Initial Comparisons Between MDCEV model and OASIS Framework

Mengyi Wang Janody Pougala Michel Bierlaire

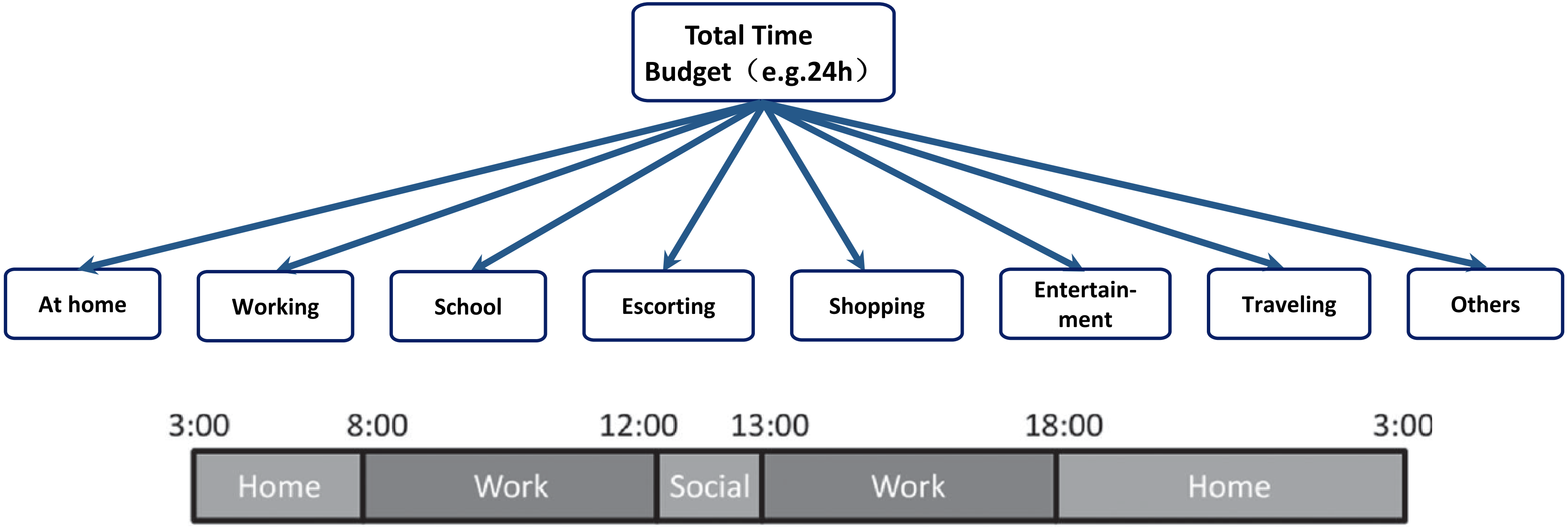
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique F´ed´erale de Lausanne

23rd, STRC, Ascona, 2023

CONTENTS

Activity-based Modeling (ABM)

- **Time** is treated as an important limited resource.
- **Time-use** and **scheduling process** are critical and central components.



