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.


Copyright 2010 by [author(s)]. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided this copyright notice appears on all such copies.
Discrete choice models, which one performs better?

Karina Gallardo and Jaebong Chang
Washington State University and Korea 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.
- Comparing discrete choice models mostly generalize preference heterogeneity. One increasingly popular strategy is the random parameter logit (RPL) model. This model relaxes the IID and TIA assumptions leading to a flexible specification and behavioral richness. RPL's open form solution requires simulations to evaluate the likelihood function and 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.

The Models

- **Multinomial Logit (MNL) Model**
  - A random utility function for consumer i choosing option j is defined by
  \[ U_{ij} = \beta_j x_{ij} + e_{ij} \]
  where \( x_{ij} \) is the estimated constant parameter for option j, \( \beta_j \) is the marginal utility of price, and \( e_{ij} \) is the error term. When assuming the stochastic term \( e_{ij} \) is IID with type i extreme value, the choice probability of an individual i choosing alternative j out of a set is defined as
  \[ P_{ij} = \frac{e^{\beta_j x_{ij}}}{\sum_j e^{\beta_j x_{ij}}} \]

- **Random Parameter Logit (RPL) Model**
  - In this application, in heterogeneous random utility \( e_{ij} \) and \( \beta_j \) are where \( i = \text{the mean alternative-specific constant for alternative } j \) and \( \sigma_i \) is the standard deviation of the distribution of \( i \) and \( \sigma_j \) is a normally distributed random disturbance. The probability that individual i choose alternative j is represented by
  \[ P_{ij} = \frac{e^{\beta_j x_{ij} + \sigma_j \varepsilon_j}}{\sum_j e^{\beta_j x_{ij} + \sigma_j \varepsilon_j}} \]

- **Error components (ECMNL) model**
  - In this model, the error component of utility is comprised by several components introducing more parsimonious distributions across random factors allowing for more flexible substitutions and correlation across alternatives. The ECMNL model is specified as
  \[ U_{ij} = \beta_j x_{ij} + \sigma_j \varepsilon_j + \varepsilon_{ij} \]
  where \( \varepsilon_{ij} \) is an alternative-specific random error component which is distributed normally with zero mean and standard deviation one and \( \sigma_j \) is the standard deviation of the error component.

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 across studies. 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.

Results

**Results - Willingness to Pay and Market Share**

### Willingness to Pay

- **December 2008**
  - MNL: 1.54, 2.05, 1.90
  - ECMNL: 1.94, 2.50, 2.18
  - RPL: 1.50, 1.98, 1.50
- **March 2009**
  - MNL: 1.88, 2.23, 2.14
  - ECMNL: 1.66, 1.48, 1.83
  - RPL: 1.33, 1.42, 1.75

**Market Share**

### December 2008

- **MNL**
  - No conditioning: 36.90, 31.60, 31.60
  - 1-day ethylene: 33.45, 31.50, 31.50
  - 2-days warm air: 36.00, 32.00, 30.50
- **ECMNL**
  - No conditioning: 36.24, 36.36, 36.82
  - 1-day ethylene: 33.45, 31.50, 31.50
  - 2-days warm air: 36.00, 32.00, 30.50
- **RPL**
  - No conditioning: 33.45, 31.50, 31.50
  - 1-day ethylene: 33.45, 31.50, 31.50
  - 2-days warm air: 36.00, 32.00, 30.50

### March 2009

- **MNL**
  - No conditioning: 29.20, 27.00, 27.00
  - 1-day ethylene: 33.45, 31.50, 31.50
  - 2-days warm air: 36.00, 32.00, 30.50
- **ECMNL**
  - No conditioning: 28.45, 26.50, 26.50
  - 1-day ethylene: 33.45, 31.50, 31.50
  - 2-days warm air: 36.00, 32.00, 30.50
- **RPL**
  - No conditioning: 33.45, 31.50, 31.50
  - 1-day ethylene: 33.45, 31.50, 31.50
  - 2-days warm air: 36.00, 32.00, 30.50

Discussion

- Three contrasting findings across datasets. First, skillfulness values signal greater explanatory power to ECMNL for the December dataset and to RPL for the March dataset. Second, prediction success index shows that for the December dataset ECMNL outperforms, while for the March dataset MNL is superior to the two other models. Third, holdout samples test 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 treatment than in March. 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 that their preferred treatment was 4 days ethylene. Whereas, a wider range of preferences is observed in March, 32 percent for 1 day ethylene plus 2 day warm air and 30 percent for 2 day 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 Hessher (2003) who concluded that context, datasets and behavioral assumptions affect RPL superiority to MNL.

Conclusions

- Our results show that ECMNL outperforms RPL and MNL when the products being tested exhibit 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 implications for issues in Train (1998) and Green and Hessher (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