Multivariate Extreme Value models

Michel Bierlaire

michel.bierlaire@epfl.ch

Transport and Mobility LAboratory

Logit

- Random utility:
$$U_{in} = V_{in} + \varepsilon_{in}$$
- ε_{in} is i.i.d. EV (Extreme Value) distributed
- ε_{in} is the maximum of many r.v. capturing unobservable attributes, measurement and specification errors.
- Key assumption: Independence

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Relax the independence assumption

$$\begin{pmatrix} U_{1n} \\ \vdots \\ U_{Jn} \end{pmatrix} = \begin{pmatrix} V_{1n} \\ \vdots \\ V_{Jn} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1n} \\ \vdots \\ \varepsilon_{Jn} \end{pmatrix}$$

that is

$$U_n = V_n + \varepsilon_n$$

and ε_n is a vector of random variables.

Assumption about the random term:
multivariate distribution

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Relax the independence assumption

A multivariate random variable ε is represented by a density function

$$f(\varepsilon_1, \ldots, \varepsilon_J)$$

and

$$P(\varepsilon \leq x) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_J} f(\varepsilon) d\varepsilon_J \ldots d\varepsilon_1$$

where $x \in \mathbb{R}^J$ is a $J \times 1$ vector of constants.

Probit model

- Multivariate normal variable $N(\mu, \Sigma)$
- $\mu \in \mathbb{R}^J$
- $\Sigma \in \mathbb{R}^{J \times J}$, definite positive
- Density function:

$$f(\varepsilon) = (2\pi)^{-\frac{J}{2}} |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(\varepsilon-\mu)^T \Sigma^{-1} (\varepsilon-\mu)}$$

Probit model

Example: trinomial model

$$
\begin{aligned}
U_1 &= V_1 + \varepsilon_1 \\
U_2 &= V_2 + \varepsilon_2 \\
U_3 &= V_3 + \varepsilon_3
\end{aligned}
$$

and $\varepsilon \sim N(0, \Sigma)$. We have $P(2) = P(U_i - U_2 \leq 0 \quad i = 1, 2, 3)$

$$
\begin{aligned}
U_1 - U_2 &= V_1 - V_2 + \varepsilon_1 - \varepsilon_2 \\
U_3 - U_2 &= V_3 - V_2 + \varepsilon_3 - \varepsilon_2
\end{aligned}
$$

Probit model

Matrix notation with

$$\Delta_2 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

$$\Delta_2 U = \begin{pmatrix} U_1 - U_2 \\ U_3 - U_2 \end{pmatrix} \sim N(\Delta_2 V, \Delta_2 \Sigma \Delta_2^T)$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Probit model

In general, we have

$$\Delta_i U \sim N(\Delta_i V, \Delta_i \Sigma \Delta_i^T)$$

and $P(i) =$

$$P(\Delta_i U \leq 0) = \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} f(\Delta_i \varepsilon) d(\Delta_i \varepsilon)_1 \ldots d(\Delta_i \varepsilon)_{J-1}$$

with

$$f(\Delta_i \varepsilon) = (2\pi)^{-\frac{J}{2}} |\Delta_i \Sigma \Delta_i^T|^{-\frac{1}{2}} e^{-\frac{1}{2}(\Delta_i \varepsilon - \Delta_i V)^T (\Delta_i \Sigma \Delta_i^T)^{-1} (\Delta_i \varepsilon - \Delta_i V)}$$

Probit model

- The integral of the density function has no closed form
- In high dimensions, numerical integration is computationally infeasible
- Therefore, the probit model with more than 5 alternatives is very difficult to use in practice

Relax the independence assumption

- If the CDF $F(\varepsilon_1, \ldots, \varepsilon_J)$ of the distribution is known

$$f(\varepsilon_1, \ldots, \varepsilon_J) = \frac{\partial^J F}{\partial \varepsilon_1 \cdots \partial \varepsilon_J}(\varepsilon_1, \ldots, \varepsilon_J)$$

- The choice probability is

$$\begin{aligned} P(i) &= P(V_1 + \varepsilon_1 \leq V_i + \varepsilon_i, \ldots, V_J + \varepsilon_J \leq V_i + \varepsilon_i) \\ &= P(\varepsilon_1 \leq V_i + \varepsilon_i - V_1, \ldots, \varepsilon_J \leq V_i + \varepsilon_i - V_J) \\ &= \int_{\varepsilon_i = -\infty}^{\infty} F_i(V_i + \varepsilon_i - V_1, \ldots, \varepsilon_i, \ldots, V_i + \varepsilon_i - V_J) d\varepsilon_i \end{aligned}$$

where $F_i = \partial F / \partial \varepsilon_i$.

Relax the independence assumption

Operational solutions:

- Generalize the logit: the Nested Logit model
- Consider a multivariate distribution such that F is known: the Multivariate Extreme Value model

Nested logit model

- Alternatives within a nest share a random term
- Random utility of alt. i in nest C_m

$$U_i = V_i + \varepsilon_i = V_i + \varepsilon_m + \varepsilon_{im}$$

- Assume that ε_m are independent across m
- ε_{im} are i.i.d. EV with scale param. μ_m for each m

Nested logit model

Assume that the nest m is given.

$$\begin{aligned} P(i|m) &= P(U_i \geq U_j, j \in C_m) \\ &= P(V_i + \varepsilon_m + \varepsilon_{im} \geq V_j + \varepsilon_m + \varepsilon_{jm}, j \in C_m) \\ &= P(V_i + \varepsilon_{im} \geq V_j + \varepsilon_{jm}, j \in C_m) \end{aligned}$$

Then we have a logit model:

$$P(i|m) = \frac{e^{\mu_m V_i}}{\sum_{j \in C_m} e^{\mu_m V_j}}$$

Nested logit model

What is the probability of choosing nest C_m?

$$P(m) = P(\max_{i \in \mathcal{C}_m} U_i \geq \max_{j \in C_k} U_j, \forall k \neq m) =$$

$$P(\varepsilon_m + \max_{i \in C_m} (V_i + \varepsilon_{im}) \geq \varepsilon_k + \max_{j \in C_k} (V_j + \varepsilon_{jk}), \forall k \neq m)$$

