Normalization and correlation of cross-nested logit models

E. Abbe, **M. Bierlaire**, **T. Toledo**

Lab. for Information and Decision Systems, Massachusetts Institute of Technology

Inst. of Mathematics, Ecole Polytechnique Fédérale de Lausanne, Switzerland

Transport Research Institute, Technion, Israël

Introduction

- GEV models
- Cross-nested formulations
- Normalization
- Variance-covariance structure
- Simple examples
- Route choice examples

GEV models

$$U = V + \varepsilon, \text{random vector}, V \in \mathbb{R}^J$$

$$F_{\varepsilon_1,\ldots,\varepsilon_J}(y_1, \ldots, y_J) = e^{-G(e^{-y_1},\ldots,e^{-y_J})}.$$

$$P(i|C) = \frac{y_i G_i(y_1, \ldots, y_J)}{\mu G(y_1, \ldots, y_J)} = \frac{e^{V_i + \ln G_i(\ldots)}}{\sum_{j \in \mathcal{C}} e^{V_j + \ln G_j(\ldots)}},$$

where $G_i = \frac{\partial G}{\partial y_i}$, $y_i = e^{V_i}$ $G : \mathbb{R}_+^J \to \mathbb{R}^+$.

GEV models

MNL: $$G(y_1,\ldots,y_J) = \sum_{j\in\mathcal{C}} y_j^{\mu}$$

NL: $$G(y_1,\ldots,y_J) = \sum_{m=1}^{M}\left(\sum_{j=1}^{J_m} y_j^{\mu_m}\right)^{\frac{\mu}{\mu_m}}$$

CNL: $$G(y_1,\ldots,y_J) = \sum_{m=1}^{M}\left(\sum_{j\in\mathcal{C}} ({\alpha_{jm}}^{1/\mu} y_j)^{\mu_m}\right)^{\frac{\mu}{\mu_m}}$$

Cross-nested logit models

Ordered GEV mode, Small (1987)

$$G(y_1,\ldots,y_J) = \sum_{r=1}^{J+M} \left(\sum_{j\in B_r} w_{r-j} y_j^{1/\rho_r} \right)^{\rho_r},$$

Vovsha (1997)

$$G(y_1,\ldots,y_J) = \sum_{m} \left(\sum_{j\in C} \alpha_{jm} y_j \right)^{\mu}$$

Cross-nested logit models

Ben-Akiva & Bierlaire (1999)

$$G(y_1,\ldots,y_J)=\sum_m\left(\sum_{j\in C}\alpha_{jm}y_j^{\mu_m}\right)^{\frac{\mu}{\mu_m}}$$

Wen & Koppelman (2001)

$$G(y_1,\ldots,y_J)=\sum_m\left(\sum_{n'\in N_m}\left(\alpha_{n'm}x_{n'}\right)^{\frac{1}{\mu_m}}\right)^{\mu_m}$$

Cross-nested logit models

Papola (2004)

$$G(y_1,\ldots,y_J) = \sum_k \left(\sum_{j \in C_k} \alpha_{jk}^{\theta_0/\theta_k} y_j^{1/\theta_k} \right)^{\frac{\theta_k}{\theta_0}}$$

- General formulation (as Ben-Akiva & Bierlaire)
- Appropriate for easy normalization (as Wen & Koppelman)
- For consistency with GEV, we note $\mu = 1/\theta_0$ and $\mu_k = 1/\theta_k$

Normalization

$$F_{\varepsilon_1,\ldots,\varepsilon_J}(y_1,\ldots,y_J) = e^{-G(e^{-y_1},\ldots,e^{-y_J})}.$$

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

Marginal distribution of ε_j:

$$F_{\varepsilon_j}(y_j) = \exp\left[-\exp\left\{ -\mu \left(y_j - \frac{\ln\left(\sum_{m=1}^{M} \alpha_{jm}\right)}{\mu} \right) \right\} \right].$$

Normalization

Marginal distribution: extreme value with location parameter

$$\frac{\ln\left(\sum_{m=1}^{M}\alpha_{jm}\right)}{\mu}$$

and scale parameter μ. Therefore,

$$E[\varepsilon_j] = \frac{\ln\left(\sum_{m=1}^{M}\alpha_{jm}\right)+\gamma}{\mu}.$$

Not necessarily constant across alternatives

Normalization: $\sum_{m=1}^{M}\alpha_{jm} = K$, ($K = 1$ makes sense)

