Conjoint Rating, Ranking and Choice: Selected Theoretical and Empirical Pitfalls

Jan Brůha

Charles University, Prague

Economic Valuation of Recreation Functions of Forest Ecosystems; December 8, 2006, Prague
Motivation

Pitfalls on Empirical Issues of Conjoint Choice, Rating and Ranking

Monte Carlo Studies

Future Research
Conjoint Choice, Ranking and Rating are widely used. Their theoretical backgrounds and empirical implementation have to be addressed.
Conjoint Choice, Ranking and Rating are widely used. Their theoretical backgrounds and empirical implementation have to be addressed.

This presentation tries to shed some (partially, new) light on the issue. This lecture concentrates on the two issues:

**Efficient Estimation** of conjoint-study data;

**Variable Selection Problem** in conjoint studies.
Multivariate LDV models

It becomes recognized that random-effect / random-parameters multivariate LDV models are theoretically correct approaches to estimation.
Efficient Estimation

Multivariate LDV models

It becomes recognized that random-effect / random-parameters multivariate LDV models are theoretically correct approaches to estimation.

Efficient estimation

Efficient estimation of MLDV models require an evaluations of multivariate integrals. What are the possibilities?
Variable Selection Problem

- There is a wide consensus about possibilities of the VSP in the linear framework;
- less pieces of knowledge are available to the LDV setting.
Variable Selection Problem

- There is a wide consensus about possibilities of the VSP in the linear framework;
- less pieces of knowledge are available to the LDV setting.

Our contribution

We evaluate selected methods for VSP in the conjoint setting using a Monte Carlo study.
Limitations of the Study

The presentation has a limited scope only

1. We do not consider behavioral biases;
The presentation has a limited scope only

1. We do not consider behavioral biases; thus we evaluate procedures as if the usual neoclassical assumptions hold.
The presentation has a limited scope only

1. We do not consider behavioral biases; thus we evaluate procedures as if the usual neoclassical assumptions hold.
2. This is by no means a complete survey of the field; rather it contributes to highlight selected issues.
The Quest

Are Contingent Ratings’ Results Identified under the Neoclassical Paradigm?
The Case for MLDV Models - I

The Quest
Are Contingent Ratings’ Results Identified under the Neoclassical Paradigm?

Answer
It depends on the estimation procedure used:

- **OLS-based** approaches (such as linear SUR or Tobit models) preclude identification of the WTP;
- **Ordered Probit** enables it.
The Case for MLDV Models - II

Even if the Ordered Probit enables the Conjoint Rating Identification, there are practical problems, such as

- How to deal with ties?
- Behavioral biases.
Even if the Ordered Probit enables the Conjoint Rating Identification, there are practical problems, such as
- How to deal with ties?
- Behavioral biases.

Therefore Conjoint Ranking or Choice might be preferable strategy.
The Case for MLDV Models - II

Even if the Ordered Probit enables the Conjoint Rating Identification, there are practical problems, such as

- How to deal with ties?
- Behavioral biases.

Therefore Conjoint Ranking or Choice might be a preferable strategy.

**Note** that even under the alternatives, responders answer several questions, thus a random-effect LDV model might be a possible choice.
An Example - Random-Effect Probit Model

Model formulation

\[ f(y^{(n)}) = \prod_{k=1}^{K} \phi(y_{k}^{(n)}|X_{kn}, c_n, \beta), \]
An Example - Random-Effect Probit Model

Model formulation

\[ f(y^{(n)}) = \prod_{k=1}^{K} \varphi(y_{kn}^{(n)} | X_{kn}, c_n, \beta), \]

\[ L(y^{(n)}|X_n, \beta) = \int_{-\infty}^{\infty} \prod_{k=1}^{K} \varphi(y_{kn}^{(n)} | X_{kn}, c_n, \beta) \frac{\phi(c_n/\sigma_n)}{\sigma_c} \, dc, \]
An Example - Random-Effect Probit Model

Model formulation

\[
f(y^{(n)}) = \prod_{k=1}^{K} \varphi(y_k^{(n)}|X_{kn}, c_n, \beta),
\]

\[
L(y^{(n)}|X_n, \beta) = \int_{-\infty}^{\infty} \prod_{k=1}^{K} \varphi(y_k^{(n)}|X_{kn}, c_n, \beta) \frac{\phi(c_n/\sigma_n)}{\sigma_c} dc,
\]

\[
L(y|X, \beta) = \prod_{n} \int_{-\infty}^{\infty} \prod_{k=1}^{K} \varphi(y_k^{(n)}|X_{kn}, c_n, \beta) \frac{\phi(c_n/\sigma_n)}{\sigma_c} dc,
\]
How to approximate the Integral?

A possible approximation:

\[ \int_{-\infty}^{\infty} \prod_{i=1}^{l} \varphi(y_i^{(n)}|X_n, c_n, \beta) \frac{\phi(c_n/\sigma_n)}{\sigma_c} \, dc \approx \]

\[ \sum_{c_n} \prod_{i=1}^{l} \varphi(y_i^{(n)}|X_n, c_n, \beta) w(c_n). \]
How to approximate the Integral?

