Divers aspects de la modélisation de la demande en transport

Michel Bierlaire

Laboratoire Transport et Mobilité, Ecole Polytechnique Fédérale de Lausanne

Laboratoire Transport et Mobilité

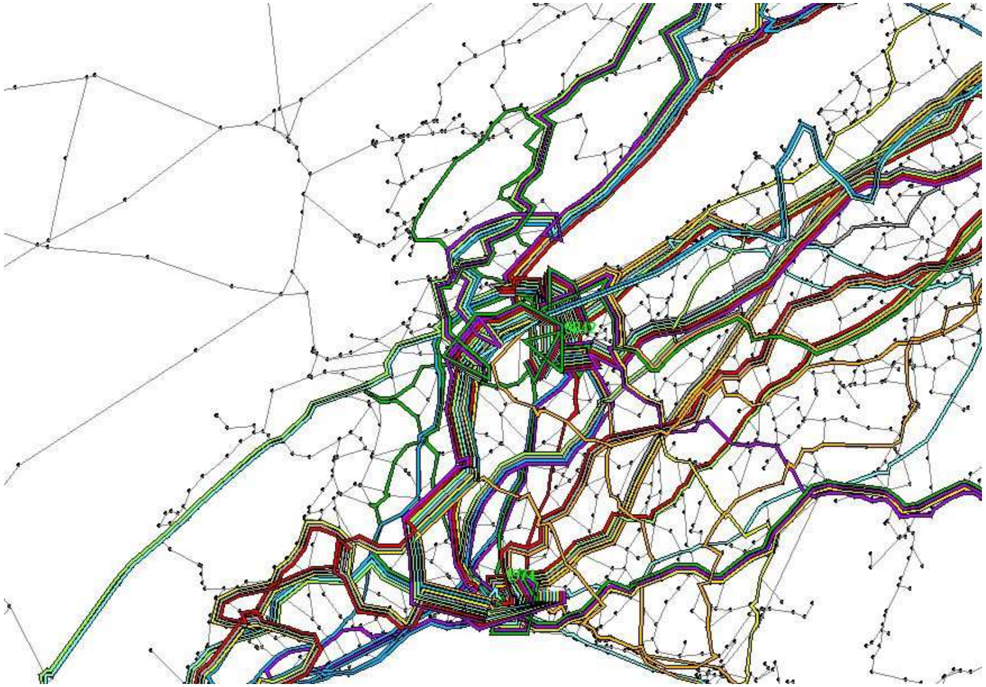
- TRANSP-OR
- Créé le 1er juillet 2006
- transp-or.epfl.ch
- Membres
  - Michel Bierlaire, directeur
  - 6 chercheurs seniors
  - 7 doctorants
  - 1 secrétaire
  - 1 informaticien

Quelques projets

- Planification des transports de la région lausannoise
  - Service de la mobilité du canton de Vaud
  - Transports publics lausannois

Quelques projets

- Analyse des choix d'itinéraires pour la tarification de la mobilité
  - Office fédéral des routes
  - IVT- ETHZ / Usi-Lugano



Quelques projets

- Rétablissement de rupture pour des lignes aériennes
  - CTI
  - APM Technologies

Quelques projets

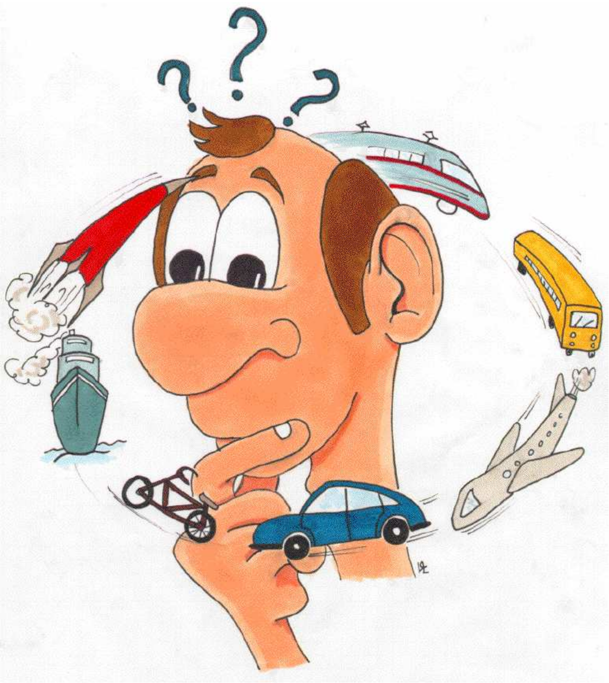
- Modèles d'aménagement du territoire et de transports pour la région bruxelloise
  - STRATEC SA, Belgique
  - University of Washington

Quelques projets

- Mouvements des piétons et des foules
  - Fonds National de la Recherche
  - Institut du Traitement des Signaux, EPFL
  - Laboratoire TOPO, INTER, EPFL

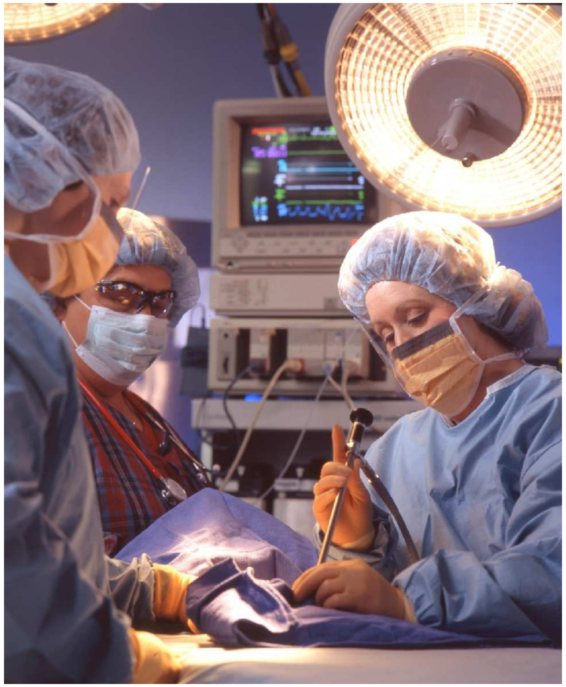
Quelques projets

- Méthodes d’optimisation pour les modèles de comportement
  - Fonds National de la Recherche



Quelques projets

- Modélisation et simulation de la congestion
  - Fonds National de la Recherche
  - Hôpitaux Universitaires de Genève



Outline

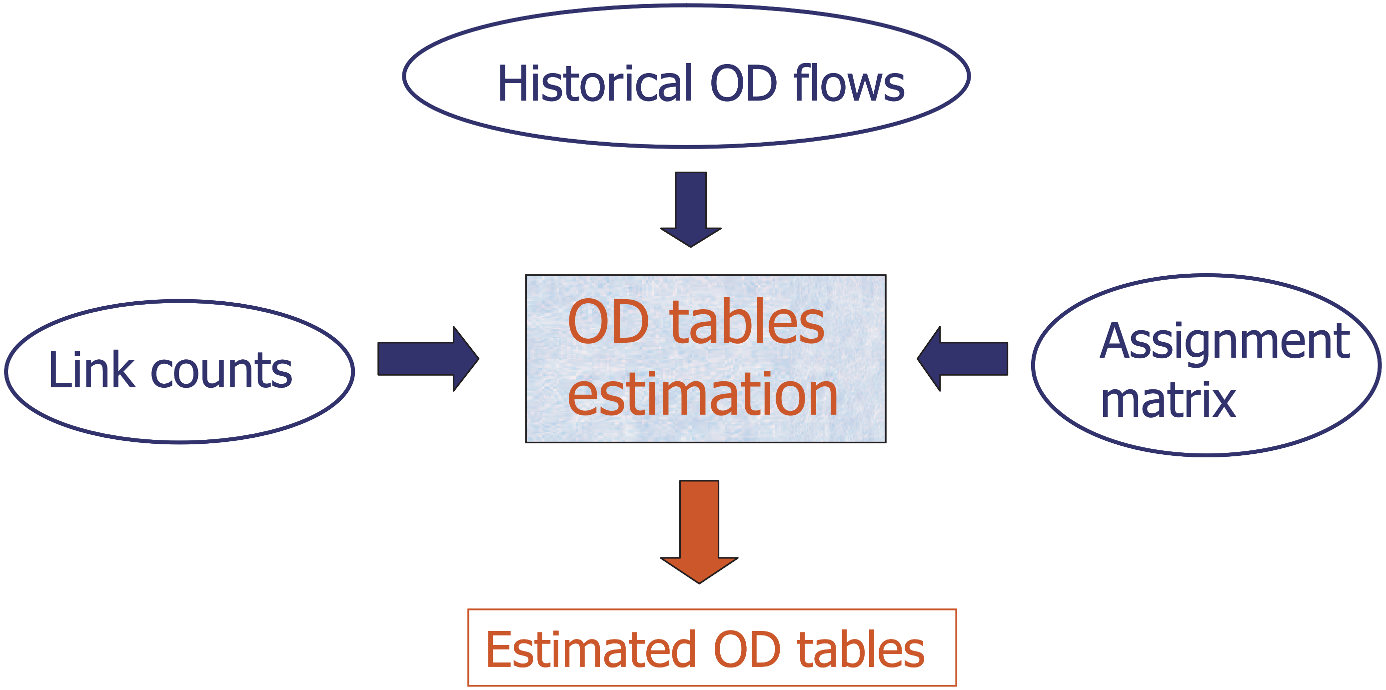
- OD estimation and prediction
- Network GEV
- Route choice

Estimation and prediction of OD tables

- Problem formulation
  - Transition and measurement equations
  - Least-square problem
- Classical and new resolution algorithms
  - Kalman Filter method
  - Lsqr algorithm
- Numerical results
  - A synthetic network
  - Irvine network

OD tables definition

OD tables x_h: amount of demand per origin, destination and departure time.