Normalization

- Wen & Koppelman (2001) "*the additional condition* $\sum_m \alpha_{mn} = 1, \forall n$ *provides a useful interpretation with respect to allocation of each alternative to each nest*"
- Condition is formally required to obtain an unbiased model
- The bias can be absorbed by the ASCs, if a full set is in the model
- Normalization for Ben-Akiva & Bierlaire:

$$\sum_{m=1}^{M} \alpha_{jm}^{\frac{\mu}{\mu_m}} = c, \quad j \in \mathcal{C}$$

Variance-covariance

Multinomial logit model

$$\mathrm{Corr}(U_i, U_j) = 0 \text{ if } i \neq j.$$

Nested logit model

$$\mathrm{Corr}(U_i, U_j) = \left(1 - \left(\frac{\mu}{\mu_m}\right)^2\right) \delta_m(i,j)$$

where

$$\delta_m(i,j) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are both in nest } m \\ 0 & \text{otherwise} \end{cases}$$

Variance-covariance

Cross-nested logit model (Papola, 2004)

$$\mathrm{Corr}(U_i, U_j) = \sum_{m=1}^{M} {\alpha_{im}}^{1/2} {\alpha_{jm}}^{1/2} \left(1 - (\frac{\mu}{\mu_m})^2\right).$$

- Conjecture
- Exact for limit cases (NL)
- Linear interpolation between limit cases
- Weights chosen so that, when $i = j$,

$$\sum_{m=1}^{M} {\alpha_{im}}^{1/2} {\alpha_{im}}^{1/2} = \sum_{m=1}^{M} \alpha_{im} = 1$$

Variance-covariance

CNL is equivalent to the model defined by

$$U_j = \max_{m=1,\ldots,M} \widehat{U}_{jm}$$

where

$$\widehat{U}_{jm} = V_j + \frac{\ln \alpha_{jm}}{\mu} + \varepsilon_{jm}$$

and ε_{jm} are such that

- ε_{jm} is independent of $\varepsilon_{\ell n}$
- CDF of ε is

$$F_{\varepsilon_{1m},\ldots,\varepsilon_{Jm}}(y_1,\ldots,y_J) = e^{-\left(\sum_{j\in C} e^{-\mu_m y_j}\right)^{\frac{\mu}{\mu_m}}}$$

Variance-covariance

The error structure of a CNL is the maximum of error terms of underlying NL models

Therefore,

$$\mathrm{Corr}(U_i, U_j) = \mathrm{Corr}\left(\max_m \widehat{\varepsilon}_{im}, \max_m \widehat{\varepsilon}_{jm}\right)$$

where

$$\mathrm{Corr}(\widehat{\varepsilon}_{im}, \widehat{\varepsilon}_{jn}) = \left(1 - \left(\frac{\mu}{\mu_m}\right)^2\right) \delta_{m,n}.$$

Variance-covariance

Equivalently,

$$\mathrm{Corr}(U_i, U_j) = \mathrm{Corr}\left(\max_m \left(\frac{\ln \alpha_{im}}{\mu} + \varepsilon_{im}\right), \max_m \left(\frac{\ln \alpha_{jm}}{\mu} + \varepsilon_{jm}\right)\right.$$

where

$$\mathrm{Corr}(\varepsilon_{im}, \varepsilon_{jn}) = \left(1 - \left(\frac{\mu}{\mu_m}\right)^2\right) \delta_{m,n}.$$

- Assumption of the conjecture: *linear*
- Actual relationship: *maximum*

Variance-covariance

Computation of the true variance-covariance from the joint CDF

$$\mathrm{Corr}(U_i, U_j) = \frac{6\mu^2}{\pi^2} \int\int_{\mathbb{R}^2} x_i x_j \partial^2_{x_i x_j} F_{\varepsilon_i,\varepsilon_j}(x_i, x_j) dx_i dx_j - \frac{6\gamma^2}{\pi^2},$$

where

$$F_{\varepsilon_i,\varepsilon_j}(x_i, x_j) = e^{-\sum_{m=1}^{M}\left((\alpha_{im}^{1/\mu} e^{-x_i})^{\mu m} + (\alpha_{jm}^{1/\mu} e^{-x_j})^{\mu m}\right)^{\frac{\mu}{\mu m}}}.$$