Note that $V_i + \varepsilon_{im}$ is EV(V_i, μ_m). Therefore

$$\max_{i \in C_m} (V_i + \varepsilon_{im}) \sim \text{EV}(\tilde{V}_m, \mu_m)$$

where

$$\tilde{V}_m = \frac{1}{\mu_m} \ln \sum_{i \in C_m} e^{\mu_m V_i}$$

See prop. 7, page 105, chap. 5

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Nested logit model

We write the random variable

$$\max_{i \in C_m}(V_i + \varepsilon_{im}) = \tilde{V}_m + \varepsilon'_m$$

Therefore,

$$\begin{aligned} P(m) &= P(\varepsilon_m + \tilde{V}_m + \varepsilon'_m \geq \varepsilon_k + \tilde{V}_k + \varepsilon'_k, \forall k \neq m) \\ &= P(\tilde{V}_m + \tilde{\varepsilon}_m \geq \tilde{V}_k + \tilde{\varepsilon}_k, \forall k \neq m) \end{aligned}$$

Looks familiar, doesn't it?

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Nested logit model

$$P(m) = P(\tilde{V}_m + \tilde{\varepsilon}_m \geq \tilde{V}_k + \tilde{\varepsilon}_k, \forall k \neq m)$$

Assume that $\tilde{\varepsilon}_m \sim \text{EV}(0, \mu)$. Then

$$P(m) = \frac{e^{\mu \tilde{V}_m}}{\sum_k e^{\mu \tilde{V}_k}}$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Nested logit model

Putting everything together:

$$P(i) = P(i|m)P(m) = \frac{e^{\mu_m V_i}}{\sum_{j \in C_m} e^{\mu_m V_j}} \frac{e^{\mu \tilde{V}_m}}{\sum_k e^{\mu \tilde{V}_k}}$$

with

$$\tilde{V}_m = \frac{1}{\mu_m} \ln \sum_{i \in C_m} e^{\mu_m V_{im}}$$

Back

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Nested logit model

Advantages

- Nest partitioning is an intuitive concept
- Direct extension of logit
- Closed form of the model

Drawbacks

- Limited correlation structure
- What is the actual density function $f(\varepsilon)$?

MEV models

Family of models proposed by McFadden (1978) (called GEV)
Idea: a model is generated by a function

$$G : \mathbb{R}_+^J \to \mathbb{R}_+$$

From G, we can build

- The cumulative distribution function (CDF)
- The probability model
- The expected maximum utility

MEV models

1. G is homogeneous of degree $\mu > 0$, that is

$$G(\alpha y) = \alpha^{\mu} G(y)$$

2. $\lim_{y_i \to +\infty} G(y_1, \ldots, y_i, \ldots, y_J) = +\infty$, for each $i = 1, \ldots, J$,

3. the kth partial derivative with respect to k distinct y_i is non negative if k is odd and non positive if k is even, i.e., for all (distinct) indices $i_1, \ldots, i_k \in \{1, \ldots, J\}$, we have

$$(-1)^k \frac{\partial^k G}{\partial y_{i_1} \ldots \partial y_{i_k}}(y) \leq 0, \ \forall y \in \mathbb{R}^J_+.$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

- CDF: $F(\varepsilon_1,\ldots,\varepsilon_J) = e^{-G(e^{-\varepsilon_1},\ldots,e^{-\varepsilon_J})}$
- Probability: $P(i|C) = \frac{e^{V_i + \ln G_i(e^{V_1},\ldots,e^{V_J})}}{\sum_{j\in C} e^{V_j+\ln G_j(e^{V_1},\ldots,e^{V_J})}}$ with $G_i = \frac{\partial G}{\partial y_i}$. This is a closed form
- Expected maximum utility: $V_C = \frac{\ln G(\ldots)+\gamma}{\mu}$ where γ is Euler's constant.
- Note: $P(i|C) = \frac{\partial V_C}{\partial V_i}$.

MEV models

Euler's constant

$$\gamma = -\int_0^{+\infty} e^{-x} \ln x dx = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln n \right)$$

Proofs

We show first that

$$F(\varepsilon_1, \ldots, \varepsilon_J) = e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_J})}$$

defines a multivariate CDF.

- F goes to zero when one ε goes to $-\infty$

$$\begin{aligned} F(\varepsilon_1, \ldots, -\infty, \ldots, \varepsilon_J) &= e^{-G(e^{-\varepsilon_1}, \ldots, e^{+\infty}, \ldots, e^{-\varepsilon_J})} \\ &= e^{-G(e^{-\varepsilon_1}, \ldots, +\infty, \ldots, e^{-\varepsilon_J})} \\ &= e^{-\infty} \\ &= 0 \end{aligned}$$

Proofs

$$F(\varepsilon_1, \ldots, \varepsilon_J) = e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_J})}$$

- F goes to one when all ε go to $+\infty$

$$\begin{aligned} F(+\infty, \ldots, +\infty) &= e^{-G(e^{-\infty}, \ldots, e^{-\infty})} \\ &= e^{-G(0, \ldots, 0)} \\ &= e^0 \\ &= 1 \end{aligned}$$

Proofs

$$F(\varepsilon_1, \ldots, \varepsilon_J) = e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_J})}$$

- The function

$$f(\varepsilon_1, \ldots, \varepsilon_J) = \frac{\partial^J F}{\partial \varepsilon_1 \cdots \varepsilon_J}(\varepsilon_1, \ldots, \varepsilon_J) \geq 0$$

  so that it defines a PDF.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Proofs

Define recursively

$$
\begin{aligned}
Q_1 &= G_1 = \partial G / \partial y_1 \geq 0 \\
Q_k &= Q_{k-1} G_k - \partial Q_{k-1} / \partial y_k
\end{aligned}
$$

We show recursively that all (signed) terms of Q_k are ≥ 0
Assume it true for Q_{k-1}

$$
Q_{k-1} = \underbrace{Q_{k-2} G_{k-1}}_{\geq 0} \underbrace{- \partial Q_{k-2} / \partial y_{k-1}}_{\geq 0}
$$