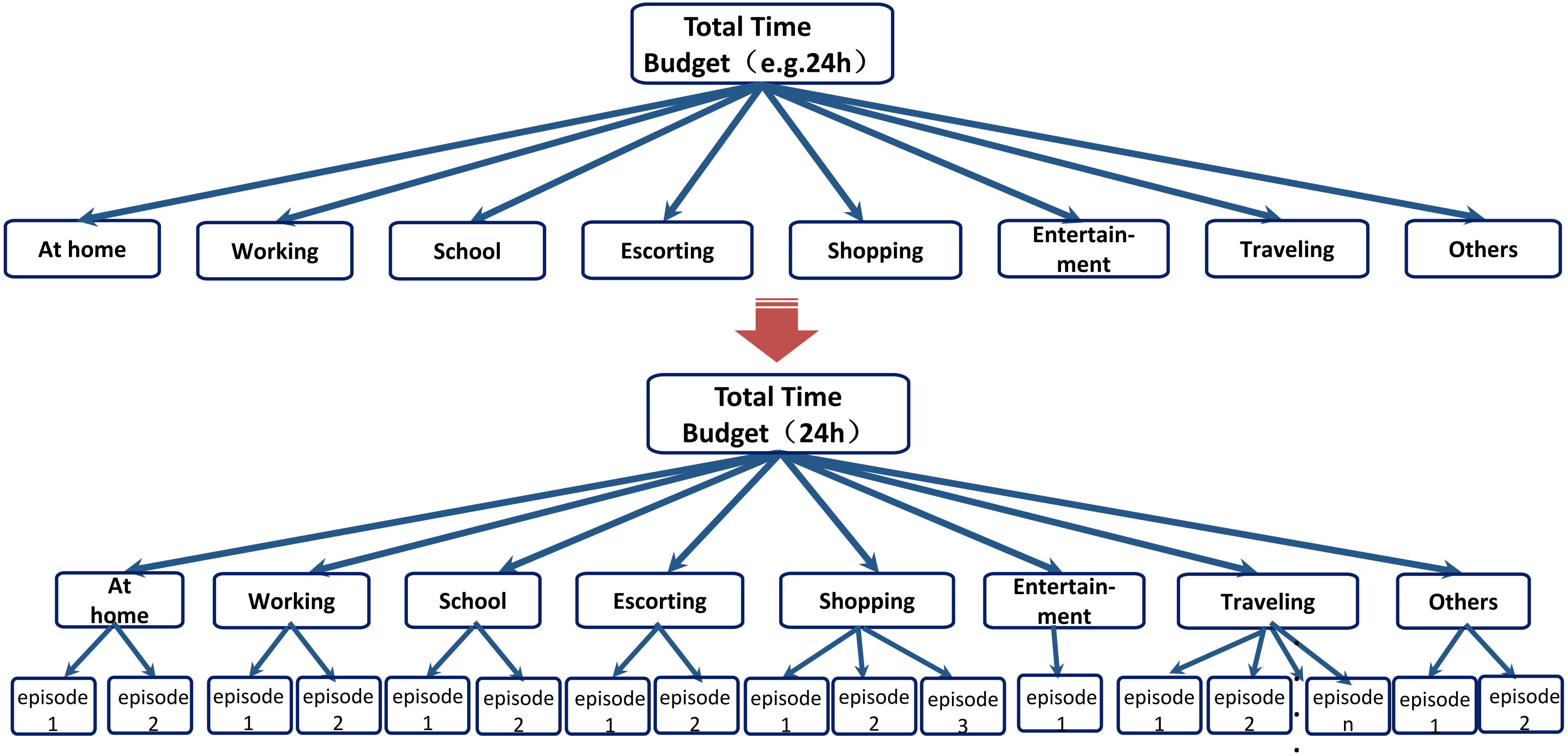
Existing Methods in Time Use

Time Use Models

- Fractional Logit Model
- Hazard-based Duration Model
- Structural Equation Model
- Tobit Model
- Multiple Discrete-Continuous Extreme Value (MDCEV) Model (Bhat, 2005)

- Multiple **activity participation** and **activity duration**
- Random utility theory with diminishing marginal utilities
- Time budget constraint
- Efficient estimation

Limitation of MDCEV model

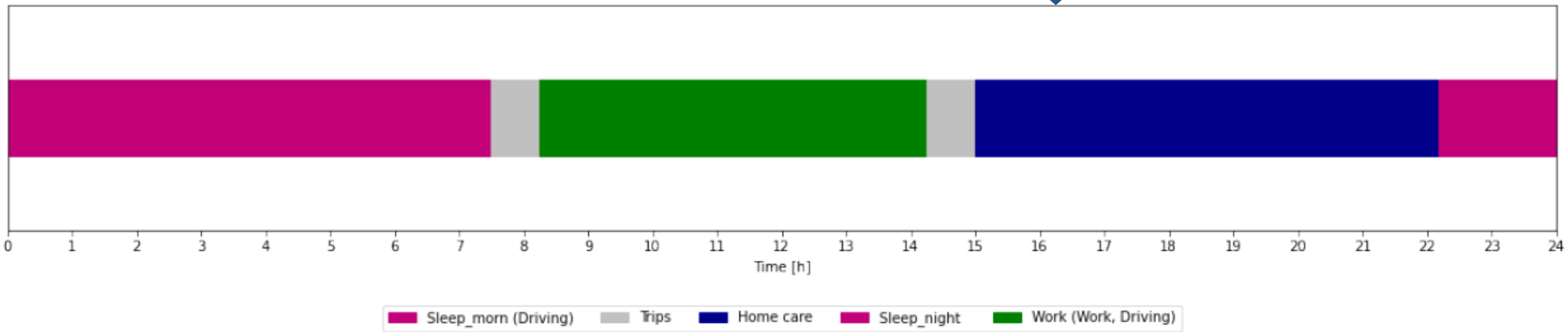


OASIS Framework (Pougala et al., 2022)

