datasets from **two** choice experiments on preferences for fresh pears under different ripening treatments. The experiments were part of sensory tests conducted in December 2008 and March 2009, at the Food Innovation Center, Oregon State University in Portland.



- During both sensory tests, participants were asked to taste pears under different treatments and to answer a questionnaire. Ripening treatments in December and March were different, given differences in time length in cold storage and fruit maturity. Having tasted the pears, respondents were asked to answer choice experiment questions where they indicated which sample, linked to a randomly assigned price, was the most preferred. Choice experiment scenarios also included a "none" option.



Date: 03/31/2009

Panelist Code:

✓ Now you are going to be presented with 13 SCENARIOS that seek to match price with quality.

✓ Pretend you are in the grocery store to buy Anjou pears. Remember that the Anjous you tasted have different EATING quality. Based on how each SAMPLE TASTED and its PRICE, choose the ONE SAMPLE SET you MOST PREFER from each of the following scenarios.

✓ If NONE of the options satisfy you, you can always go home with NO Anjous.

✓ Within each senario read down the columns and put a "✓" in the BOX that BEST MATCHES your combination of price and quality.

Scenario 1	
PEAR SAMPLE	I WOULD CHOOSE CHECK "✓" THE BOX THAT BEST MATCHES YOUR PREFERENCE
123	☐ $1.79 /lb
245	☐ $1.79 /lb
348	☐ $1.59 /lb
489	☐ $2.19 /lb
567	☐ $1.79 /lb
NONE	☐ None of the above

Scenario 2	
PEAR SAMPLE	I WOULD CHOOSE CHECK "✓" THE BOX THAT BEST MATCHES YOUR PREFERENCE
567	☐ $1.39 /lb
489	☐ $1.79 /lb
348	☐ $2.19 /lb
245	☐ $2.19 /lb
123	☐ $2.19 /lb
NONE	☐ None of the above

Scenario 3	
PEAR SAMPLE	I WOULD CHOOSE CHECK "✓" THE BOX THAT BEST MATCHES YOUR PREFERENCE
245	☐ $1.39 /lb
348	☐ $1.79 /lb
489	☐ $2.19 /lb
567	☐ $1.59 /lb
123	☐ $1.99 /lb
NONE	☐ None of the above

PLEASE SEE THE BACK OF THIS PAGE FOR NEXT SCENARIOS

The Models

- **Multinomial Logit (MNL) Model**

  A random utility function for consumer i choosing option j is defined by

$$U_{ij} = \alpha_j + \beta P_{ij} + \epsilon_{ij}$$

where α_j is the estimated constant parameter for ripening treatment j, β is the marginal utility of price, and P_{ij} is the price. When assuming the stochastic term (ϵ_{ij}) is IID with type I extreme value, the choice probability of an individual i choosing alternative j out of a set J is expressed as

$$P_{ij} = \frac{exp(\alpha_j + \beta P_{ij})}{\sum_{k=1}^{J} exp\,(\alpha_k + \beta P_{ik})}$$

- **Random Parameter Logit (RPL) Model**

  In this application, the alternative-specific constant α_j terms are assumed as variant parameters across individuals and expressed as:

$$\alpha_{ij} = \bar{\alpha}_j + \sigma_j \omega_{ij}$$

where α_j is the mean alternative-specific constant for alternative j, σ_j is the standard deviation of the distribution of α_{ij} ,and ω_{ij} is a normally distributed random disturbance. The probability that individual i choose alternative j is represented by:

$$P_{ij} = \int \frac{exp(\alpha_{ij} + \beta P_{ij})}{\sum_{k=1}^{J} exp\,(\alpha_{ik} + \beta P_{ik})} f(\alpha_{ij}) d\alpha_{ij}$$

where $f(\alpha_{ij})$ is the density function.