Variance-covariance

Parameter identification from a given variance-covariance

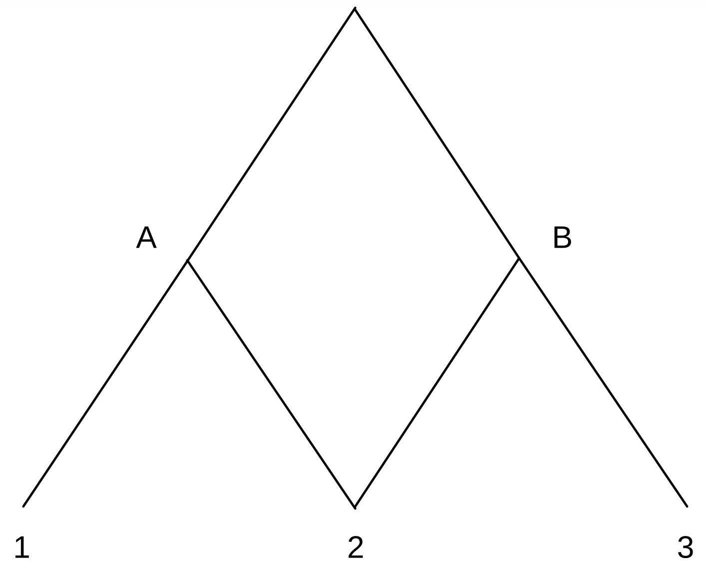
- System of equations:

$$\begin{array}{rcllr} \mathrm{Corr}(U_i, U_j) & = & c_{ij} & J(J-1)/2 & \text{equations} \\ \sum_m \alpha_{im} & = & 1 & J & \text{equations} \end{array}$$

  Total: $(J^2 + J)/2$ equations

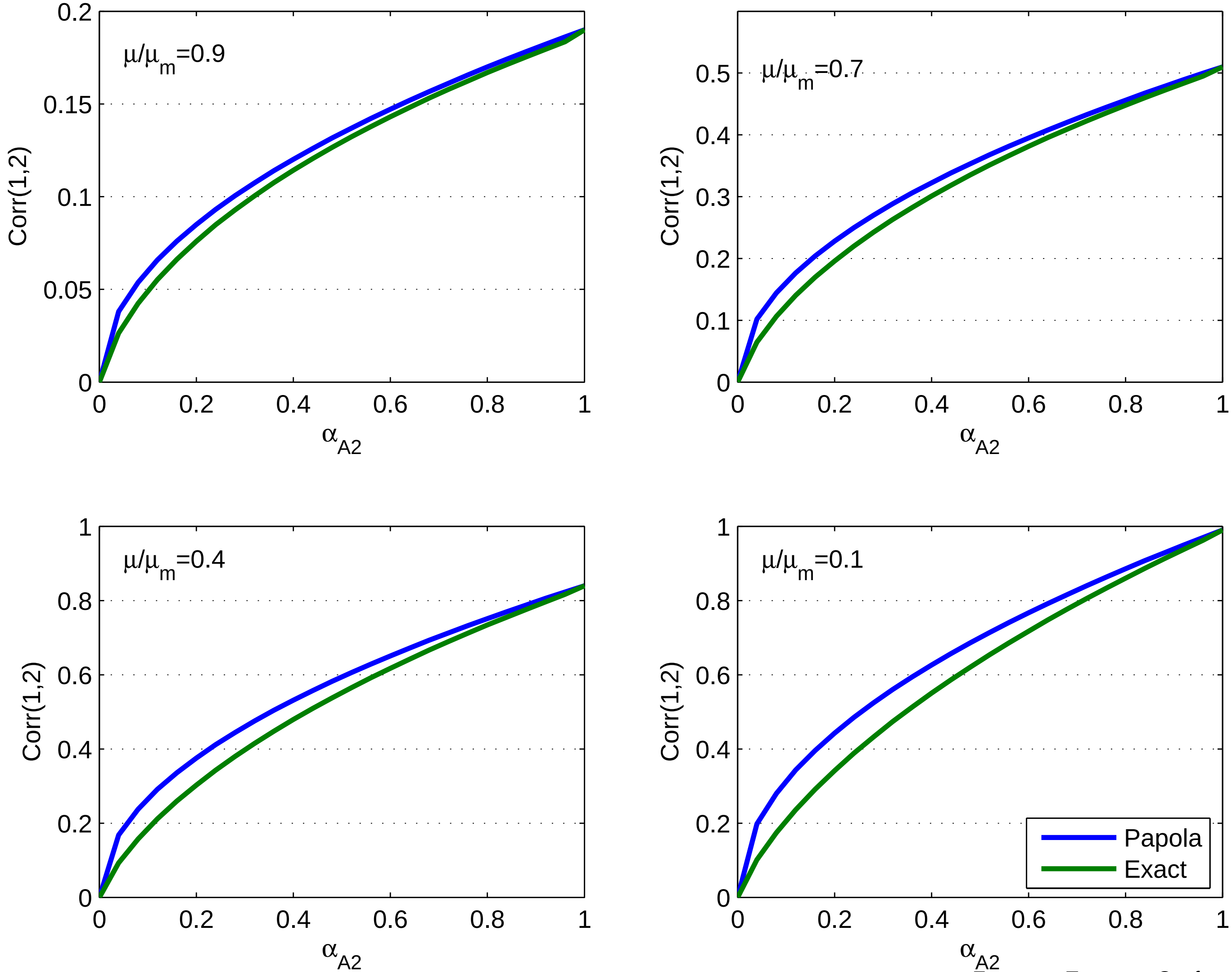
- Assume a full specification: JM α's, M μ's
- #equations = #unknowns if $M = J/2$.

