Estimation techniques for MEV models with sampling of alternatives

Ricardo Hurtubia, Gunnar Flötteröd, Michel Bierlaire

European Transport Conference

Glasgow, October 11-13, 2010

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Motivation

- Discrete choice models with large or unknown choice sets require sampling of alternatives
- Consistent estimation is possible for MNL (McFadden, 1978)
- Sampling in non-logit models: can't be directly extended from MNL case
- Asymptotically unbiased estimator for nested logit proposed by Guevara and Ben-Akiva (2010)
- Bias can be reduced using bootstrapping techniques and importance sampling

Outline

1. Sampling of alternatives for MNL
2. MEV models
3. Sampling of alternatives for MEV models
4. Proposed techniques for reduced-bias estimation
    - 4.1 Bootstrapping
    - 4.2 Importance Sampling
5. Conclusions

Sampling of alternatives for MNL

- Choice probability with full choice set:

$$P(i) = \frac{e^{V_{ni}}}{\sum_{j \in C_n} e^{V_{nj}}}$$

- Choice probability with a sample D_n (McFadden, 1978):

$$P(i|D_n) = \frac{e^{\mu V_{ni} + \ln \pi(D_n|i)}}{\sum_{j \in D_n} e^{\mu V_{nj} + \ln \pi(D_n|j)}}$$

- Extension of this results for non-Logit models is not straightforward

MEV models

- Generating function $G(e^{V_1}, ..., e^{V_J})$
- Choice probability:

$$P_n(i) = \frac{e^{V_{in} + \ln G_i}}{\sum_{j \in C_n} e^{V_{jn} + \ln G_j}}$$

- where $G_i = \frac{\partial G(e^{V_{1n}}, e^{V_{2n}}, \ldots, e^{V_{J_n n}})}{\partial e^{V_{in}}}$

MEV models

- Different G functions generate different models:
  - MNL: $G = \sum_{j \in C_n} e^{\mu V_{jn}}$
  - Nested Logit: $G = \sum_{m=1}^{M} \left(\sum_{j \in C_{mn}} e^{\mu_m V_{jn}} \right)^{\frac{\mu}{\mu_m}}$
  - Cross-nested Logit: $G = \sum_{m=1}^{M} \left(\sum_{j \in C_{mn}} \left(\alpha_{jm} e^{V_{in}} \right)^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$

Sampling of alternatives for MEV models

- Choice probability when considering a sample D_n (Bierlaire, Bolduc and McFadden, 2008):

$$P_n(i|D_n) = \frac{e^{V_{in}+\ln G_i+\ln \pi(D_n|i)}}{\sum_{j\in D_n} e^{V_{jn}+\ln G_j+\ln \pi(D_n|j)}}$$

- In many cases (NL, CNL) $\ln G_i$ still depends on the full choice set C_n

Sampling of alternatives for MEV models

- in the Nested Logit Case:

$$\ln G_{in} = \left(\frac{\mu}{\mu_{m(i)}} - 1 \right) \left(\ln \sum_{j \in C_{m(i)n}} e^{\mu_{m(i)} V_{jn}} \right) + \ln \mu + (\mu_{m(i)} - 1) V_{in}$$

- Logsum approximation (Guevara and Ben-Akiva, 2010):

$$\left(\ln \sum_{j \in C_{m(i)n}} e^{\mu_{m(i)} V_{jn}} \right) \approx \left(\ln \sum_{j \in D_{m(i)n}} w_{jn} e^{\mu_{m(i)} V_{jn}} \right)$$

- with $w_{jn} = \frac{\widetilde{n}_{jn}}{E_n(j)}$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Sampling of alternatives for MEV models