As $G_k = \partial G / \partial y_k \geq 0$, we have

$$
Q_{k-1} G_k \geq 0
$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Proofs

$$Q_{k-1} = \underbrace{Q_{k-2} G_{k-1}}_{\geq 0} \underbrace{- \partial Q_{k-2} / \partial y_{k-1}}_{\geq 0}$$

$$\begin{array}{rcl} \partial Q_{k-1} / \partial y_k & = & \partial Q_{k-2} / \partial y_k \;\; G_{k-1} \\ & + & Q_{k-2} \;\; \partial G_{k-1} / \partial y_k \\ & - & \partial^2 Q_{k-2} / \partial y_{k-1} \partial y_k \end{array}$$

By assumption, each increase of the order of derivatives imposes a change of sign, so that

$$\partial Q_{k-1} / \partial y_k \leq 0$$

Therefore

$$Q_k = \underbrace{Q_{k-1} G_{k-1}}_{\geq 0} \underbrace{- \partial Q_{k-1} / \partial y_k}_{\geq 0}$$

Proofs

$$F(\varepsilon_1, \ldots, \varepsilon_J) = e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_J})}$$

We show recursively that

$$\frac{\partial^J F}{\partial \varepsilon_1 \cdots \varepsilon_J} = e^{-\varepsilon_1} \cdots e^{-\varepsilon_J} Q_J F \geq 0$$

By direct derivation, we have

$$\frac{\partial F}{\partial \varepsilon_1} = e^{-\varepsilon_1} G_1 F = e^{-\varepsilon_1} Q_1 F$$

Proofs

For $1 < k <= J$, assume that

$$\frac{\partial^{k-1} F}{\partial \varepsilon_1 \cdots \varepsilon_{k-1}} = e^{-\varepsilon_1} \cdots e^{-\varepsilon_{k-1}} Q_{k-1} F \geq 0$$

Define $y_i = e^{-\varepsilon_i}$

$$\begin{aligned}
\frac{\partial^k F}{\partial \varepsilon_1 \cdots \varepsilon_k} &= \frac{\partial}{\partial y_k} \left(e^{-\varepsilon_1} \cdots e^{-\varepsilon_{k-1}} Q_{k-1} F\right) \frac{\partial y_k}{\partial \varepsilon_k} \\
&= e^{-\varepsilon_1} \cdots e^{-\varepsilon_{k-1}} \left(\frac{\partial Q_{k-1}}{\partial y_k} F + Q_{k-1} \frac{\partial F}{\partial G} \frac{\partial G}{\partial y_k} \right) (-e^{-\varepsilon_k})
\end{aligned}$$

Proofs

$$
\begin{aligned}
\frac{\partial^k F}{\partial \varepsilon_1 \cdots \varepsilon_k} &= e^{-\varepsilon_1} \cdots e^{-\varepsilon_{k-1}} \left(\frac{\partial Q_{k-1}}{\partial y_k} F + Q_{k-1} \frac{\partial F}{\partial G} \frac{\partial G}{\partial y_k} \right) (-e^{-\varepsilon_k}) \\
&= e^{-\varepsilon_1} \cdots e^{-\varepsilon_{k-1}} \left(\frac{\partial Q_{k-1}}{\partial y_k} F + Q_{k-1} (-F) G_k \right) (-e^{-\varepsilon_k}) \\
&= e^{-\varepsilon_1} \cdots e^{-\varepsilon_{k-1}} e^{-\varepsilon_k} \left(Q_{k-1} G_k F - \frac{\partial Q_{k-1}}{\partial y_k} F \right) \\
&= e^{-\varepsilon_1} \cdots e^{-\varepsilon_{k-1}} e^{-\varepsilon_k} Q_k F \\
&\geq 0
\end{aligned}
$$

Proofs

Marginal distributions for i

$$F(\varepsilon_1, \ldots, \varepsilon_J) = e^{-G(e^{-\varepsilon_1},\ldots,e^{-\varepsilon_J})}$$

$$\begin{array}{lrcl} \varepsilon_j \to +\infty, \forall j \neq i & F & = & e^{-G(0,\ldots,e^{-\varepsilon_i},\ldots,0)} \\ & & = & e^{-e^{-\mu\varepsilon_i} G(0,\ldots,1,\ldots,0)} \\ & & = & e^{-\alpha e^{-\mu\varepsilon_i}} \\ & & = & e^{-e^{-\mu\varepsilon_i + \ln\alpha}} \end{array}$$

This is an extreme value distribution

F is a multivariate extreme value distribution

Proofs

$$F(\varepsilon_1,\ldots,\varepsilon_J) = e^{-G(e^{-\varepsilon_1},\ldots,e^{-\varepsilon_J})}$$

$$F_i = e^{-\varepsilon_i} G_i F$$

Probability of the first alternative

$$\begin{aligned} P(1) &= \int_{\varepsilon_1=-\infty}^{+\infty} F_1(\varepsilon_1, V_1 - V_2 + \varepsilon_1, \ldots, V_1 - V_J + \varepsilon_1) d\varepsilon_1 \\ &= \int_{\varepsilon_1=-\infty}^{+\infty} \quad e^{-\varepsilon_1} \\ & \qquad G_1(e^{-\varepsilon_1}, e^{-V_1+V_2-\varepsilon_1}, \ldots, e^{-V_1+V_J-\varepsilon_1}) \\ & \qquad e^{-G(e^{-\varepsilon_1}, e^{-V_1+V_2-\varepsilon_1}, \ldots, e^{-V_1+V_J-\varepsilon_1})} \\ & \qquad d\varepsilon_1 \end{aligned}$$

Proofs

- G is homogeneous of degree μ ($\alpha = e^{-V_1-\varepsilon_1}$)

$$
\begin{aligned}
& G(e^{-\varepsilon_1}, e^{-V_1+V_2-\varepsilon_1}, \ldots, e^{-V_1+V_J-\varepsilon_1}) \\
= \quad & e^{-\mu(V_1+\varepsilon_1)} G(e^{V_1}, e^{V_2}, \ldots, e^{V_J}) \\
= \quad & e^{-\mu(V_1+\varepsilon_1)} G
\end{aligned}
$$