Example 1



- Scale parameters equal: $\mu_A = \mu_B = \mu_m$
- $\alpha_{A1} = \alpha_{B3} = 1$
- $\alpha_{B2} = 1 - \alpha_{A2}$

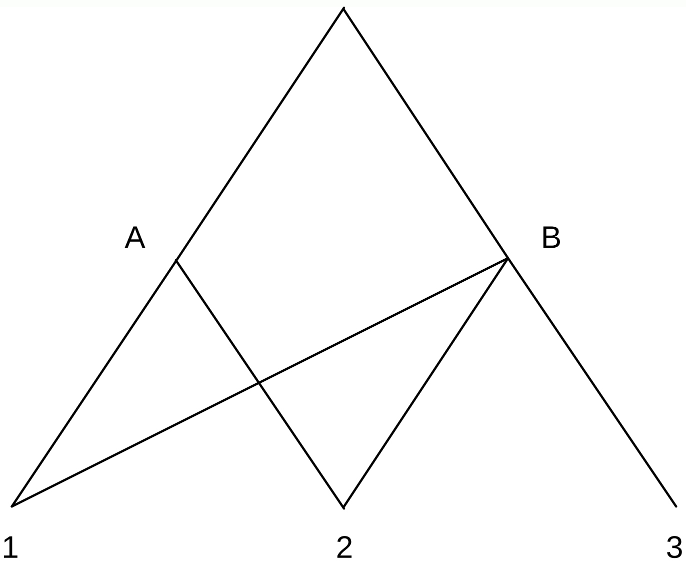
Example 1


μ/μm=0.9
μ/μm=0.7
μ/μm=0.4
μ/μm=0.1
Corr(1,2)
αA2
Papola
Exact
0
0.05
0.1
0.15
0.2
0.2
0.4
0.6
0.8
1