- D_{mn} includes the chosen alternative and randomly sampled (without replacement) elements of the nest m
- Estimation through maximum log-likelihood with the following choice probability

$$P_n(i|D_n) = \frac{e^{V_{in} + \ln G_i(D_{m(i)n}) + \ln \frac{|C_{m(i)}|}{|D_{m(i)n}|}}}{\sum_{j \in D_n} e^{V_{jn} + \ln G_j(D_{m(j)n}) + \ln \frac{|C_{m(j)}|}{|D_{m(j)n}|}}}$$

- where $\ln G_i(D_{m(i)n})$ considers the approximated logsum

Sampling of alternatives for MEV models

- The approximated logsum generates asymptotically unbiased estimates
  $\rightarrow$ biased parameters when the sample size is small (even when the true choice probabilities are used to calculate w_{jn})
- Possible improvements for the approximated logsum:
  - Correction of the bias using Bootstrapping
  - Importance sampling of the elements in the logsum

Bootstrapping

- Simulation based technique for statistical inference of the properties of an estimator, from a sub-sample of observations
- Application to the approximated logsum:
    1. Estimation using the approximated logsum
    2. Re-sampling (with replacement) from the original sample of alternatives
    3. Re-calculation of the logsum with the new sample
    4. Calculation of the bias
    5. Re-estimation correcting for the bias

Bootstrapping

- Boostrap estimator of the bias:

$$\rho_{mn} = \frac{1}{B} \sum_{b}^{B} \left(\ln \sum_{j \in D_{mn}^{b}} w_{jn} e^{\mu_m^0 V_{jn}(\beta^0)} \right) - \left(\ln \sum_{j \in D_{mn}} w_{jn} e^{\mu_m^0 V_{jn}(\beta^0)} \right)$$

- where
  - β^0, μ^0 : set of parameters from the original estimation
  - D_{mn}^b is the set of alternatives in each re-sampling instance (b)
  - B is the number of re-sampling instances

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Bootstrapping

- Estimation through maximum log-likelihood with the following choice probability

$$P_n(i|D_n) = \frac{e^{V_{in} + \ln \hat{G}_i(D_{m(i)n}) + \ln \frac{|C_{m(i)}|}{|D_{m(i)n}|}}}{\sum_{j \in D_n} e^{V_{jn} + \ln \hat{G}_j(D_{m(j)n}) + \ln \frac{|C_{m(j)}|}{|D_{m(j)n}|}}}$$

where:

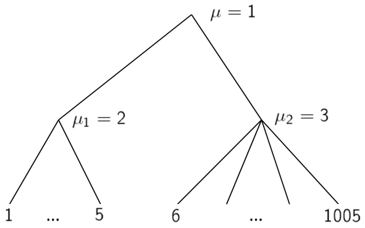
$$\ln \hat{G}_i(D_{m(i)n}) = \left(\frac{\mu}{\mu_{m(i)}} - 1\right) \left(\left(\ln \sum_{j \in D_{m(i)n}} w_{jn} e^{\mu_{m(i)} V_{jn}} \right) - \rho_{m(i)n} \right) + \ln \mu + (\mu_{m(i)} - 1) V_{in}$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Bootstrapping: Experiment

- Nested logit:



- Utility: $V_{in} = \beta_a a_{in} + \beta_b b_{in}$
- Attributes: $a_{in}, b_{in} \sim U(-1, 1)$
- True parameters $\beta_a = 1$, $\beta_b = 1$, $\mu_1 = 2$, $\mu_2 = 3$
- Sampling of alternatives within nest 2

Bootstrapping: Results

- Results from Monte Carlo experiment using approximated logsum (sample size = 5):

parameter	average value	std-error	true value	t-test
β_a	0.855	0.082	1	1.773
β_b	0.843	0.068	1	2.288 *
μ_1	2.569	0.581	2	0.978
μ_2	3.622	0.272	3	2.290 *

* Biased estimates

Bootstrapping: Results

- Results after bootstrapping (sample size = 5):

parameter	average value	std-error	true value	t-test
β_a	0.953	0.079	1	0.595
β_b	0.957	0.079	1	0.548
μ_1	2.264	0.517	2	0.511
μ_2	3.224	0.229	3	0.974