- G_1 is homogeneous of degree $\mu - 1$

$$
\begin{aligned}
& G_1(e^{-\varepsilon_1}, e^{-V_1+V_2-\varepsilon_1}, \ldots, e^{-V_1+V_J-\varepsilon_1}) \\
= \quad & e^{-\mu(V_1+\varepsilon_1)} e^{V_1+\varepsilon_1} G_1(e^{V_1}, e^{V_2}, \ldots, e^{V_J}) \\
= \quad & e^{-\mu(V_1+\varepsilon_1)} e^{V_1+\varepsilon_1} G_1
\end{aligned}
$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Proofs

$$P(1) = \int_{\varepsilon_1=-\infty}^{+\infty} e^{-\varepsilon_1} e^{-\mu(V_1+\varepsilon_1)} e^{V_1+\varepsilon_1} G_1 e^{-Ge^{-\mu(V_1+\varepsilon_1)}} d\varepsilon_1$$

$$= e^{V_1} G_1 \int_{\varepsilon_1=-\infty}^{+\infty} e^{-\mu(V_1+\varepsilon_1)} e^{-Ge^{-\mu(V_1+\varepsilon_1)}} d\varepsilon_1$$

$$t = -e^{-\mu V_1 - \mu\varepsilon_1} \quad dt = \mu e^{-\mu V_1 - \mu\varepsilon_1} d\varepsilon_1$$

$$= e^{V_1} G_1 \frac{1}{\mu} \int_{-\infty}^{0} e^{tG} dt = \frac{e^{V_1} G_1}{\mu G}$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Proofs

$$P(i) = \frac{e^{V_i} G_i(e^{V_1}, e^{V_2}, \ldots, e^{V_J})}{\mu G(e^{V_1}, e^{V_2}, \ldots, e^{V_J})}$$

Euler's theorem: $\mu G(y_1, \ldots, y_J) = \sum_{j=1}^{J} y_j G_j$

$$\boxed{P(i) = \frac{e^{V_i + \ln G_i(e^{V_1}, e^{V_2}, \ldots, e^{V_J})}}{\sum_j e^{V_j + \ln G_j(e^{V_1}, e^{V_2}, \ldots, e^{V_J})}}}$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Proofs

Expected maximum utility V_C

- Notation:

$$\langle x_j \rangle_j = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_J \end{pmatrix}$$

- For each realization of ε, there is one i which corresponds to the maximum utility
- If i is the maximum utility, the EMU is

$$\bar{V}_i = \int_{\varepsilon_i=-\infty}^{\infty} (V_i + \varepsilon_i) F_i \left(\langle V_i + \varepsilon_i - V_j \rangle_j \right) d\varepsilon_i$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Proofs

- Note $a = G(\langle e^{V_j} \rangle_j)$
- We have $F_i = e^{-\varepsilon_i} G_i F$

$$F_i\left(\langle V_i + \varepsilon_i - V_j \rangle_j\right) =$$

$$e^{-\varepsilon_i} G_i\left(\langle e^{-V_i - \varepsilon_i + V_j} \rangle_j\right) e^{\left(-G(\langle e^{-V_i - \varepsilon_i + V_j} \rangle_j)\right)} =$$

$$e^{-\varepsilon_i} e^{-(\mu-1)(V_i+\varepsilon_i)} G_i\left(\langle e^{V_j} \rangle_j\right) e^{\left(-e^{-\mu(V_i+\varepsilon_i)} G(\langle e^{V_j} \rangle_j)\right)} =$$

$$e^{-\mu(V_i+\varepsilon_i)} e^{V_i} G_i\left(\langle e^{V_j} \rangle_j\right) e^{-a e^{-\mu(V_i+\varepsilon_i)}}$$

Proofs

$$\bar{V}_i = \int_{\varepsilon_i=-\infty}^{\infty} (V_i + \varepsilon_i) F_i \left(\langle V_i + \varepsilon_i - V_j \rangle_j \right) d\varepsilon_i =$$

$$\int_{\varepsilon_i=-\infty}^{\infty} (V_i + \varepsilon_i) e^{-\mu(V_i+\varepsilon_i)} e^{V_i} G_i \left(\langle e^{V_j} \rangle_j \right) e^{-ae^{-\mu(V_i+\varepsilon_i)}} d\varepsilon_i$$

$$w = V_i + \varepsilon_i$$

$$\bar{V}_i = \int_{w=-\infty}^{\infty} w e^{-\mu w} e^{V_i} G_i \left(\langle e^{V_j} \rangle_j \right) e^{-ae^{-\mu w}} dw$$

$$V_C = \sum_{i=1}^{J} \bar{V}_i$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Proofs

$$V_C = \sum_{i=1}^{J} \int_{w=-\infty}^{\infty} we^{-\mu w} e^{V_i} G_i \left(\langle e^{V_j} \rangle_j \right) e^{-ae^{-\mu w}} dw =$$

$$\int_{w=-\infty}^{\infty} we^{-\mu w} \sum_{i=1}^{J} (e^{V_i} G_i \left(\langle e^{V_j} \rangle_j \right)) e^{-ae^{-\mu w}} dw$$

Euler's theorem: $\sum e^{V_i} G_i = \mu G = \mu a$

$$V_C = \int_{w=-\infty}^{\infty} we^{-\mu w} \mu a e^{-ae^{-\mu w}} dw$$

Proofs

$$V_C = \int_{w=-\infty}^{\infty} we^{-\mu w}\mu a e^{-ae^{-\mu w}} dw$$

$$t = ae^{-\mu w} \quad dt = -\mu a e^{-\mu w} dw \quad w = \frac{1}{\mu}(\ln a - \ln t)$$

$$V_C = -\frac{1}{\mu}\int_{t=+\infty}^{0} (\ln a - \ln t) e^{-t} dt =$$

$$-\frac{1}{\mu}\int_{t=0}^{+\infty} \ln t e^{-t} dt + \frac{1}{\mu} \ln a \int_{t=0}^{+\infty} e^{-t} dt$$

$$V_C = \frac{\gamma + \ln a}{\mu} = \frac{\gamma + \ln G(\langle e^{V_j} \rangle_j)}{\mu}$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV vs GEV