- **Error components (ECMNL) model**

  In this model , the unobserved portion of utility is comprised by several components introducing more parsimonious distributions across random factors allowing flexible substitution patterns and correlation across alternatives . The ECMNL model is specified as,

$$U_{ij} = \alpha_j + \beta P_{ij} + \theta_j \gamma_{ij} + \epsilon_{ij}$$

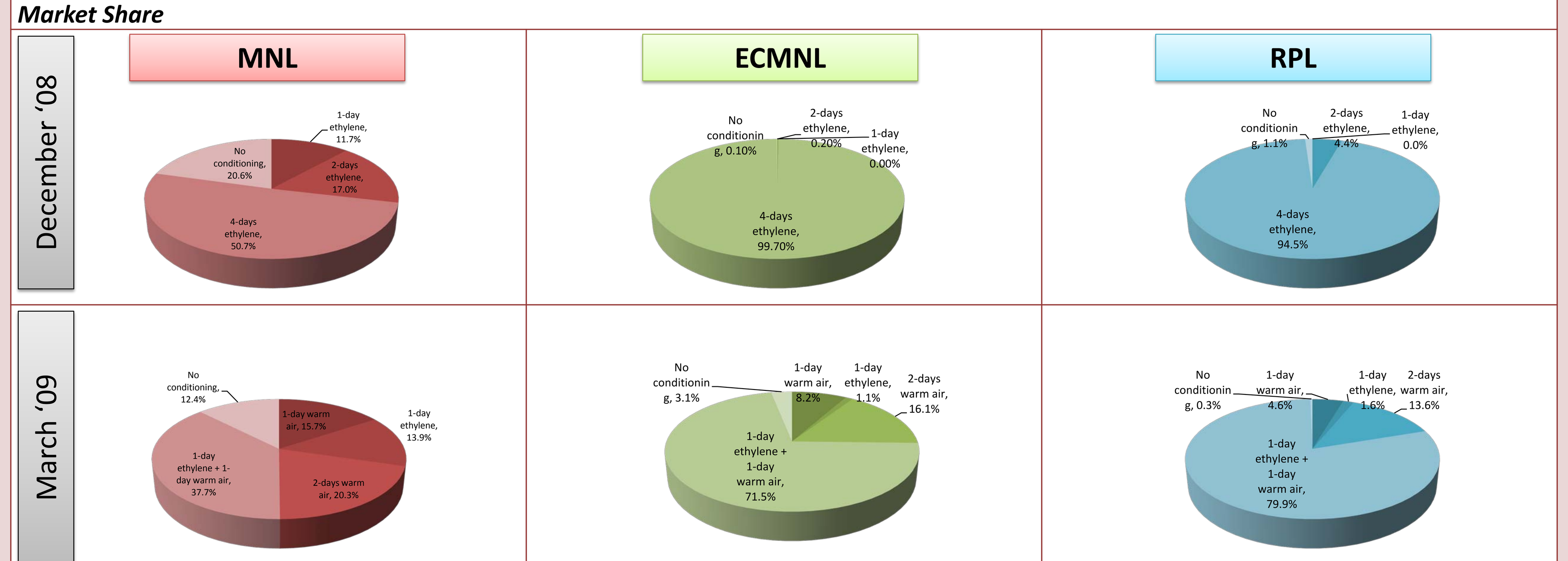
where γ_{ij} is a alternative-specific random error component which is distributed normally with zero mean and standard deviation one and θj is the standard deviation of the error component.

Results - Willingness to Pay and Market Share

	December 2008			March 2009		
	MNL	ECMNL	RPL	MNL	ECMNL	RPL
Willingness-to-pay						
1-day ethylene	$2.06 (0.08) [a]	$0.97 (0.07)	$1.56 (0.08)	-	-	-
2-days ethylene	$2.18 (0.08)	$1.57 (0.04)	$1.92 (0.04)	-	-	-
4-days ethylene	$2.53 (0.10)	$2.01 (0.03)	$2.14 (0.04)	-	-	-
No conditioning	$2.24 (0.08)	$1.56 (0.06)	$1.83 (0.04)	-	-	-
1-day warm air	-	-	-	$1.92 (0.06)	$1.66 (0.06)	$1.57 (0.07)
1-day ethylene	-	-	-	$1.88 (0.06)	$1.42 (0.04)	$1.48 (0.18)
2-days warm air	-	-	-	$2.01 (0.06)	$1.73 (0.03)	$1.67 (0.05)
1-day ethylene + 1-day warm air	-	-	-	$2.23 (0.06)	$1.90 (0.04)	$1.84 (0.05)
No conditioning	-	-	-	$1.84 (0.06)	$1.54 (0.04)	$1.33 (0.07)

[a] Numbers in parentheses are standard errors obtained via parametric bootstrapping.