- significant reduction of the bias

Importance sampling

- The bias can be reduced with a better sample for the approximation of the logsum
- The sample of alternatives in the logsum does not have to be the same as the sample of alternatives for the choice set
- Method:
  1. Random sampling of alternatives for the elements in the logsum
  2. Estimation using approximated logsum $\rightarrow \beta^0, \mu^0$
  3. Importance sampling of alternatives for the logsum following $P(\beta^0, \mu^0)$
  4. Re-estimation

Importance sampling

- First estimation:

$$P_n(i|D_n) = \frac{e^{V_{in} + \ln G_i(L_{m(i)n}) + \ln \frac{|C_{m(i)}|}{|D_{m(i)n}|}}}{\sum_{j \in D_n} e^{V_{jn} + \ln G_j(L_{m(j)n}) + \ln \frac{|C_{m(j)}|}{|D_{m(j)n}|}}}$$

- where $L_{m(i)n}$ is randomly generated sample of alternatives for the logsum ($|L_{m(i)n}| = |D_{m(i)n}|$)
- From this first estimation we get β^0, μ^0

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Importance sampling

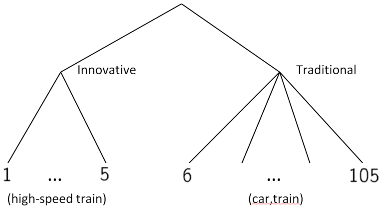
- The sampling for the logsum is done again, following a MNL inside the nest using the previously estimated parameters (β^0, μ^0)

$$g_n(i|m) = \frac{e^{V_{ni}(\beta^0,\mu^0)}}{\sum_{j \in C_m} e^{V_{nj}(\beta^0,\mu^0)}}$$

- The elements in the sample for the choice set remain the same
- A re-estimation is performed

Importance sampling: Experiment

- Synthetic data generated from a survey to evaluate a high speed train in Switzerland



- $V_{hs} = \beta_{cost} C_{hs} + \beta_{time\_T} TT_{hs} + \beta_{headway} HE_{hs}$
- $V_{car} = \beta_{cost} C_{car} + \beta_{time\_C} TT_{car}$
- $V_{train} = \beta_{cost} C_{train} + \beta_{time\_T} TT_{train} + \beta_{headway} HE_{train}$

Importance sampling:Results

- Results from Monte Carlo experiment using approximated logsum (sample size = 5):

parameter	average value	std-error	true value	t-test
β_{cost}	-1.253	0.152	-0.849	2.666 *
β_{time_C}	-2.958	0.359	-1.760	3.388 *
β_{time_T}	-2.708	0.306	-1.840	2.835 *
$\beta_{headway}$	-0.967	0.217	-0.496	2.165 *
μ_1(innovative)	1.220	0.160	2	4.873 *
μ_2(traditional)	3.146	0.368	4	2.318 *

* Biased estimates

Importance sampling:Results

- Results from Monte Carlo experiment using importance sampling (sample size = 5):

parameter	average value	std-error	true value	t-test
β_{cost}	-0.930	0.135	-0.849	0.560
β_{time_C}	-1.997	0.321	-1.760	0.736
β_{time_T}	-2.008	0.314	-1.840	0.535
$\beta_{headway}$	-0.592	0.143	-0.496	0.672
μ_1(innovative)	1.766	0.359	2	0.652
μ_2(traditional)	3.503	0.430	4	1.155

- Significant reduction of the bias

Conclusions

- Two methods that reduce the bias for MEV model estimation were presented
- Bootstrapping reduces the bias of the approximated logsum
  - bootstrapped results will depend on the quality of the original estimator
- Importance sampling of the elements in the logsum allows to find unbiased estimates
  - different sample for the choice set and the elements in the logsum
- Further work:
  - Test other correlation structures (e.g. Cross-nestedlogit)
  - Estimation over real data

Thank you