- McFadden introduces the General Extreme Value model (GEV)
- In statistics, a Generalized Extreme Value distribution (Jenkinson, 1955) is a univariate distribution with CDF

$$F_X(x) = \begin{cases} e^{-(1+\xi((x-\mu)/\sigma))^{-1/\xi})} & -\infty < x \leq \mu - \sigma/\xi \quad \text{for } \xi < 0 \\ & \mu - \sigma/\xi \leq x < \infty \quad \text{for } \xi > 0 \\ e^{-e^{-(x-\mu)/\sigma}} & -\infty < x < \infty \quad \text{for } \xi = 0 \end{cases}$$

$\xi = 0$	Type 1 EV distribution
$\xi > 0$	Type 2 EV distribution
$\xi < 0$	Type 3 EV distribution

MEV models

Example: $G(y) = \sum_{i=1}^{J} y_i^{\mu}$

1. $G(\alpha y) = \sum_{i=1}^{J} (\alpha y_i)^{\mu} = \alpha^{\mu} \sum_{i=1}^{J} y_i^{\mu} = \alpha^{\mu} G(y)$
2. $\lim_{y_i \to +\infty} G(y) = +\infty$, $i = 1, \ldots, J$
3. $\frac{\partial G}{\partial y_i} = \mu y_i^{\mu-1}$ and $\frac{\partial^2 G}{\partial y_i \partial y_j} = 0$

$\boxed{G \text{ complies with the theory}}$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

Example: $G(y) = \sum_{i=1}^{J} y_i^{\mu}$

$$\begin{aligned} F(\varepsilon_1, \ldots, \varepsilon_J) &= e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_J})} \\ &= e^{-\sum_{i=1}^{J} e^{-\mu\varepsilon_i}} \\ &= \prod_{i=1}^{J} e^{-e^{-\mu\varepsilon_i}} \end{aligned}$$

Product of i.i.d EV

Logit Model

MEV models

Example: $G(e^{V_1}, \ldots, e^{V_J}) = \sum_{i=1}^{J} e^{\mu V_i}$

$$P(i) = \frac{e^{V_i + \ln G_i(e^{V_1}, \ldots, e^{V_J})}}{\sum_{j \in C} e^{V_j + \ln G_j(e^{V_1}, \ldots, e^{V_J})}}$$ with $G_i(x) = \mu x_i^{\mu-1}$

$$\begin{aligned} e^{V_i + \ln G_i(e^{V_1}, \ldots, e^{V_J})} &= e^{V_i + \ln \mu + (\mu-1) \ln e^{V_i}} \\ &= e^{\ln \mu + \mu V_i} \end{aligned}$$

$$P(i) = \frac{e^{\ln \mu + \mu V_i}}{\sum_{j \in C} e^{\ln \mu + \mu V_j}} = \frac{e^{\mu V_i}}{\sum_{j \in C} e^{\mu V_j}}$$

MEV models

Example: $G(e^{V_1}, \ldots, e^{V_J}) = \sum_{i=1}^{J} e^{\mu V_i}$

$$
\begin{aligned}
V_C &= \tfrac{1}{\mu}\left(\ln G(e^{V_1}, \ldots, e^{V_J}) + \gamma\right) \\
&= \tfrac{1}{\mu} \ln \sum_{i=1}^{J} e^{\mu V_i} + \frac{\gamma}{\mu}
\end{aligned}
$$

Remember the NL formulation?

MEV models

Example: Nested logit

$$G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

1. $G(\alpha y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} (\alpha y_i)^{\mu_m} \right)^{\frac{\mu}{\mu_m}} = \alpha^{\mu} \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$
2. $\lim_{y_i \to +\infty} G(y) = +\infty$, $i = 1, \ldots, J$

MEV models

Example: $G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$

3.

$$\frac{\partial G}{\partial y_i} = \frac{\mu}{\mu_m} \mu_m y_i^{\mu_m - 1} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m} - 1} \geq 0$$

$$\frac{\partial^2 G}{\partial y_i y_j} = \mu \mu_m y_i^{\mu_m - 1} y_j^{\mu_m - 1} (\frac{\mu}{\mu_m} - 1) \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m} - 2} \leq 0$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

- The logit model is a MEV model
- The nested logit model is also a MEV model

$$G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

- If $\frac{\mu}{\mu_m} \leq 1$, then G complies with the theory
- Are there other such models?