A possible approximation:

\[
\int_{-\infty}^{\infty} \prod_{i=1}^{l} \varphi(y_i^{(n)}|X_n, c_n, \beta) \frac{\phi(c_n/\sigma_n)}{\sigma_c} \, dc \quad \approx \quad (1)
\]

\[
\approx \sum_{c_n} \prod_{i=1}^{l} \varphi(y_i^{(n)}|X_n, c_n, \beta) w(c_n).
\]

How to generate \( c_n, w(c_n) \)?
- quadrature-base rules,
- pseudo-Monte Carlo integration (pseudo random sequence)
- quasi-Monte Carlo integration (low discrepancy methods)
How to approximate the Integral?

A possible approximation:

\[
\int_{-\infty}^{\infty} \prod_{i=1}^{l} \varphi(y_{i}^{(n)}|X_n, c_n, \beta) \frac{\phi(c_n/\sigma_n)}{\sigma_c} \, dc \approx \sum_{c_n} \prod_{i=1}^{l} \varphi(y_{i}^{(n)}|X_n, c_n, \beta) w(c_n).
\]

How to generate \(c_n, w(c_n)\)?

- quadrature-base rules,
- pseudo-Monte Carlo integration (pseudo random sequence)
- quasi-Monte Carlo integration (low discrepancy methods)
Design of the Monte Carlo Study I

**Design**

We use Lusk (2002) AJAE model extended by the random effect parameter.

We compare three methods of approximations of (1):

1. pseudo MC based on Halton sequences;
2. quasi Monte Carlo;
3. quasi Monte Carlo (McFadden, 1989 simulator).
Accuracy of Numerical Integration (N = 100; A = 5)

Elapsed Time (N = 100; A = 5)

WTP Estimation (N = 100; A = 5)
Results of the Monte Carlo Study I

Findings - I

Contrary to Train (1998) and (1999) we find that the qMC methods does not seem to outperform the pMC approaches significantly (both in the expected accuracy and the approximation variance).
Results of the Monte Carlo Study I

Findings - I

Contrary to Train (1998) and (1999) we find that the qMC methods does not seem to outperform the pMC approaches significantly (both in the expected accuracy and the approximation variance).

A Conjecture

The product rule in (1) drives down the advantage of low discrepancy methods.
Results of the Monte Carlo Study I

Findings - I
Contrary to Train (1998) and (1999) we find that the qMC methods do not seem to outperform the pMC approaches significantly (both in the expected accuracy and the approximation variance).

A Conjecture
The product rule in (1) drives down the advantage of low discrepancy methods.

Future research
To try qMC methods other than low-discrepancy approaches (such as Weyl or Haber sequences).
We confirm that the McFadden (1989) approach to simulated maximum likelihood outperforms the pMC integration with identical draws over individuals.
Conclusion I

For a large class of MLDV, the conventional pMC is a sufficient tool and there seems to be a little efficiency using more elaborated algorithm.

- This holds especially if the pMC integration is done efficiently;
- on the other hand, the only costs of the qMC are fixe-costs of learning by the researcher.
A Never-ending Battle with Model Selection

The Quest

How to Select Correct Regressors to a Model
A Never-ending Battle with Model Selection

The Quest
How to Select Correct Regressors to a Model

- The purpose matter: forecasting versus structural analysis
- A neglected issue in non-linear models (because of complexity?)
Design of the Monte Carlo Study II

Design

We use again Lusk (2002) AJAE model extended by ‘false’ variables correlated with the true ones.
Design of the Monte Carlo Study II

Design

We use again Lusk (2002) AJAE model extended by ‘false’ variables correlated with the true ones.

The procedures are programmed in GOAT TOOLBOX™ programmed by me.
Approach of this study

It is hard to evaluate model selection procedures analytically: we use the Monte Carlo approach.
Approach of this study

It is hard to evaluate model selection procedures analytically: we use the Monte Carlo approach.

We use the four methods

- Penalized Likelihood (AIC and BIC penalties)
- Cross Validation Approaches (leave one CV\(_1\) and leave 25% CV\(_{0.25}\) )
Results of the Monte Carlo Study - II

**Results**

Penalized Likelihoods are too conservatives: they choose larger models than necessary.

This is surprising for the BIC penalty.
Results of the Monte Carlo Study - II

**Results**

Penalized Likelihoods are too conservatives: they choose larger models than necessary.

This is surprising for the BIC penalty.

**Results**

The CV$_{0.25}$ outperforms CV$_1$.

The results confirm Shao (1996, 1997) analyses for linear systems.
AIC-based penalty

BIC-based penalty

CV_{0.25%}

CV_{1}
The qMC applications usually use the low-discrepancy methods, such as Halton sequences, to approximate integrals in MLDV likelihoods.
Planned future research - MLDV models

The qMC applications usually use the low-discrepancy methods, such as Halton sequences, to approximate integrals in MLDV likelihoods.

Future research

To try quadrature and qMC methods other than low-discrepancy approaches.
Future research

We will try to extend the GOAT TOOLBOX™ to include not only variable selection (given functional form) procedures, but also selection procedures to choose a functional form.
Future research

We will try to extend the GOAT TOOLBOX™ to include not only variable selection (given functional form) procedures, but also selection procedures to choose a functional form.

Non-parametric problem

To check $CV_1$ against $CV_{0.25}$ in non-parametric setting (latent class models, mixed logit).

Forecasting problems

We plan to incorporate White (2000) reality check for data snooping (against a parsimonious specifications).
Comments are welcomed.

Jan Brůha
Jan_bruha@yahoo.co.uk