Example 1

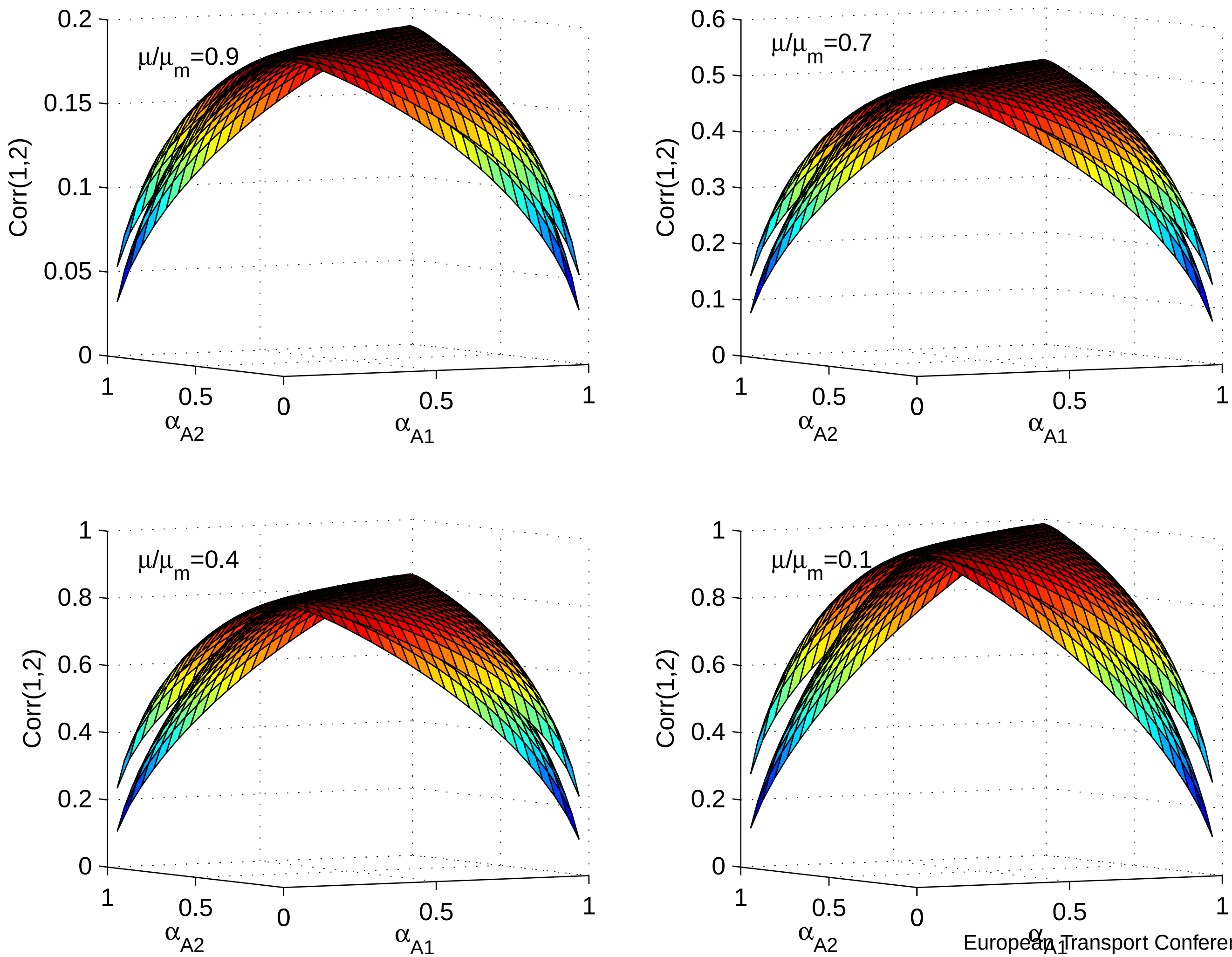
- Papola's conjecture overestimates the correlation
- Overestimation increases with μ_m
- Limit cases $\alpha_{A2} = 0$ and $\alpha_{A1} = 0$ are exact.