Cross-Nested logit model

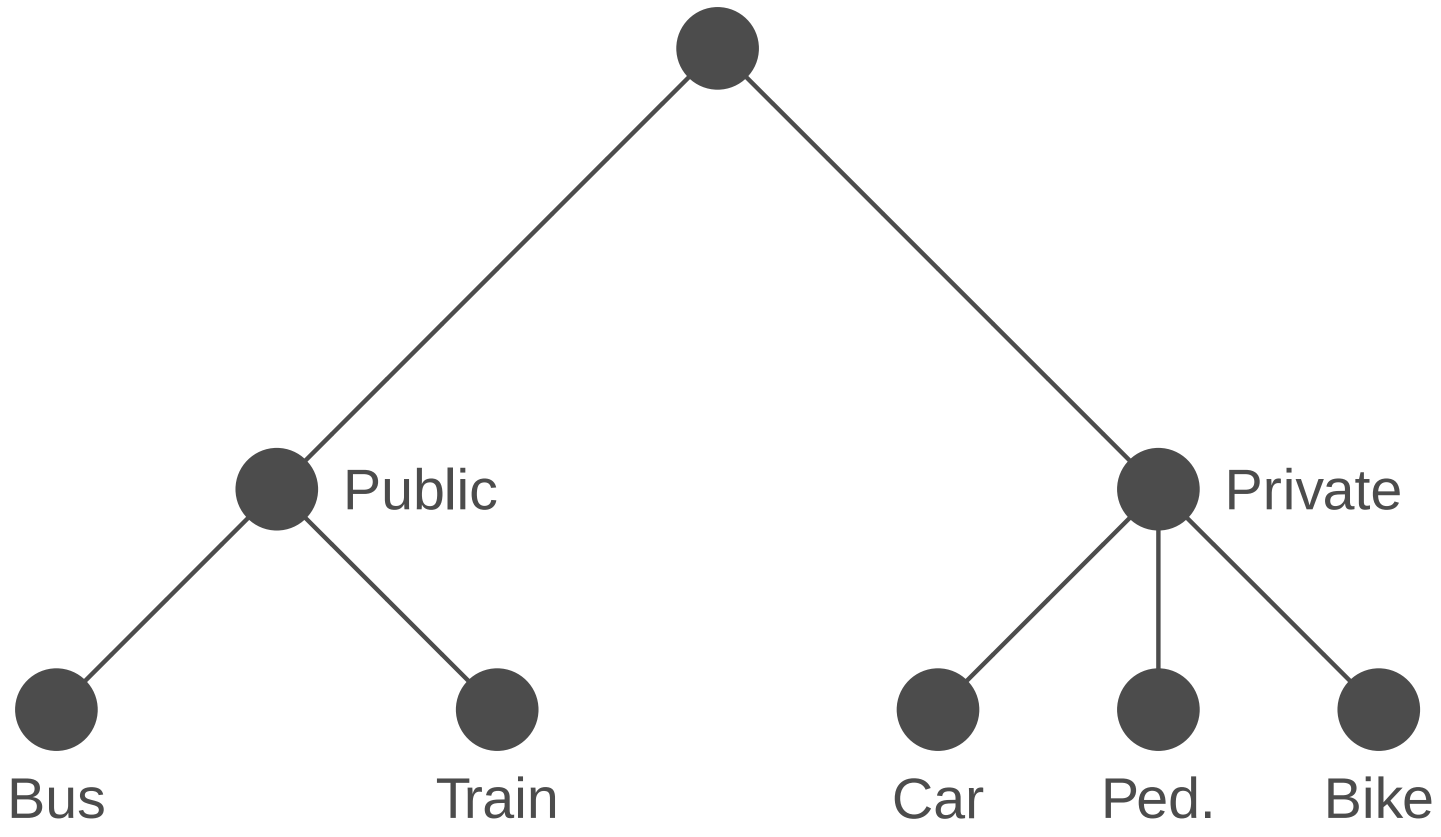
- MEV model with

$$G(y_1,\ldots,y_J) = \sum_{m=1}^{M} \left(\sum_j ({\alpha_{jm}}^{1/\mu} y_j)^{\mu_m} \right)^{\frac{\mu}{\mu_m}},$$

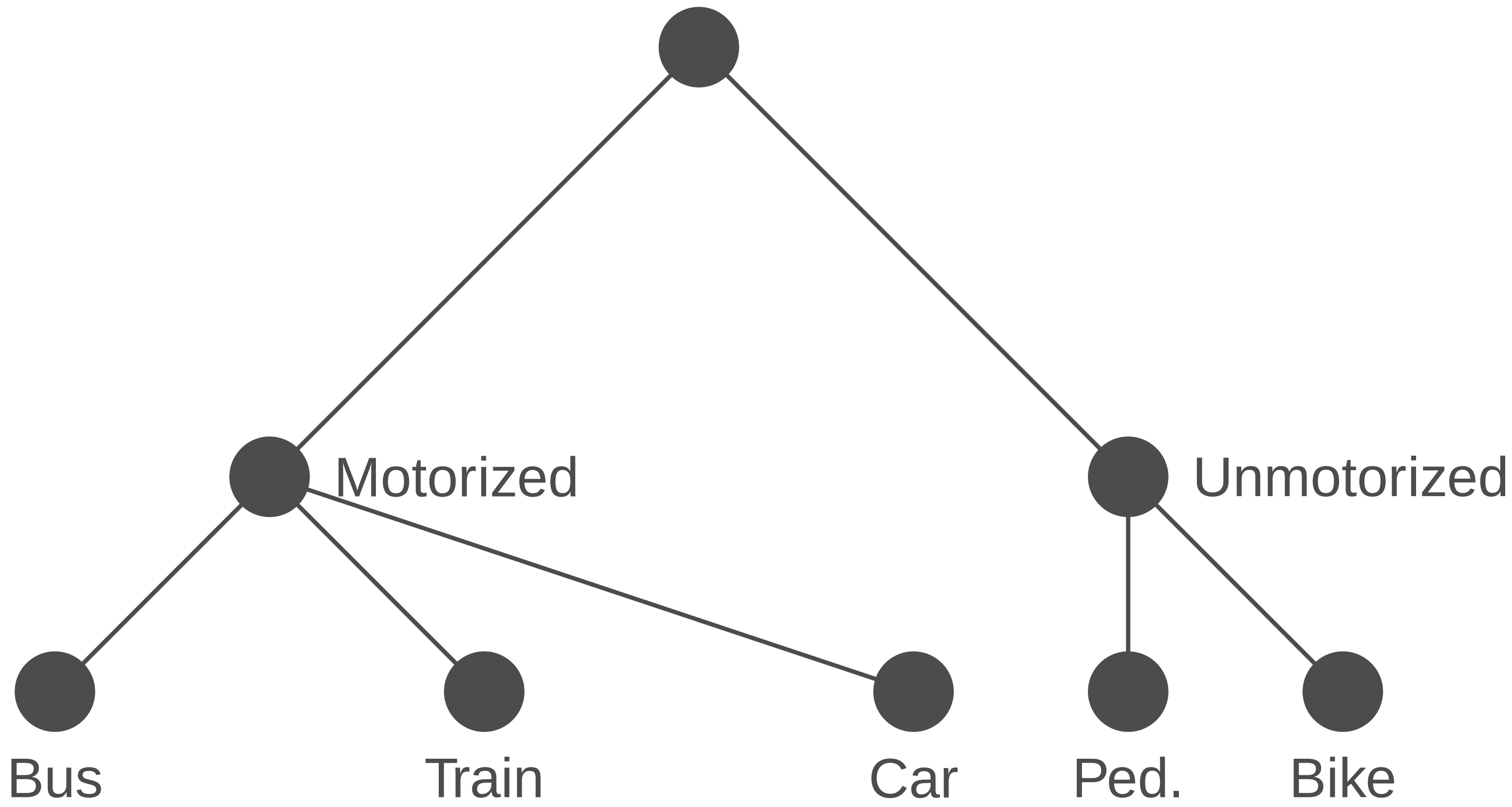
  with $\frac{\mu}{\mu_m} \leq 1$, $\alpha_{jm} \geq 0$, and $\forall j, \exists m$ s.t. $\alpha_{jm} > 0$