Formulation

→ Ashok and Ben-Akiva, 1993

Transition equation:

$$\partial x_h = \sum_{p=h-q'}^{h-1} f_h^p \partial x_p + w_h,$$

- $\partial x_h = x_h - x_h^H$ deviations between actual and historical OD flows
- f_h^p auto-regressive process
- w_h random variable capturing the error
- q' number of former time intervals influencing ∂x_h

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Formulation

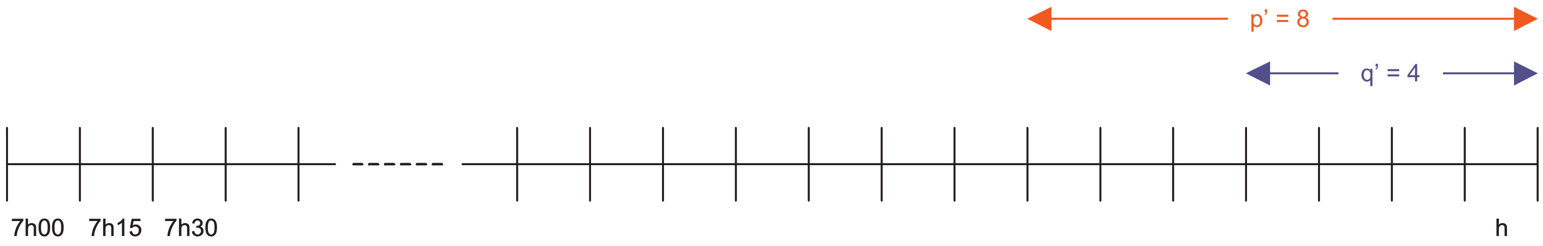
$\rightarrow$ Ashok and Ben-Akiva, 1993

Measurement equation:

$$\partial y_h = \sum_{p=h-p'}^{h} a_h^p \partial x_p + v_h,$$

- $\partial y_h = y_h - \sum_{p=h-p'}^{h} a_h^p x_p^H$ deviations of the counts
- a_h^p assignment matrix for interval h departing during time interval p
- v_h random variable capturing the error
- p' maximum number of time intervals needed to travel between any OD pair

Least-squares formulation



OD estimation and prediction is a least-squares problem

$$\min_{x} \sum_{i=1}^{h} \|C_i x - d_i\|$$

Characteristics:

- Size increases continually
- Sparsity

$\Longrightarrow$ solution algorithms: incremental or direct!

Kalman filter method

INCREMENTAL ALGORITHM USING NORMAL EQUATIONS $\longrightarrow$ Kalman, 1960

Advantages

- Incremental
- Discard past data
- Constant size

Drawbacks

- Ignore the sparsity
- Limited to medium-scale problems
- Constant computation time, even with a good starting solution

LSQR algorithm

MATRIX-FREE ALGORITHM $\longrightarrow$ Paige and Saunders, 1982

Advantages

- Exploits the sparsity
- Numerically efficient, especially for large-scale problems
- Iterative method: exploits starting point
- Theoretically converge in n iterations, but ...

Drawbacks

- Size increases continually
- Needs past data

LSQR algorithm

MATRIX-FREE ALGORITHM → Paige and Saunders, 1982

Advantages

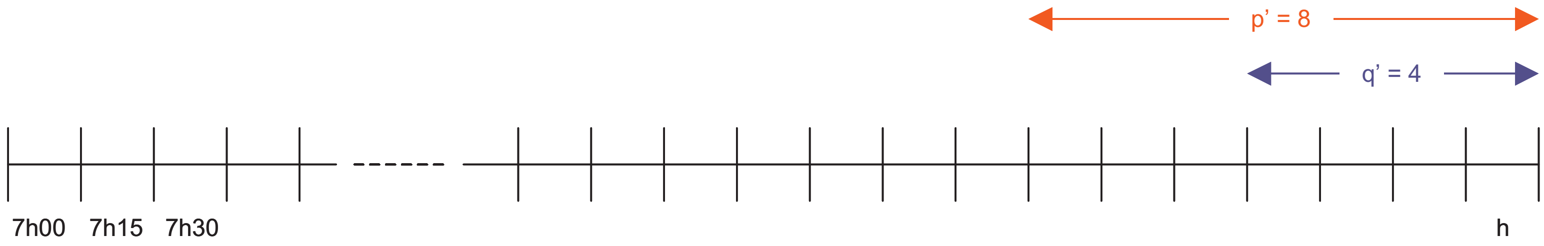
- Exploits the sparsity
- Numerically efficient, especially for large-scale problems
- Iterative method: exploits starting point
- Theoretically converge in n iterations, but ...

Drawbacks

- Size increases continually $\Longrightarrow$ adaptation
- Needs past data

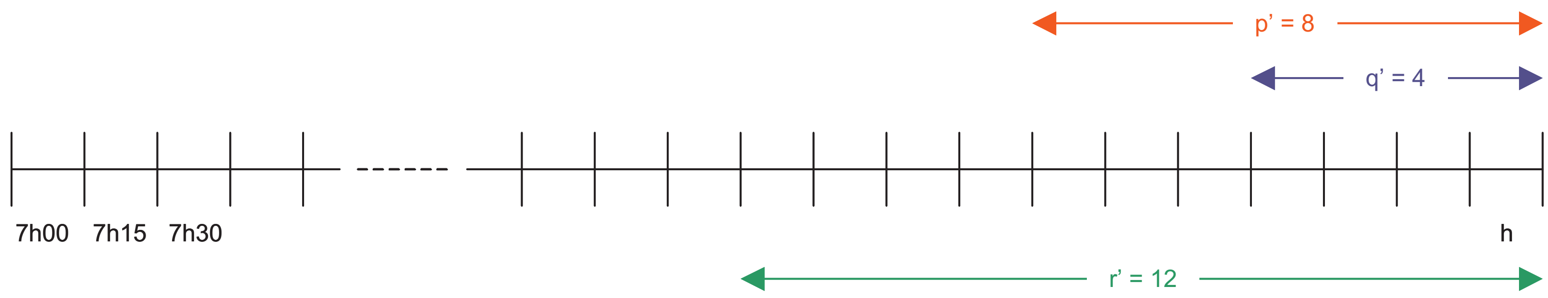
LSQR algorithm adaptation

→ Paige and Saunders, 1982



LSQR algorithm adaptation

→ Paige and Saunders, 1982



Adaptation:

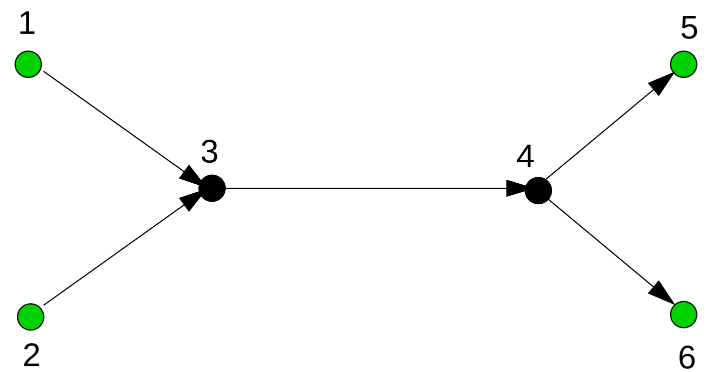
- Take into account only equations from $h - r'$ to h

Consequence:

- ignores the estimation error of time intervals 1 to $h - r' - 1$ by not propagating the variance-covariance

Numerical impact of the adaptation

Numerical comparison on a synthetic network

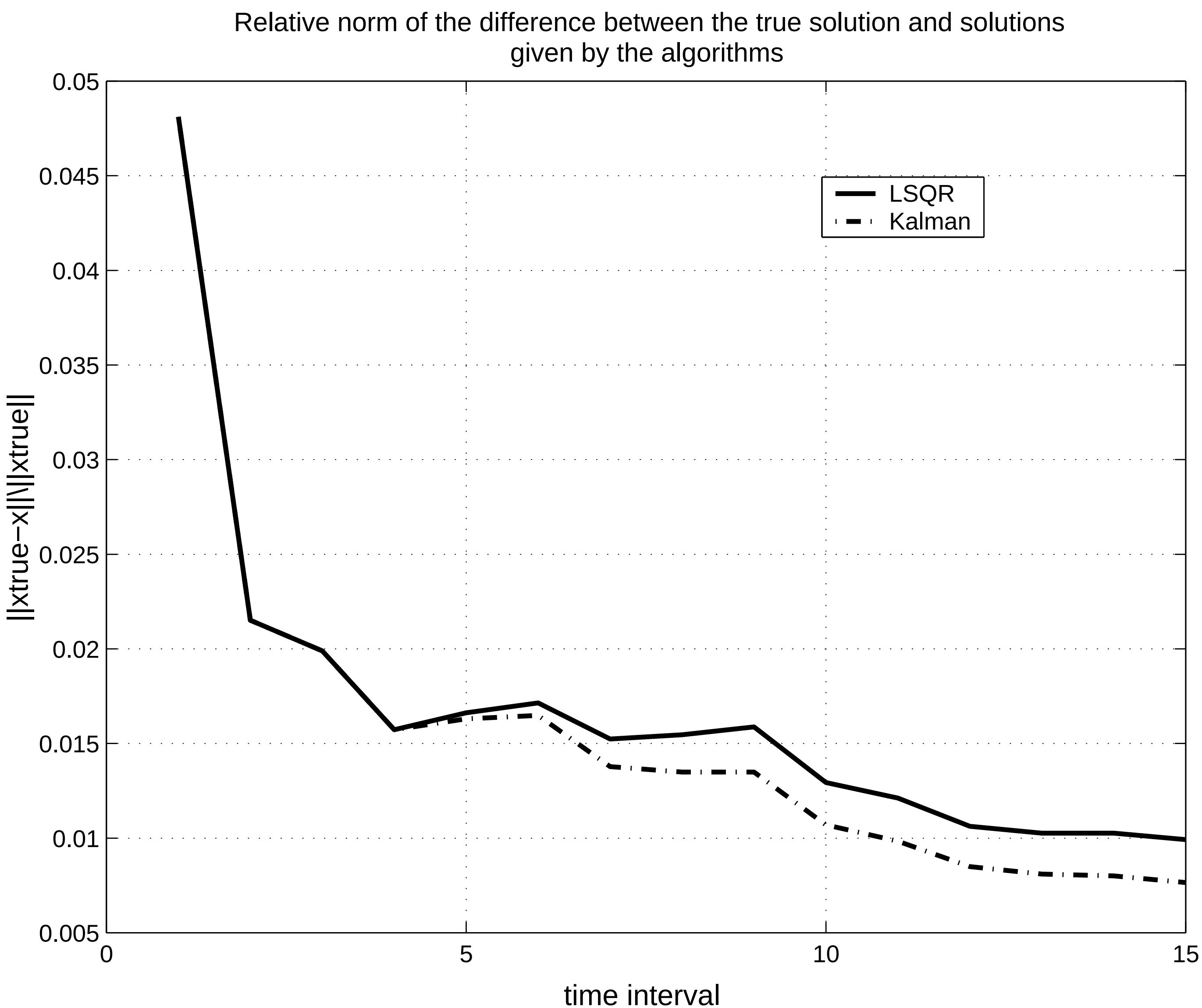


- True OD tables known
- Counts given by assignment of disturbed OD flows
- 15 time intervals
- $p' = 3$, $q' = 2$, $r' = 3$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Numerical impact of this adaptation