Example 2



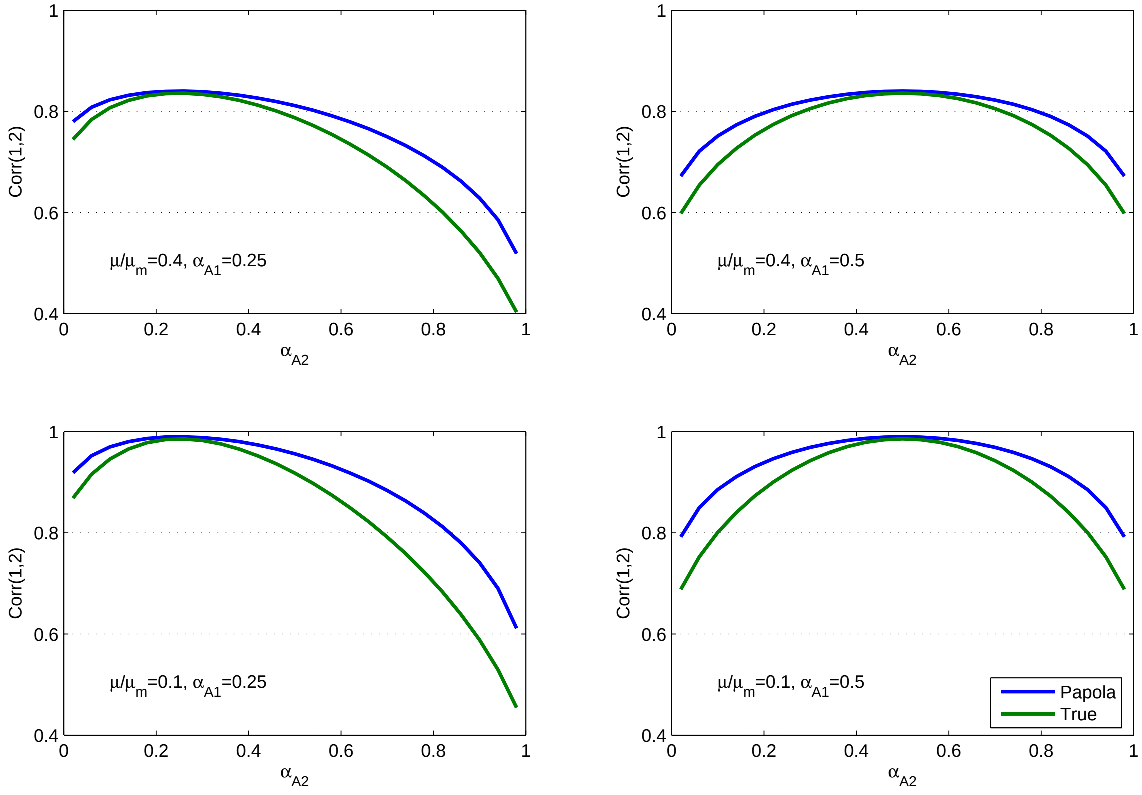
- Scale parameters equal: $\mu_A = \mu_B = \mu_m$
- $\alpha_{B3} = 1$
- $\alpha_{B1} = 1 - \alpha_{A1}$
- $\alpha_{B2} = 1 - \alpha_{A2}$

Example 2



μ/μm=0.9
Corr(1,2)
αA2
αA1
μ/μm=0.7
Corr(1,2)
αA2
αA1
μ/μm=0.4
Corr(1,2)
αA2
αA1
μ/μm=0.1
Corr(1,2)
αA2
αA1