- Generalization of the nested-logit model

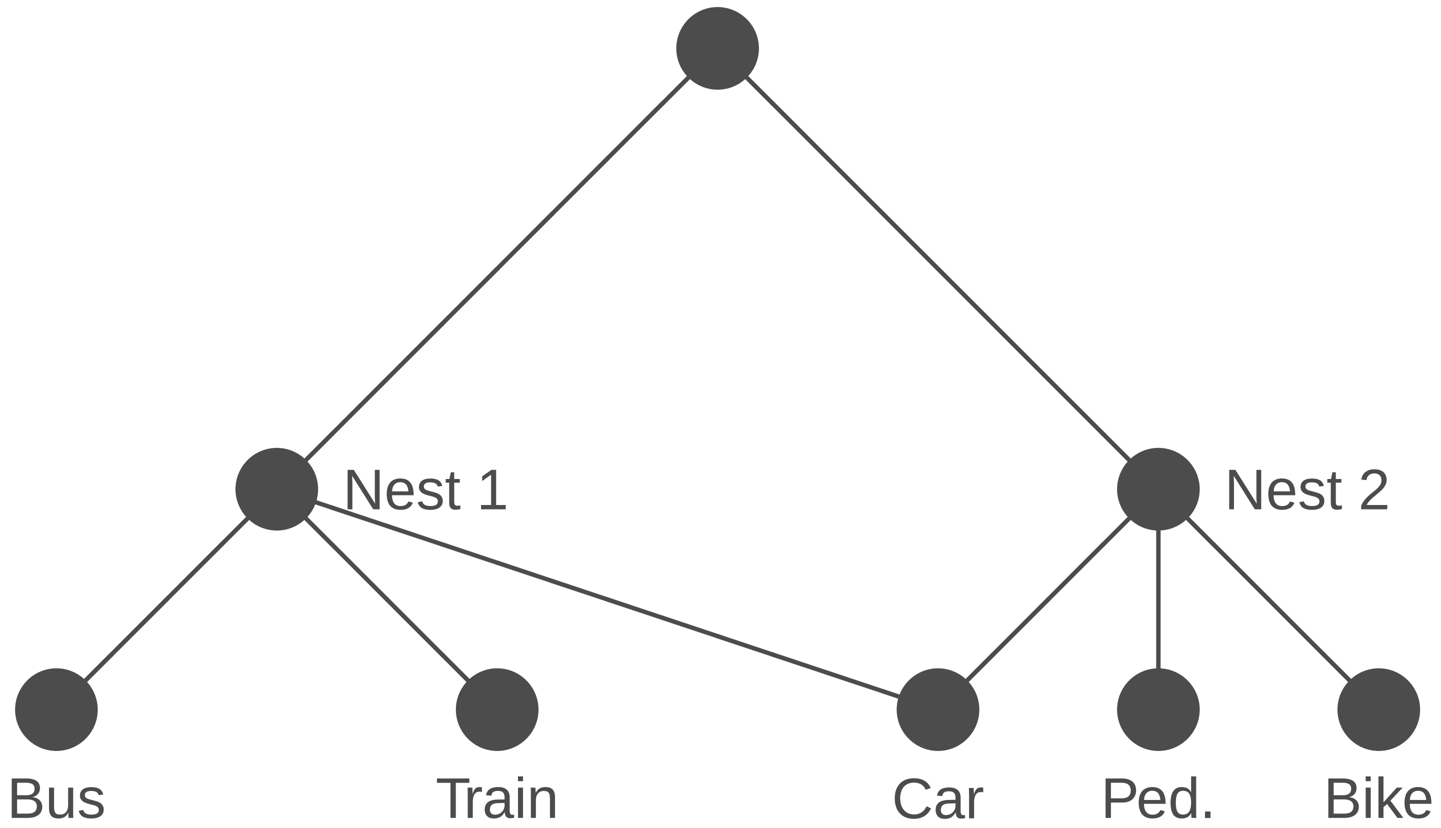
Nested Logit Model



Nested Logit Model



Cross-Nested Logit Model



Cross-Nested Logit Model

$$P(i|\mathcal{C}) = \sum_{m=1}^{M} \frac{\left(\sum_{j\in\mathcal{C}} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}\right)^{\frac{\mu}{\mu_m}}}{\sum_{n=1}^{M}\left(\sum_{j\in\mathcal{C}} \alpha_{jn}^{\mu_n/\mu} e^{\mu_n V_j}\right)^{\frac{\mu}{\mu_n}}} \frac{\alpha_{im}^{\mu_m/\mu} e^{\mu_m V_i}}{\sum_{j\in\mathcal{C}} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}}.$$

which can nicely be interpreted as

$$P(i|\mathcal{C}) = \sum_{m} P(m|\mathcal{C}) P(i|m).$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

- Provide a great deal of flexibility
- Require significant imagination
- Require heavy proofs

Network MEV

Bierlaire (2002), Daly & Bierlaire (2006)
Motivations:

- Extension of the tree representation for Nested Logit
- Investigate new MEV models
- Provide the proof once for all