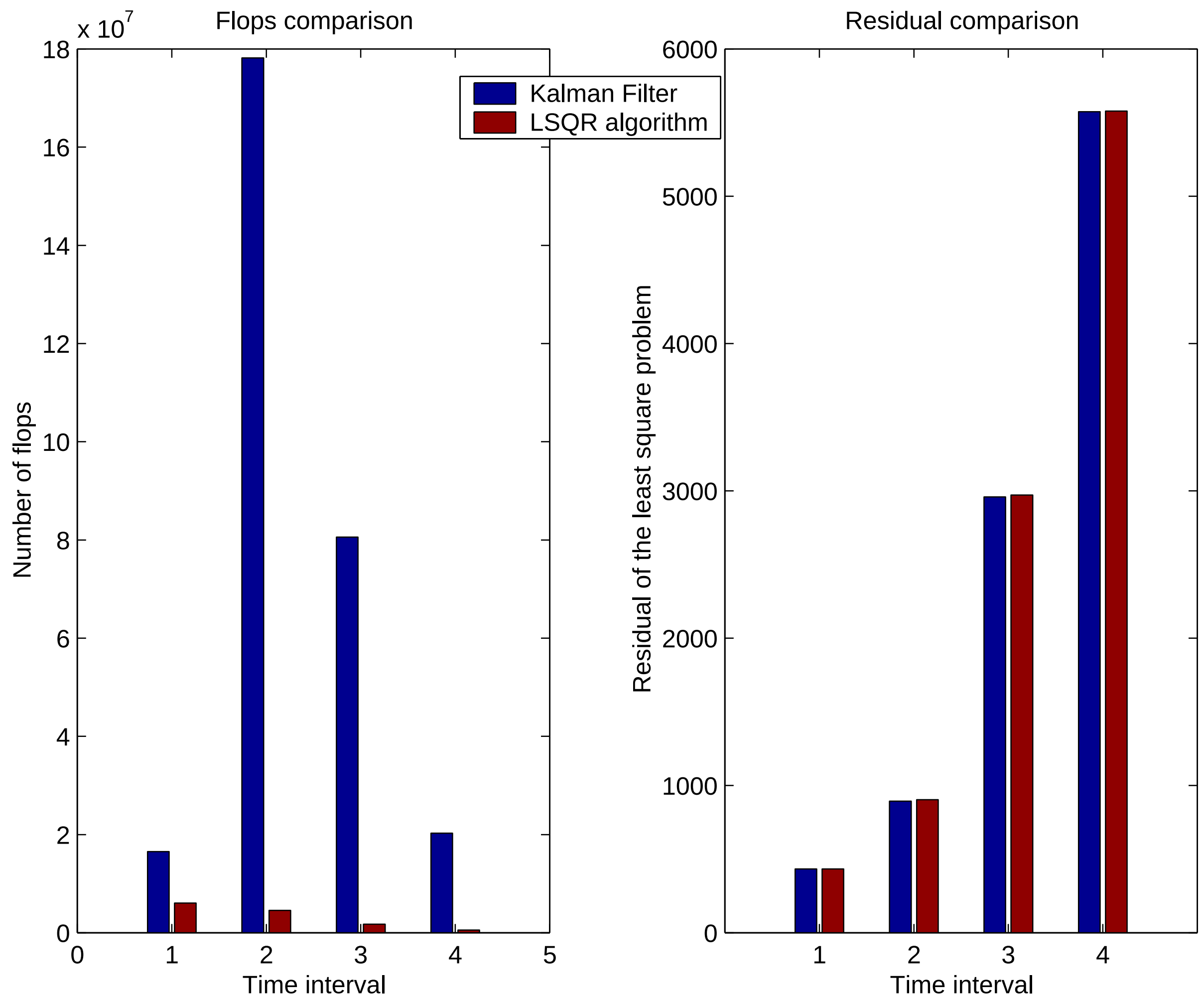


Numerical Results

Irvine Network (input data obtain from DynaMIT)

- 627 OD pairs
- 217 links counts
- Simulation between 7:15am and 8h15am
- 4 intervals of 15 minutes
- $p' = 0$, $q' = r' = 1$

Numerical Results



Conclusion

- General Formulation
  - difference between formulation and resolution algorithm
  - easily extendible
- Kalman Filter
  - well adapted for small to medium size problems
- LSQR
  - significant computation improvement
  - exploits good starting points

Bierlaire, M., and Crittin, F. (2004). An efficient algorithm for real-time estimation and prediction of dynamic OD table **Operations Research** 52(1):116-127.

2006 Best paper award, Transportation Science & Logistics Society, INFORMS

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

- OD estimation and prediction
- Network GEV
- Route choice

MEV models

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

$$G : \mathbb{R}^J \rightarrow \mathbb{R}$$

From G, we can build

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

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

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

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 $G_i = \frac{\partial G}{\partial x_i}$. 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

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

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: $G(x) = \sum_{i=1}^{J} x_i^{\mu}$

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

G complies with the theory

MEV models

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

$$
\begin{aligned}
F(x_1, \ldots, x_J) &= e^{-G(e^{-x_1}, \ldots, e^{-x_J})} \\
&= e^{-\sum_{i=1}^{J} e^{-\mu x_i}} \\
&= \prod_{i=1}^{J} e^{-e^{-\mu x_i}}
\end{aligned}
$$

Product of i.i.d EV

Multinomial Logit Model

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

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

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

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

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

MEV models

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

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

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.

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

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

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

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

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

- 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

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

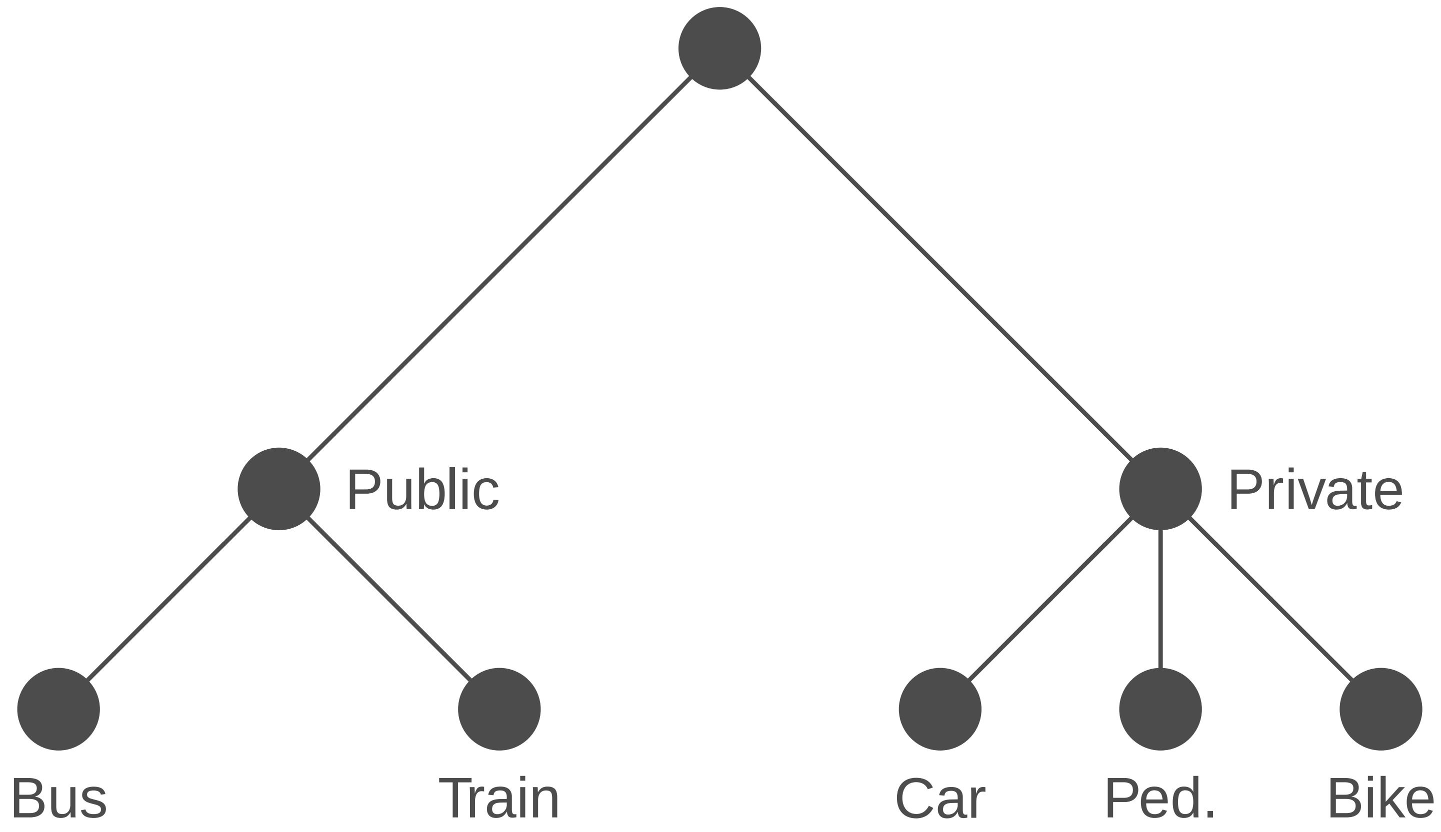
with $\frac{\mu}{\mu_m}\leq 1$, $\alpha_{jm}\geq 0$, and $\forall j,\exists m$ s.t. $\alpha_{jm}>0$