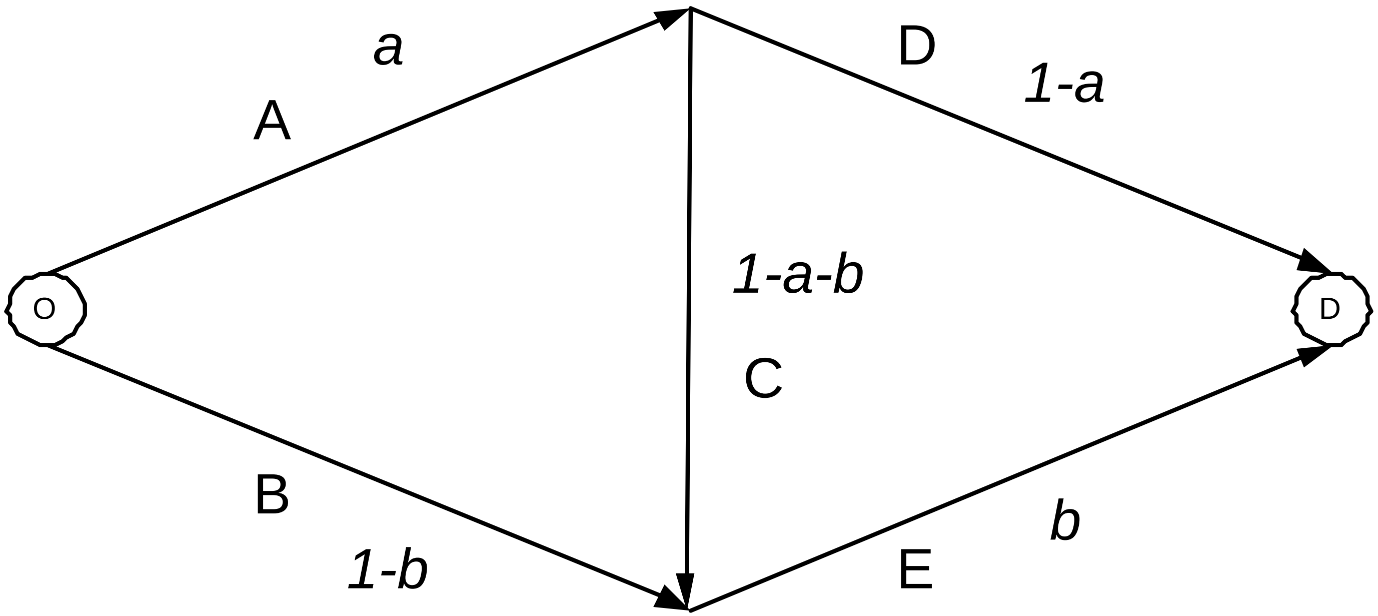
Example 2


Corr(1,2)
μ/μ_m=0.4, α_{A1}=0.25
μ/μ_m=0.4, α_{A1}=0.5
μ/μ_m=0.1, α_{A1}=0.25
μ/μ_m=0.1, α_{A1}=0.5
α_{A2}
Papola
True
0
0.2
0.4
0.6
0.8
1

Example 2

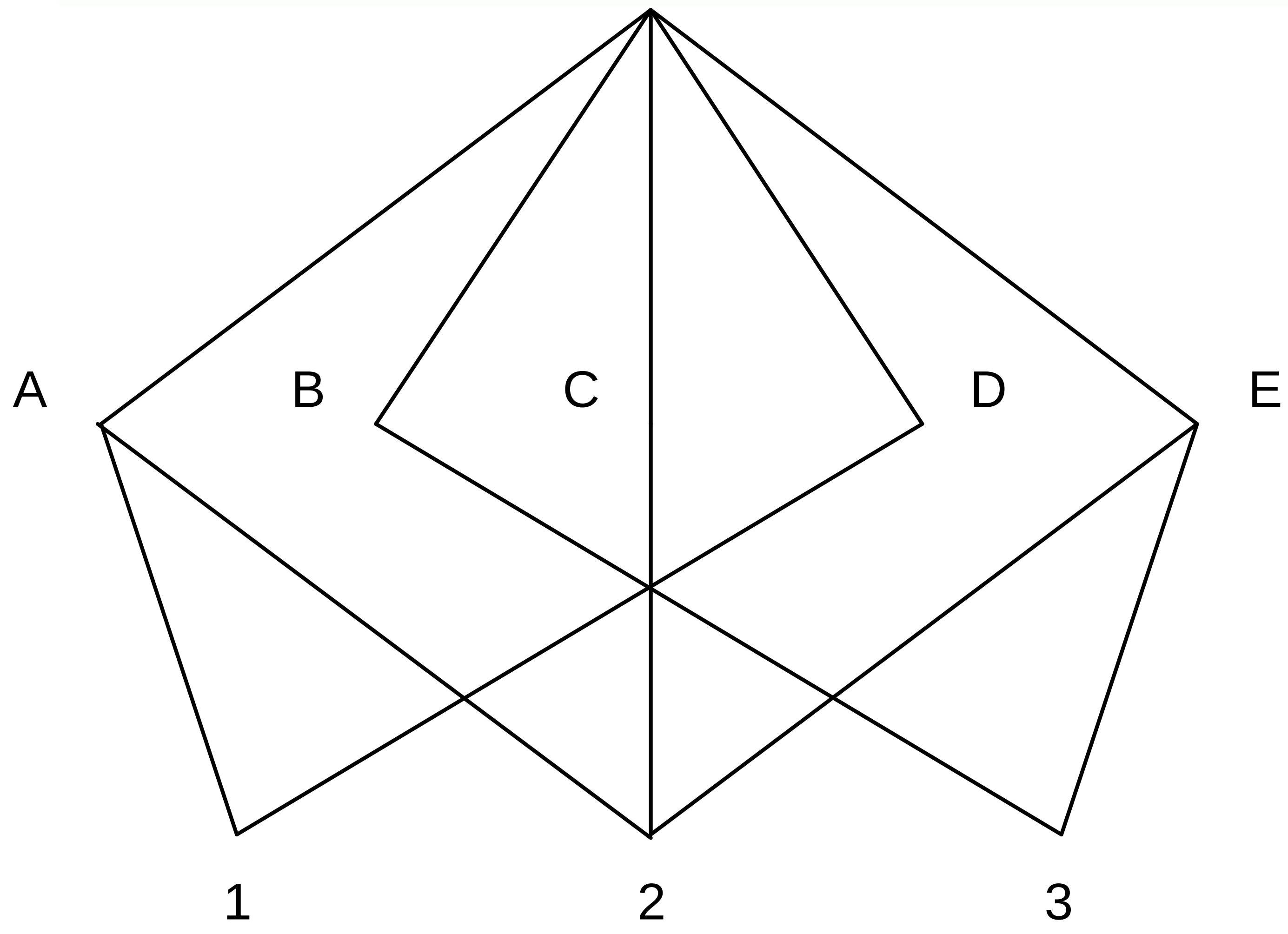
- Similar conclusions
- Papola's conjecture exact for $\alpha_{A1} = \alpha_{A2}$, where the model is equivalent to a NL.