Market Share



Market share estimates calculated by assuming price is at $1.50/lb for all options.

Results

Prediction success index from within sample and prediction tests from twenty models over random hold-out samples

Data Set	Model		
	MNL	ECMNL	RPL
Log likelihood			
December 2008	-1049.84	-494.51	-497.59
March 2009	-1039.58	-547.26	-522.34
Overall prediction success index from within sample			
December 2008	0.038	0.042	0.030
March 2009	0.093	0.053	0.090
Prediction test from twenty models over random samples and hold-out samples			
Mean rank			
December 2008	1.550	2.050	1.900
March 2009	1.750	2.200	1.700
Average percentage correctly predicted			
December 2008	36.920 [22.36, 51.17] [a]	35.120 [21.96, 51.17]	36.240 [21.96, 48.96]
March 2009	31.250 [21.15, 42.82]	30.230 [19.08, 42.54]	31.660 [21.15, 45.11]

[a] Numbers in brackets are minimum and maximum average percentages of the number of correctly predicted choice sets over the hold-out samples.

Discussion

- Three contrasting findings across datasets. First, likelihood values signal greater explanatory power to ECMNL for the December dataset and to RPL for the March dataset. Second, prediction success indexes shows that for the December dataset ECMNL outperforms, while for the March dataset MNL is superior to the other two models. Third, holdout samples tests reveal superior prediction ability for MNL in the December dataset but for the March dataset it is RPL the model with the highest prediction ability.
- An explanation for the differences across datasets is that product attributes influence model performance. Different treatments led to different eating quality characteristics that were perceived by consumers. In the December trial, participants were more homogeneous in their preferences for each treatment than in March. Indeed, in December, 50 percent of respondents agreed in that their preferred sample was treatment 4 days ethylene. Whereas, a wider range of preferences is observed in March, 32 percent for 1 day ethylene plus 1 day warm air and 30 percent for 2 days warm air. We hypothesize that these differences in the distribution of preferences explains the differences in prediction ability across datasets. These claims agree with Train (1998) and Greene and Hensher (2003) who concluded that context, datasets and behavioral assumptions affect RPL superiority to MNL.

Conclusions

- Our results show that ECMNL outperformed RPL and MNL when the products being tested exhibited heterogeneous quality characteristics quickly perceived by respondents. Whereas when differences were not easily perceived, RPL outperformed MNL and RPL. Interestingly, MNL outperformed for the holdout sample prediction when using the December dataset and exhibited a higher prediction success index than RPL and ECMNL when using the March dataset. This result supports the claim in Chang et al. (2009) that more parsimonious models often exhibit a greater predictive ability. Overall, findings in this study raise similar issues to Train (1998) and Green and Hensher (2003) in that further studies controlling for context and dataset nature are needed since they are determinant for measuring the predictive performance of models more flexible than MNL.

Related Studies

- Chang, J. B., Lusk, J. L. and Norwood, F. B. (2009) How closely do hypothetical surveys and laboratory experiments predict field behavior? American Journal of Agricultural Economics, 91, 518-34.
- Greene, W., and Hensher, D. (2003) A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B, 37, 681-98.
- Revelt, D. and Train, K. (1998) Mixed logit with repeated choices: Households' Choices of Appliance Efficiency Level. The Review of Economics and Statistics, 80, 647-57.
- Train, K.(1998) Recreation demand models with taste differences over people. Land Economics, 74, 230-39.