- Generalization of the nested-logit model

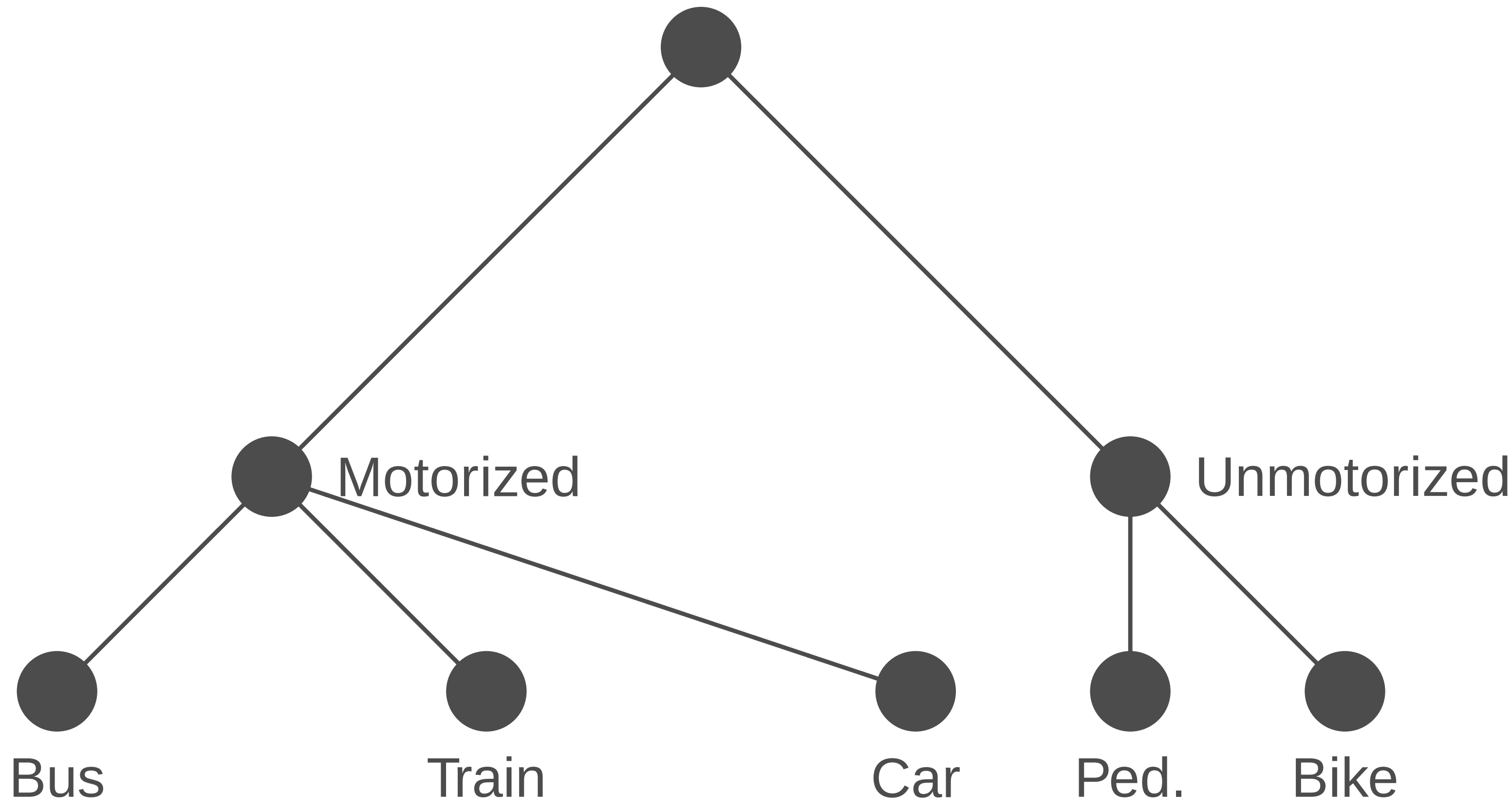
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

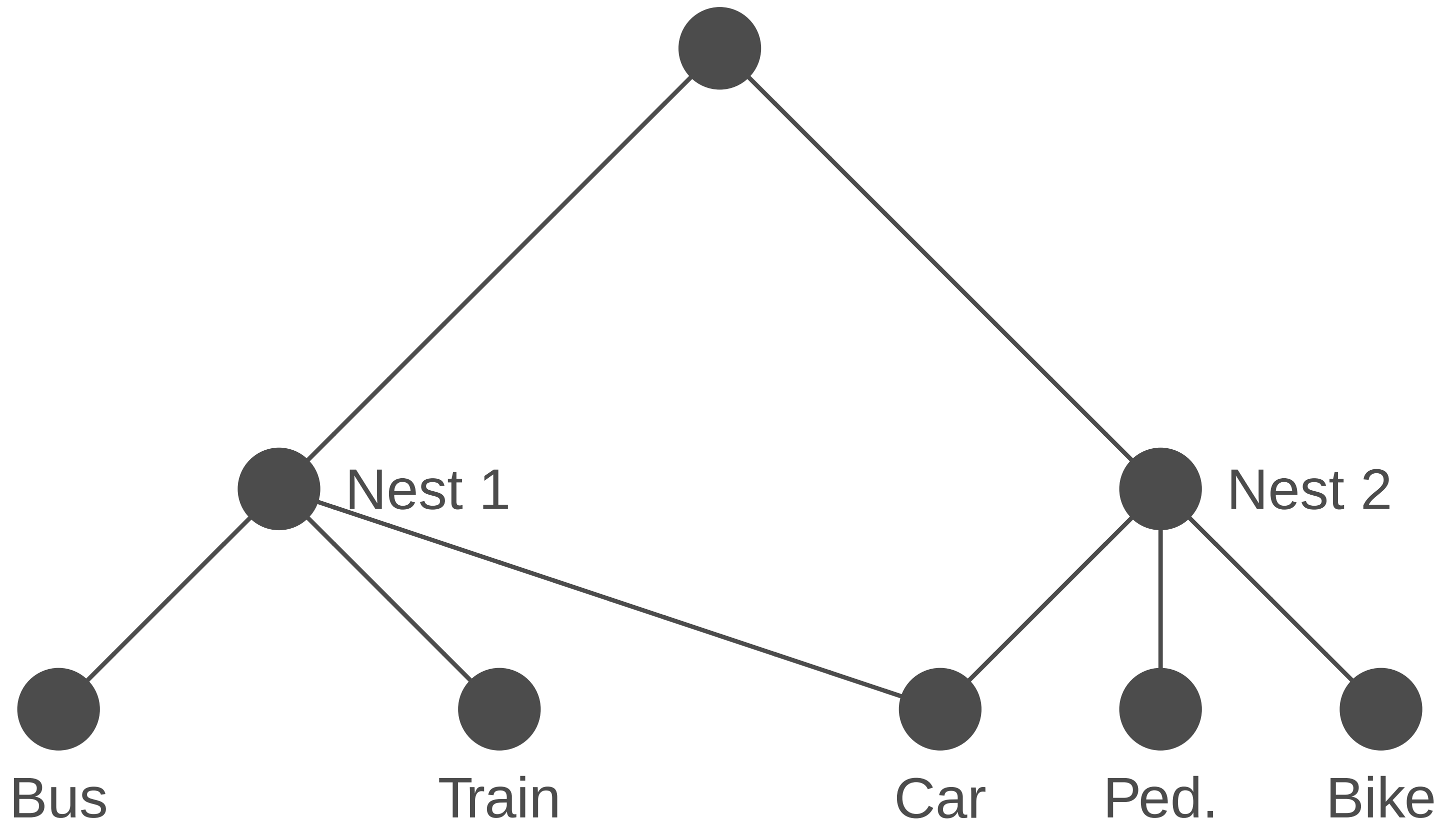
Nested Logit Model



Nested Logit Model



Cross-Nested Logit Model



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

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

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

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

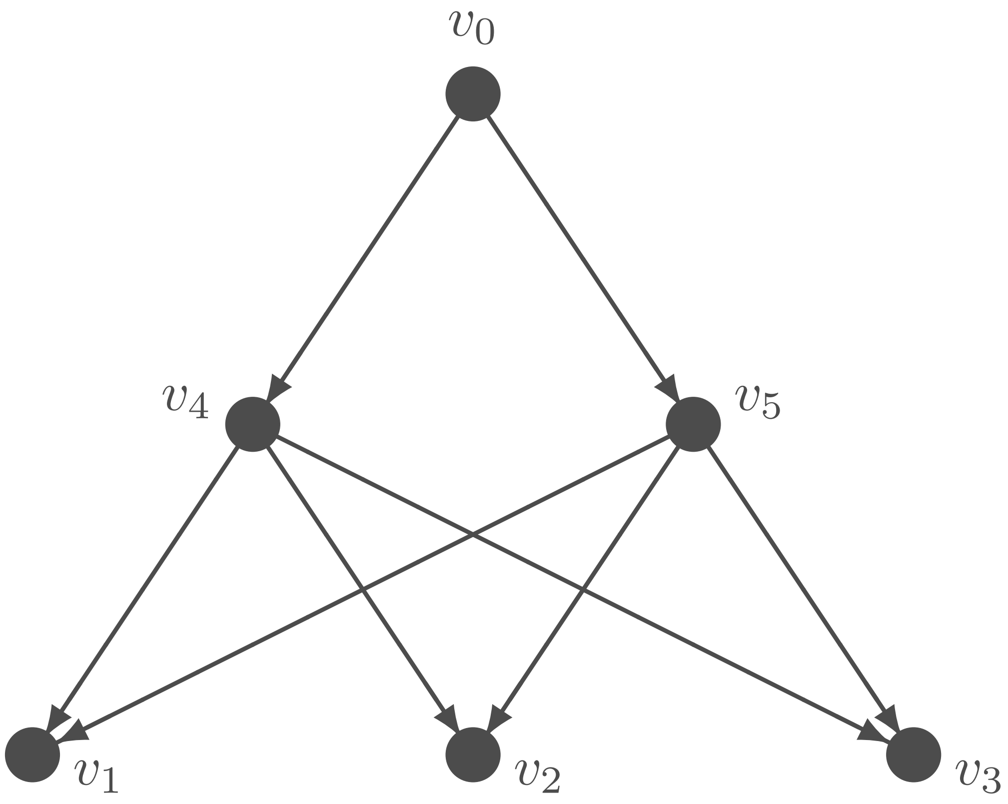
Network MEV

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 $\prod_{k=1}^{P} \alpha_{(i_{k-1}, i_k)} > 0$.