Network MEV

Let (V, E) be a network with link parameters $\alpha_{(i,j)} \geq 0$

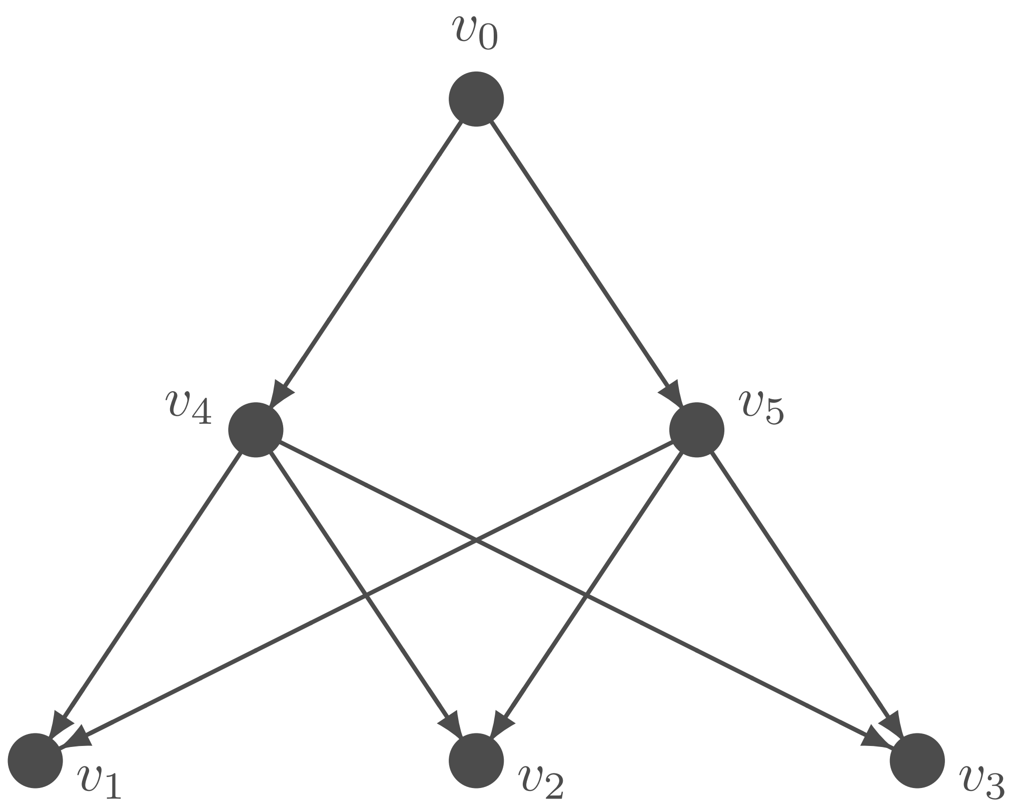
Assumptions:

1. No circuit.
2. One node without predecessor: *root*.
3. J nodes without successor: *alternatives*.
4. For each node v_i, there exists at least one path from the root to v_i such that $\prod_{k=1}^{P} \alpha_{(i_{k-1}, i_k)} > 0$.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Network MEV



Network MEV

For each node v_i, we define

- a set of indices $I_i \subseteq \{1, \ldots, J\}$ of J_i relevant alternatives,
- a homogeneous function $G^i : \mathbb{R}^{J_i} \longrightarrow \mathbb{R}$, and
- a parameter μ_i.

Recursive definition of I_i:

- $I_i = \{i\}$ for alternatives,
- $I_i = \bigcup_{j \in \text{succ}(i)} I_j$ for all other nodes.

Network MEV

Recursive definition of G^i:

For alternatives:

$$G^i \quad : \quad \mathbb{R} \longrightarrow \mathbb{R} \quad : \quad G^i(y_i) = y_i^{\mu_i} \quad i = 1, \ldots, J$$

For all others:

$$G^i \quad : \quad \mathbb{R}^{J_i} \longrightarrow \mathbb{R} \quad : \quad G^i(y) = \sum_{j \in \text{succ}(i)} \alpha_{(i,j)} G^j(y)^{\frac{\mu_i}{\mu_j}}$$

Theorem

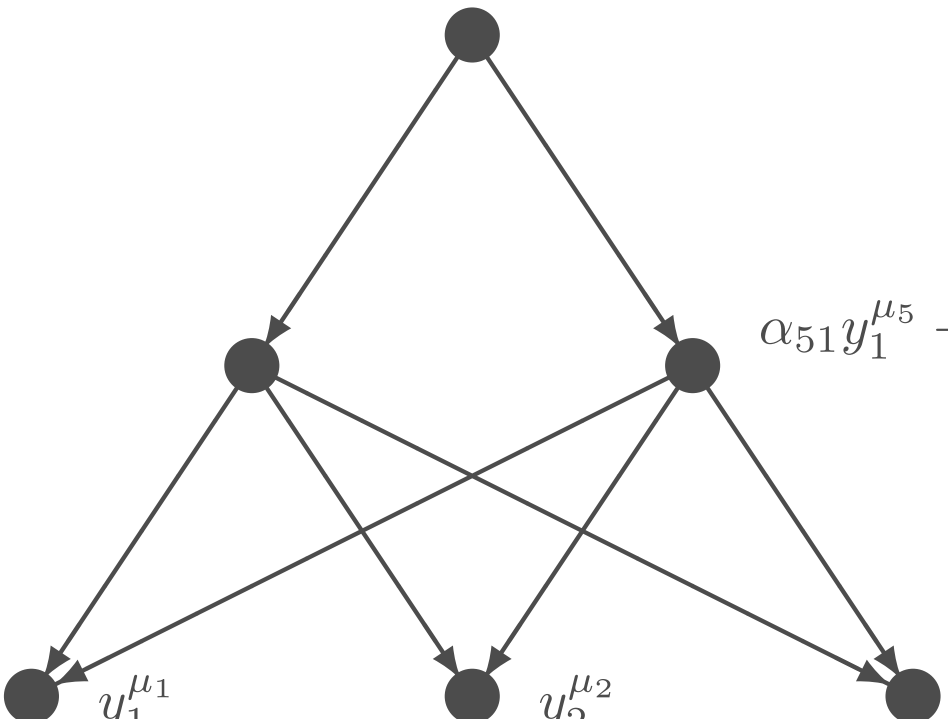
If all $G^j(y)$ are MEV generating functions, so is G^i

Network MEV

Example: Cross-Nested Logit

$$G = \sum_m \left(\sum_{j \in C} \alpha_{jm} y_j^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

$$\sum_{i=4,5} \alpha_{0i} \left(\alpha_{i1} y_1^{\mu_i} + \alpha_{i2} y_2^{\mu_i} + \alpha_{i3} y_3^{\mu_i} \right)^{\frac{\mu_0}{\mu_i}}$$



$$\alpha_{51} y_1^{\mu_5} + \alpha_{52} y_2^{\mu_5} + \alpha_{53} y_3^{\mu_5}$$

$y_1^{\mu_1}$ $y_2^{\mu_2}$ $y_3^{\mu_3}$

Network MEV

Similar idea: Daly (2001) *Recursive Nested EV Model*

Advantages :

- Easy to design
- No more proof necessary

Summary

- Need to relax the independence assumption
- Probit
- Nested logit
- MEV family
- CNL