Route choice



Correlation matrix: $\begin{pmatrix} 1 & & \\ a & 1 & \\ 0 & b & 1 \end{pmatrix}$

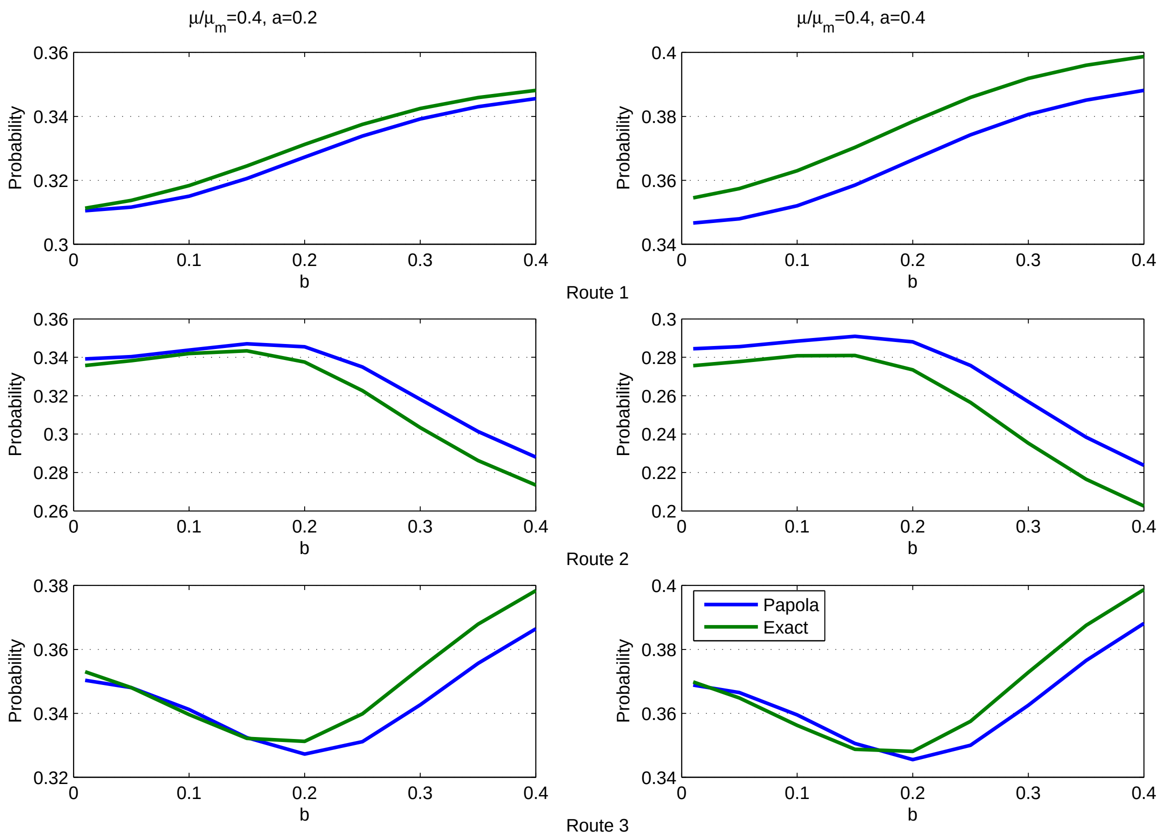
Route choice



Route choice

- Parameters: 7 α's, 1 μ/μ_m
- Equations: 2 from the matrix, 3 from the normalization
- Missing: 3 equations.
- We arbitrarily set
  - $\alpha_{A2} = 1/3$
  - $\alpha_{C2} = 1/3$
  - $\mu/\mu_m = 0.4$
- Compare the probabilities

Route choice



Conclusions

- Normalization of CNL: formal motivation
- Correlation structure:
  - correct formulation
  - comparison with Papola's conjecture
  - Advise: use the correct formulation
  - Other issues: see next paper by Papola...