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:

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

For all others:

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

Theorem

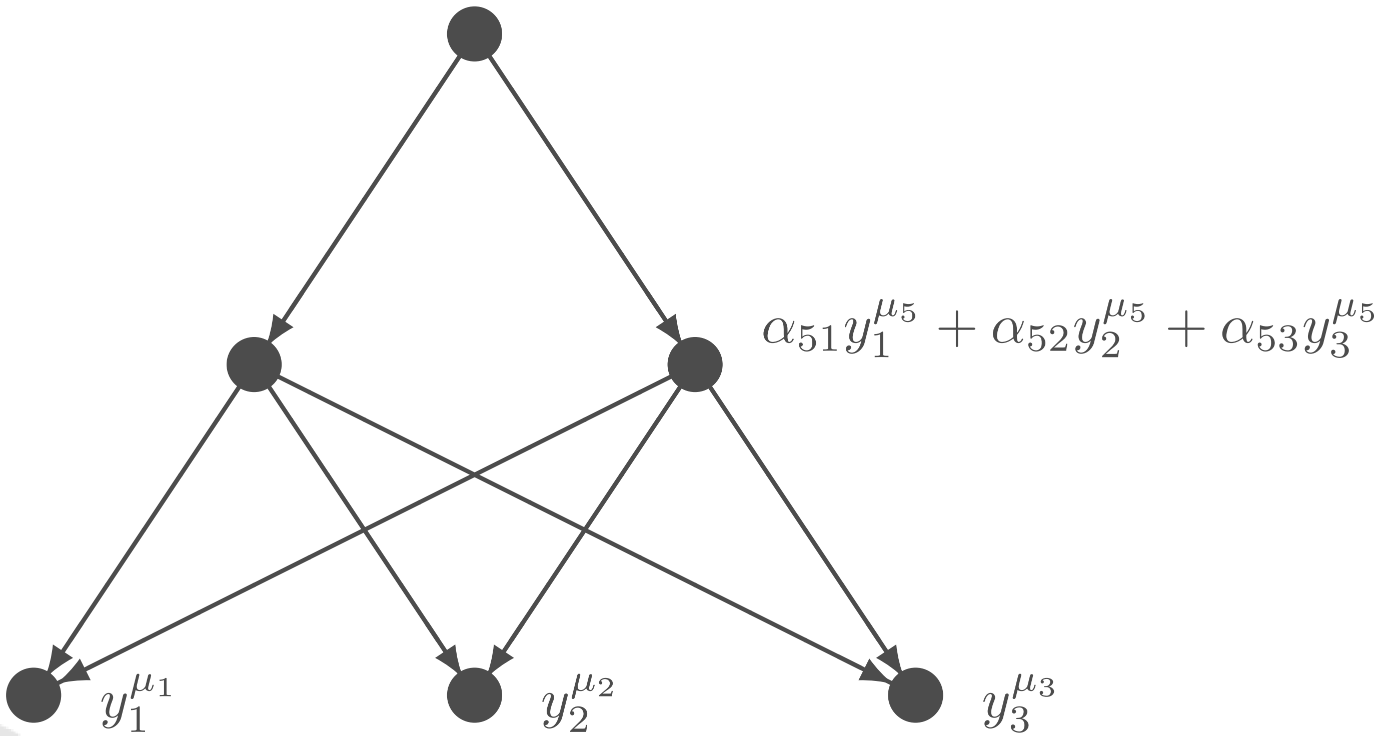
If all $G^j(x)$ are MEV generating functions, so is G^i

Network MEV

Example: Cross-Nested Logit

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

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



Network MEV

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

Advantages :

- Easy to design
- No more proof necessary

Daly, A., and Bierlaire, M. (2006). A General and Operational Representation of Generalised Extreme Value Models, **Transportation Research Part B: Methodological** 40(4):285-305.

Outline

- OD estimation and prediction
- Network GEV
- Route choice

Recent developments in route choice modeling
with **Emma Frejinger**

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

- Introduction
  - Problem description
  - Existing models
- Subnetwork approach
- Latent route choice model
- Future work

Route choice problem

Given a transportation network composed of nodes, links, origin and destinations.
For a given transportation mode and origin-destination pair, which is the chosen route?

Applications

- Intelligent transportation systems
- GPS navigation
- Transportation planning

Issues

- The choice set is unknown
- There are many (feasible) alternatives available
- The alternatives are often highly correlated due to overlapping paths
- Choice data is difficult to obtain

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Existing Approaches

- Assumption: Travelers use the shortest (with regard to any arbitrary generalized cost) route among all
  - Behaviorally unrealistic
- Random utility models (discrete choice models)

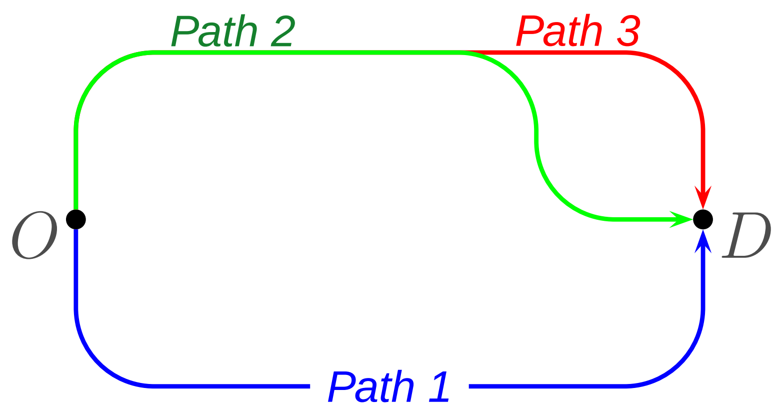
Existing Approaches - MNL

- Random terms are assumed to be i.i.d. Extreme Value

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

Alternatives are assumed to be independent. This assumption is (in general) not valid in a route choice context due to overlapping paths.

Existing Approaches



Travel time is the only considered attribute and

$V_1 = V_2 = V_3 = T$ then

$P(1|\{1,2,3\}) = P(2|\{1,2,3\}) = P(3|\{1,2,3\}) = \frac{1}{3}$

- Unrealistic path choice probabilities for correlated alternatives (overlapping paths)

Existing Approaches

- Few models explicitly capturing correlation have been used on large-scale route choice problems
  - C-Logit (Cascetta et al., 1996)
  - Path Size Logit (Ben-Akiva and Bierlaire, 1999)
  - Link-Nested Logit (Vovsha and Bekhor, 1998)
  - Logit Kernel model adapted to route choice situation (Bekhor et al., 2002)
- Probit model (Daganzo, 1977) permits an arbitrarily covariance structure specification but cannot be applied in a large-scale route choice context

Existing Approaches

- Link based path-multilevel logit model (Marzano and Papola, 2005)
  - Illustrated on simple examples and not estimated on real data

Subnetwork approach

Emma Frejinger and Michel Bierlaire

Subnetworks

How can we explicitly capture the most important correlation structure without considerably increasing the model complexity?

Subnetworks

How can we explicitly capture the most important correlation structure without considerably increasing the model complexity?

- Which are the behaviorally important decisions?

Subnetworks

How can we explicitly capture the most important correlation structure without considerably increasing the model complexity?

- Which are the behaviorally important decisions?
- Our hypothesis: choice of specific parts of the network (e.g. main roads, city center)
- Concept: subnetwork

Subnetworks

- Subnetwork approach designed to be behaviorally realistic and convenient for the analyst
- Subnetwork component is a set of links corresponding to a part of the network which can be easily labeled
- Paths sharing a subnetwork component are assumed to be correlated even if they are not physically overlapping

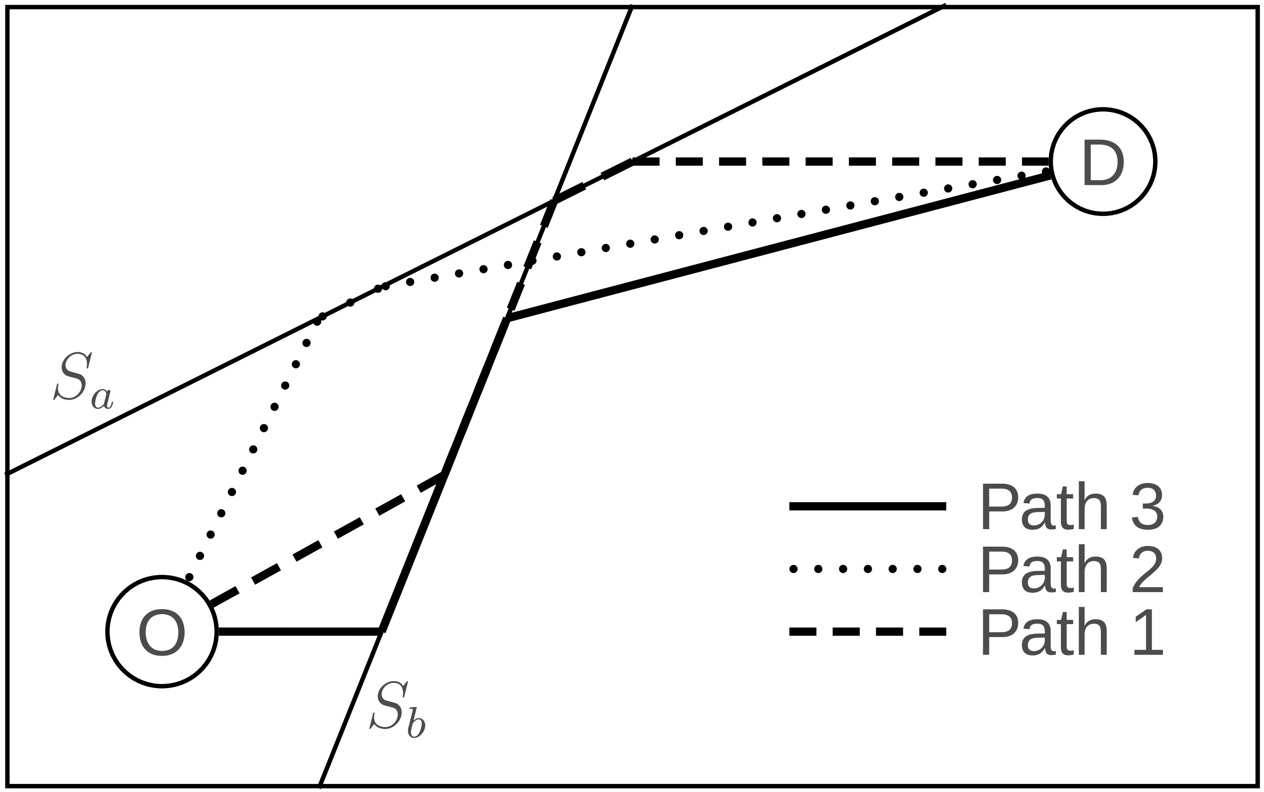
Subnetworks - Methodology

- Factor analytic specification of an error component model (based on model presented in Bekhor et al., 2002)

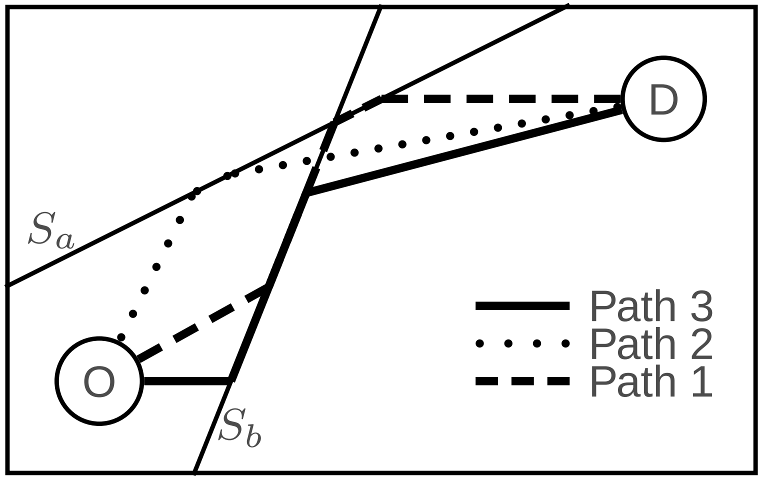
$$\mathbf{U}_n = \beta^T \mathbf{X}_n + \mathbf{F}_n \mathbf{T} \zeta_n + \nu_n$$

  - $\mathbf{F}_{n\ (J\text{x}Q)}$: factor loadings matrix
  - $(f_n)_{iq} = \sqrt{l_{niq}}$
  - $\mathbf{T}_{(Q\text{x}Q)} = \text{diag}\,(\sigma_1, \sigma_2, \ldots, \sigma_Q)$
  - $\zeta_{n\ (Q\text{x}1)}$: vector of i.i.d. N(0,1) variates
  - $\nu_{(J\text{x}1)}$: vector of i.i.d. Extreme Value distributed variates

Subnetworks - Example



Subnetworks - Example



$$U_1 = \beta^T X_1 + \sqrt{l_{1a}}\sigma_a\zeta_a + \sqrt{l_{1b}}\sigma_b\zeta_b + \nu_1$$

$$U_2 = \beta^T X_2 + \sqrt{l_{2a}}\sigma_a\zeta_a + \nu_2$$

$$U_3 = \beta^T X_3 + \sqrt{l_{3b}}\sigma_b\zeta_b + \nu_3$$

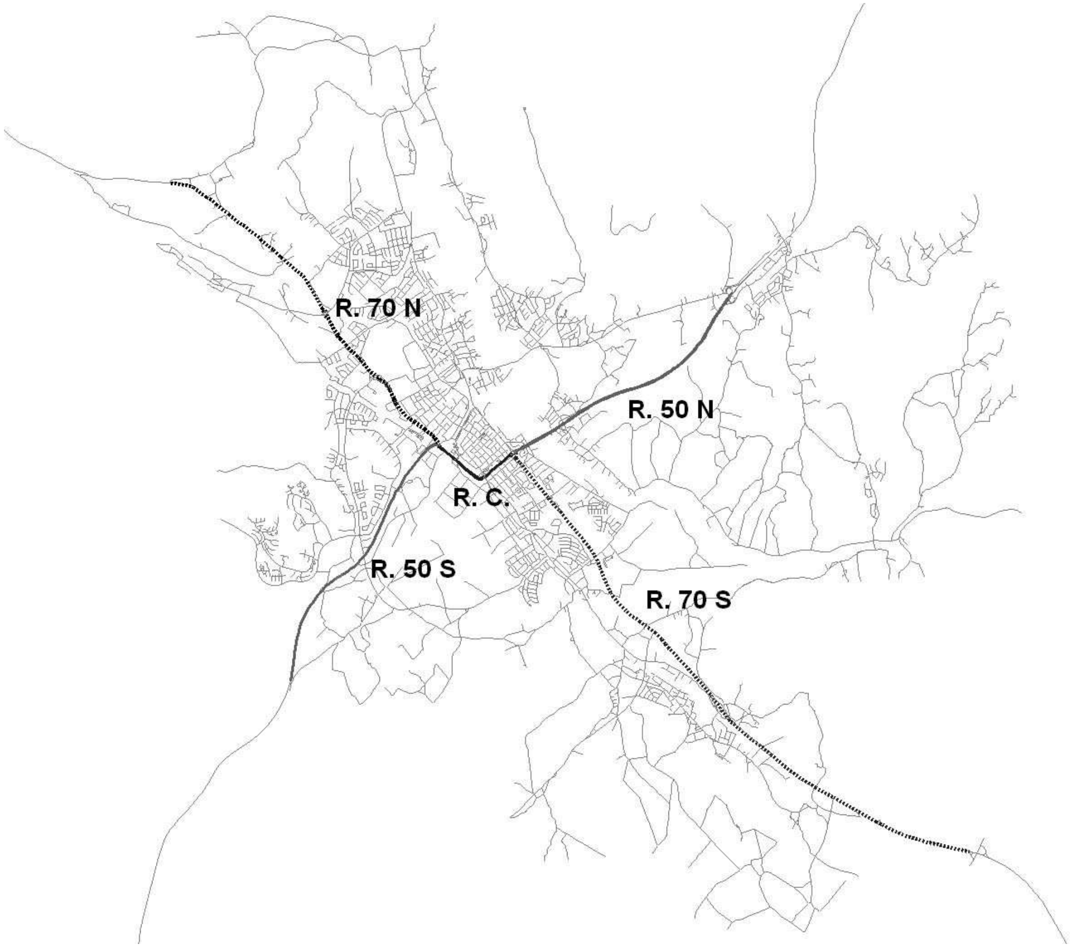
$$\mathbf{F}\mathbf{T}\mathbf{T}^T\mathbf{F}^T =$$

$$\begin{bmatrix} l_{1a}\sigma_a^2 + l_{1b}\sigma_b^2 & \sqrt{l_{1a}}\sqrt{l_{2a}}\sigma_a^2 & \sqrt{l_{1b}}\sqrt{l_{3b}}\sigma_b^2 \\ \sqrt{l_{1a}}\sqrt{l_{2a}}\sigma_a^2 & l_{2a}\sigma_a^2 & 0 \\ \sqrt{l_{3b}}\sqrt{l_{1b}}\sigma_b^2 & 0 & l_{3b}\sigma_b^2 \end{bmatrix}$$

Empirical Results

- The approach has been tested on three datasets: Boston (Ramming, 2001), Switzerland, and Borlänge
- Deterministic choice set generation
  Link elimination
- GPS data from 24 individuals
  2978 observations, 2179 origin-destination pairs
- Borlänge network
  3077 nodes and 7459 links
- BIOGEME (biogeme.epfl.ch, Bierlaire, 2003) has been used for all model estimations

Borlänge Road Network



Subnetwork Components

	R.50 S	R.50 N	R.70 S	R.70 N	R.C.
Component length [m]	5255	4966	11362	7028	1733
Nb. of Observations	173	153	261	366	209
Weighted Nb. of Observations (N_q) $N_q = \sum_{o \in O} \frac{l_{oq}}{L_q}$	36	88	65	73	116

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Model Specifications

- Six different models: MNL, PSL, EC_1, EC'_1, EC_2 and EC'_2
- EC_1 and EC'_1 have a simplified correlation structure
- EC'_1 and EC'_2 do not include a Path Size attribute
- Deterministic part of the utility

$$
\begin{aligned}
V_i = \; & \beta_{\text{PS}} \ln(\text{PS}_i) + \beta_{\text{EstimatedTime}} \text{EstimatedTime}_i + \\
& \beta_{\text{NbSpeedBumps}} \text{NbSpeedBumps}_i + \beta_{\text{NbLeftTurns}} \text{NbLeftTurns}_i + \\
& \beta_{\text{AvgLinkLength}} \text{AvgLinkLength}_i
\end{aligned}
$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Estimation Results

- Parameter estimates for explanatory variables are stable across the different models
- Path size parameter estimates

Parameter	PSL	EC_1	EC_2
Path Size	-0.28	-0.49	-0.53
Scaled estimate	-0.33	-0.53	-0.56
Rob. T-test 0	-4.05	-5.61	-5.91

- All covariance parameters estimates in the different models are significant except the one associated with R.50 S

Estimation Results

Model	Nb. σ Estimates	Nb. Estimated Parameters	Final L-L	Adjusted Rho-Square
MNL	-	12	-4186.07	0.152
PSL	-	13	-4174.72	0.154
EC_1 (with PS)	1	14	-4142.40	0.161
EC'_1	1	13	-4165.59	0.156
EC_2 (with PS)	5	18	-4136.92	0.161
EC'_2	5	17	-4162.74	0.156
1000 pseudo-random draws for Maximum Simulated Likelihood estimation 2978 observations Null log likelihood: -4951.11 BIOGEME (biogeme.epfl.ch) has been used for all model estimations.				

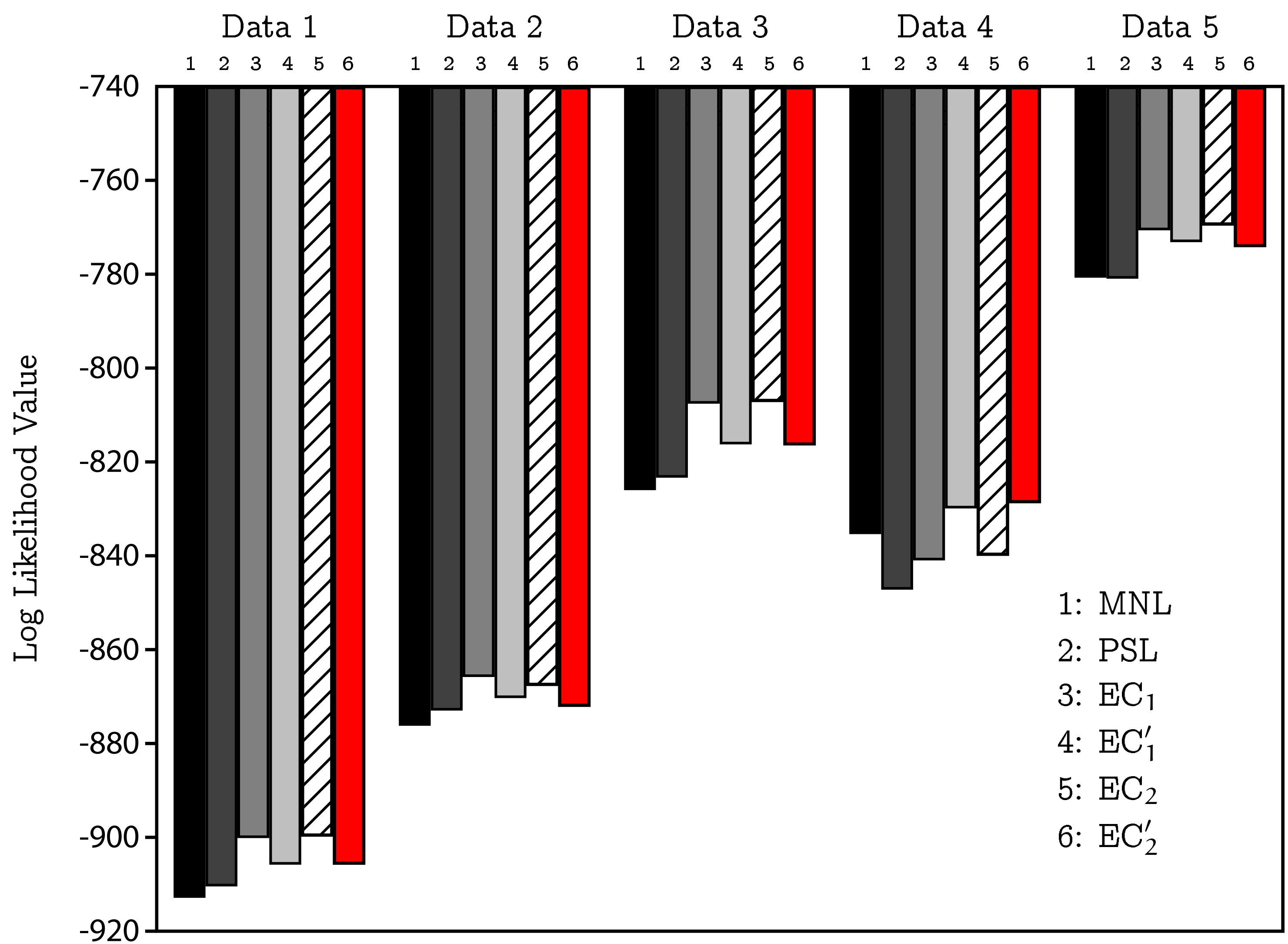
Forecasting Results

- Comparison of the different models in terms of their performance of predicting choice probabilities
- Five subsamples of the dataset
  - Observations corresponding to 80% of the origin destination pairs (randomly chosen) are used for estimating the models
  - The models are applied on the observations corresponding to the other 20% of the origin destination pairs
- Comparison of final log-likelihood values

Forecasting Results

- Same specification of deterministic utility function for all models
- Same interpretation of these models as for those estimated on the complete dataset
- Coefficient and covariance parameter values are stable across models

Forecasting Results



Conclusion - Subnetworks

- Models based on subnetworks are designed for route choice modeling of realistic size
- Correlation on subnetwork is explicitly captured within a factor analytic specification of an Error Component model
- Estimation and prediction results clearly shows the superiority of the Error Component models compared to PSL and MNL

Conclusion - Subnetworks

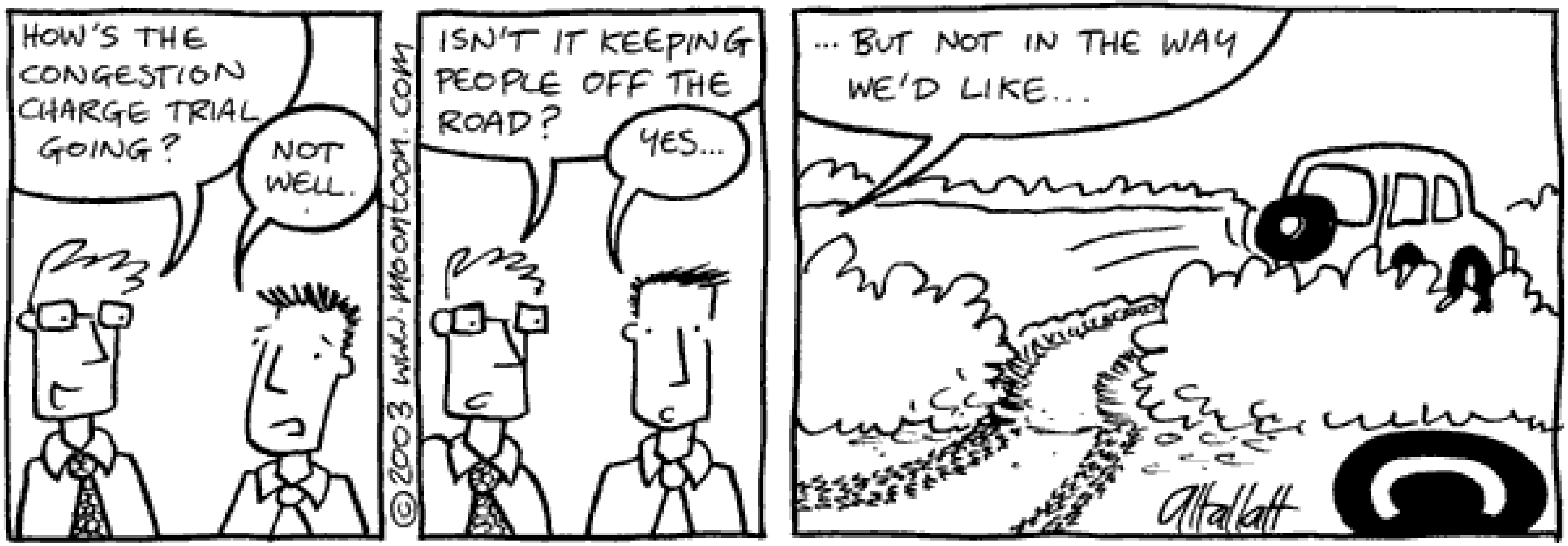
- The subnetwork approach is flexible and the model complexity can be controlled by the analyst
- Paper to appear:

Frejinger, E., and Bierlaire, M. (2007). Capturing correlation with subnetworks in route choice models, **Transportation Research Part B: Methodological** 41(3):363-378.

A latent route choice model

Michel Bierlaire, Emma Frejinger and Jelena Stojanovic

Mobility Pricing



Swiss Mobility Pricing Project

- A part of a major study on various mobility pricing scenarios in Switzerland
- A collaboration with ETH Zurich and USI Lugano
- Revealed Preferences (RP) and Stated Preferences (SP) data has been collected
- RP data concern long distance route choice by car
  - Route descriptions are approximative
  - Route choices are latent

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Objective

- Estimate route choice models based on latent chosen routes
- Literature on latent choice models
  - Ben-Akiva et al. (1984), label path approach
  - Ben-Akiva and Lerman (1985), destination choice
  - Ben-Akiva et al. (2006), lane choice

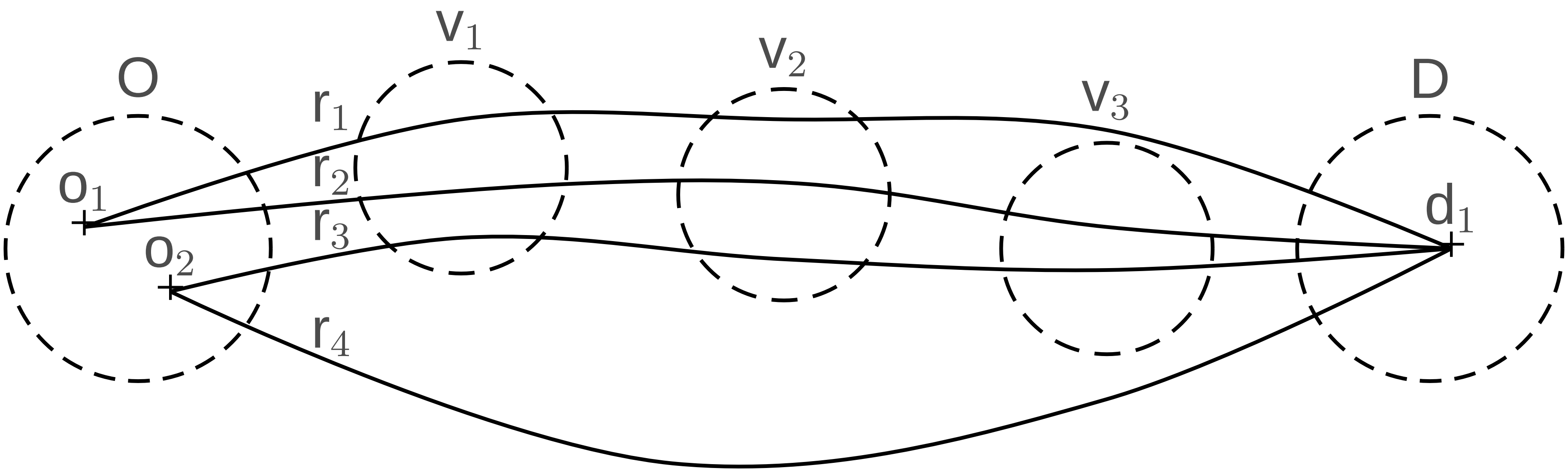
Observations

- Exact descriptions of chosen routes are difficult and expensive to obtain
- The concept of path and network as we need for modeling is abstract for respondents
- Here, a chosen route is described by a sequence of cities and locations
- *Aggregate observations* (several paths in the network can correspond to the same observation)

Observations

- Better quality of the observations
- Travelers do not need to refer to the network used by the analyst
- Exact origin-destination pairs are not necessarily known
- Exact route is not known

Observations - Example



Modeling Approach

- Several possible modeling approaches
  - Construction of paths from the aggregate observations
    - Involves subjective judgments and generate noise
  - Alternatives in the model are aggregates instead of physical paths
    - Estimated model is of little use in practice
- Our approach: compute the likelihood of an aggregate observation for a classical route choice model

Modeling Approach

- Probability of an aggregate observation i:

$$P(i) = \sum_{s \in S_i} P(s|S_i) \sum_{r \in C_s} \delta_{ri} P(r|C_s)$$

  - s: origin-destination pair
  - S_i: set of all origin-destination pairs for observation i
  - r: route
  - C_s: set of all routes for origin-destination pair s
  - $\delta_{ri} = \begin{cases} 1 & \text{if } r \text{ corresponds to } i \\ 0 & \text{otherwise} \end{cases}$

Modeling Approach

- Probability of an aggregate observation i:

$$P(i) = \sum_{s \in S_i} P(s|S_i) \sum_{r \in C_s} \delta_{ri} P(r|C_s)$$

  - $P(s|S_i)$ can be modeled in several ways
  - $P(r|C_s)$: route choice model that is identifiable if
    1. at least one of the routes in C_s crosses the observed zones, and
    2. at least one route in C_s does not cross the observed zones.
- This type of models can be estimated with BIOGEME

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Empirical Results

- Simplified Swiss network (39411 links and 14841 nodes)
- RP data collection through telephone interviews
- Long distance car travel
- The chosen routes are described with the origin and destination cities as well as 1 to 3 cities or locations that the route pass by
- 940 observations available after data cleaning and verification

Empirical Results

Empirical Results

- This application is one of few presented in the literature that are based on RP data
- The network is to our knowledge the largest one used for evaluation of route choice modeling approaches

Empirical Results

- No information available on the exact origin destination pairs

$$P(s|i) = \frac{1}{|S_i|} \; \forall s \in S_i$$

- Two origin-destination pairs are randomly chosen for each observation

Empirical Results

- 46 routes per choice set are generated with a choice set generation algorithm
- After choice set generation 780 observations are available
  - 160 observations were removed because either all or none of the generated routes crossed the observed zones
- We estimate Path Size Logit (Ben-Akiva and Bierlaire, 1999) and Subnetwork (Frejinger and Bierlaire, 2006) models

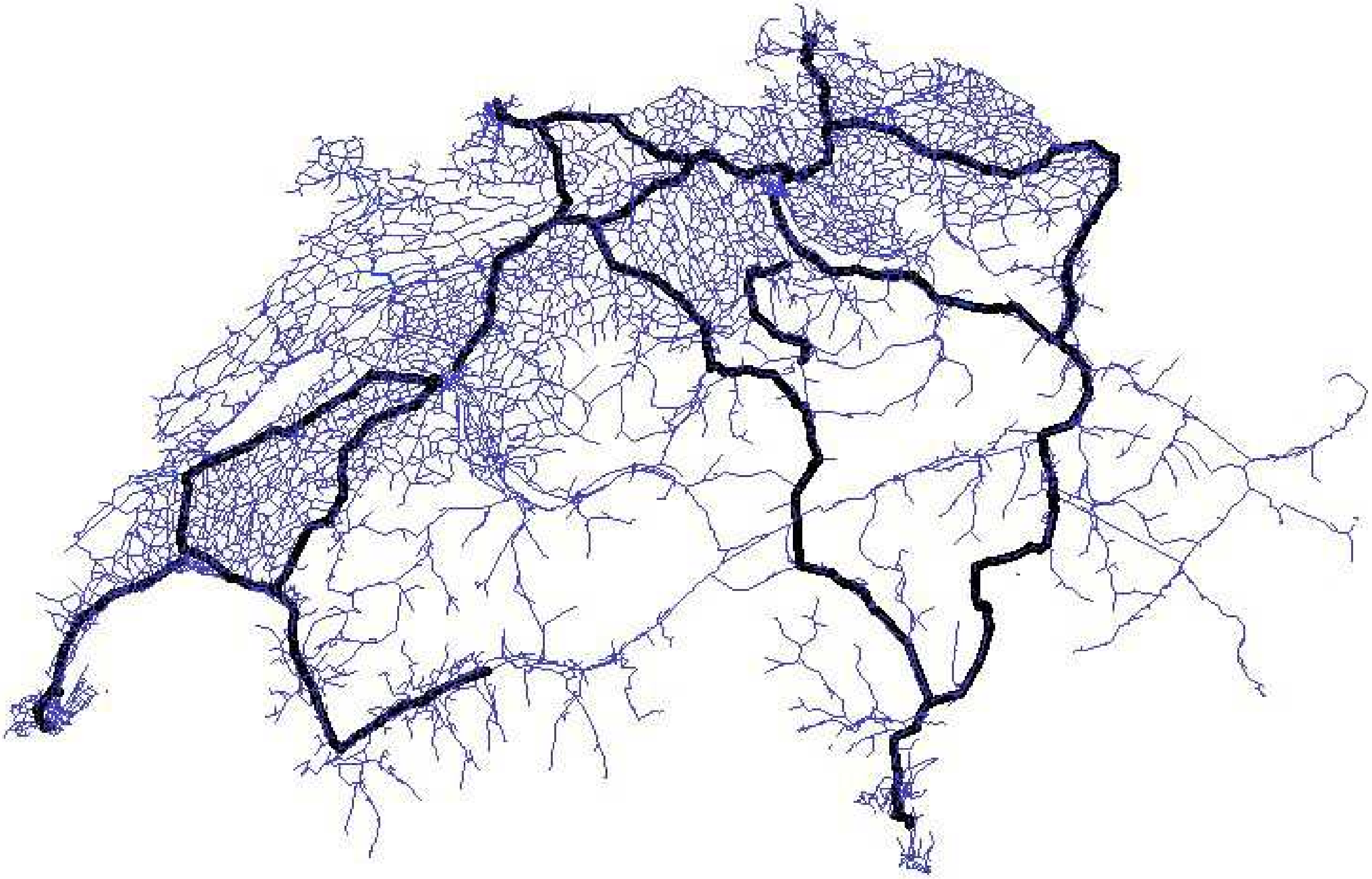
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Empirical Results - Subnetwork

- Subnetwork: main motorways in Switzerland
- Correlation among routes is explicitly modeled on the subnetwork
- Combined with a Path Size attribute
- Linear-in-parameters utility specifications

Empirical Results - Subnetwork



Parameter	PSL		Subnetwork	
ln(path size) based on free-flow time	**1.04**	(0.134) 7.81	**1.10**	(0.141) 7.78
Scaled Estimate	*1.04*		*1.04*	
Freeway free-flow time 0-30 min	**-7.12**	(0.877) -8.12	**-7.45**	(0.984) -7.57
Scaled Estimate	*-7.12*		*-7.04*	
Freeway free-flow time 30min - 1 hour	**-1.69**	(0.875) -1.93	**-2.26**	(1.03) -2.19
Scaled Estimate	*-1.69*		*-2.14*	
Freeway free-flow time 1 hour +	**-4.98**	(0.772) -6.45	**-5.64**	(1.00) -5.61
Scaled Estimate	*-4.98*		*-5.33*	
CN free-flow time 0-30 min	**-6.03**	(0.882) -6.84	**-6.25**	(0.975) -6.41
Scaled Estimate	*-6.03*		*-5.91*	
CN free-flow time 30 min +	**-1.87**	(0.331) -5.64	**-2.16**	(0.384) -5.63
Scaled Estimate	*-1.87*		*-2.04*	
Main free-flow travel time 10 min +	**-2.03**	(0.502) -4.05	**-2.46**	(0.624) -3.95
Scaled Estimate	*-2.03*		*-2.33*	
Small free-flow travel time	**-2.16**	(0.685) -3.16	**-2.75**	(0.804) -3.42
Scaled Estimate	*-2.16*		*-2.60*	
Proportion of time on freeways	**-2.2**	(0.812) -2.71	**-2.31**	(0.865) -2.67
Scaled Estimate	*-2.2*		*-2.18*	
Proportion of time on CN	**0 fixed**		**0 fixed**	
Proportion of time on main	**-4.43**	(0.752) -5.88	**-4.40**	(0.800) -5.51
Scaled Estimate	*-4.43*		*-4.16*	
Proportion of time on small	**-6.23**	(0.992) -6.28	**-6.02**	(1.03) -5.83
Scaled Estimate	*-6.23*		*-5.69*	
Covariance parameter			**0.217**	(0.0543) 4.00
Scaled Estimate			*0.205*	

Empirical Results

	PSL	Subnetwork
Covariance parameter		0.217
(Rob. Std. Error) Rob. T-test		(0.0543) 4.00
Number of simulation draws	-	1000
Number of parameters	11	12
Final log-likelihood	-1164.850	-1161.472
Adjusted rho square	0.145	0.147
Sample size: 780, Null log-likelihood: -1375.851		

Empirical Results

- All parameters have their expected signs and are significantly different from zero
- The values and significance level are stable across the two models
- The subnetwork model is significantly better than the Path Size Logit (PSL) model

Conclusion - Latent route choice

- Aggregate observations are convenient to report paths
- They can be used for estimating route choice models
- Care must be taken about the level of aggregation
- Parameters of the RP model are significant and meaningful
- Available in Biogeme / Bioroute

Future work

- Choice set generation
  - Stochastic path generation algorithm
- Analysis of sensitivity of the modeling results regarding the choice set definition