Dimensions:

- Activity participation;
- Activity duration;
- Activity location;
- Time of day;
- Travel mode

Person	Schedule	Activity	Start time (hh:mm)	Duration (hh:mm)	Location	Mode	Participation mode
Sara	Isolated indiv./ Fam. of 2/ Fam. of 3	Sleep_morn	00:00	6:30	Home	No trip	Solo
		Sleep_morn	00:00	6:30	Home	PT	Solo
		Sleep_morn	00:00	6:30	Home	Driving	Solo
		Personal care	6:30	2:00	Home	No trip	Solo
		Personal care	6:30	2:00	Home	PT	Solo
		Personal care	6:30	2:00	Home	Driving	Solo
		Work	8:30	6:00	Work	PT	Solo
		Work	8:30	6:00	Work	Driving	Solo
		Work	8:30	6:00	Home	No trip	Solo
		Work	8:30	6:00	Home	PT	Solo
		Work	8:30	6:00	Home	Driving	Solo
		Home care	14:30	7:40	Home	No trip	Solo
		Home care	14:30	7:40	Home	PT	Solo
		Home care	14:30	7:40	Home	Driving	Solo
		Sleep_night	22:10	1:50	Home	No trip	Solo



Motivation

- It is valuable to analyze strengths, weaknesses and application scenarios of **MDCEV** and **OASIS**.
- Currently, there is no **comparative study** of them in the literature.
- An initial attempt to compare the **forecasting performance** in two **common dimensions** of the models: **activity participation** and **activity duration**.

CONTENTS

Comparison List

Features		MDCEV	OASIS
Model Structure	**Principle**	Constrained optimization problem with random utility theory	
	Input	Individual and household attributes; Model parameters; Error terms	A set of considered activities with locations and transport modes; Scheduling preferences; Activity flexibility; Model parameters; Error terms
	Output	Time allocation of different activities	Activity schedule for whole day
	Modeling Dimensions	Activity participation; Activity duration , Activity frequency; Chronological order of episodes from the same activity category	Activity participation; Activity start time; Activity duration; Activity frequency; Activity sequence; Activity location; Travel mode; Travel time
	Utility Function	Non-linear	Linear
	Explanatory Variables	Individual and household attributes	Activity-travel attributes
	Choice Set	All activity categories	All valid schedules
Model Estimation		Maximum likelihood estimation	Maximum likelihood estimation relied on sampled choice set
Model Simulation		Specific algorithm by certain rules	Standard mathematical programming algorithm
Model Extension		In-home activities; Household level; Different distributions of error terms; Random coefficients	
Research Focus		Exploring time allocation mechanism by sophisticated utility functions; Describing influence factors by estimation results	Exploring daily scheduling mechanism by integrating all dimensions; Capturing trade-offs between different choices by model simulation

Model Structure—Principle

MDCEV

$$\max\ U(\boldsymbol{t}) = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left(\frac{t_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\}$$

$$\text{s.t.} \sum_{k=1}^{K} t_k = T$$

Where： ψ_k— baseline marginal utility；

（$\psi_k > 0$, $\psi(\boldsymbol{z}_k, \varepsilon_k) = \exp(\boldsymbol{\beta}' \boldsymbol{z}_k + \varepsilon_k)$）

γ_k and α_k— satiation parameters.

（$\gamma_k > 0$，$0 < \alpha_k < 1$）

OASIS

$$\max_{\omega,z,x,\tau} U(\omega, z, x, \tau, \epsilon) + \sum\nolimits_{a=0}^{A} \omega_a (V_a^1 + V_a^2 + V_a^3) + \sum\nolimits_{a=0}^{A} \sum\nolimits_{b=0}^{A} z_{ab} (V_{ab}^4 + V_{ab}^5).$$

s.t.

$$\begin{cases}
\sum_a \sum_b (\omega_a \tau_a + z_{ab} \rho_{ab}) = T, \\
\sum_a \sum_b (\omega_a c_a + z_{ab} \kappa_{ab}) \le B, \\
\omega_{\text{dawn}} = \omega_{\text{dusk}} = 1, \\
\tau_a \ge \omega_a \tau_a^{\min}, \quad \forall a \in A, \\
\tau_a \le \omega_a T, \quad \forall a \in A, \\
z_{ab} + z_{ba} \le 1, \quad \forall a, b \in A, a \ne b, \\
z_{a,\text{dawn}} = z_{\text{dusk},a} = 0, \quad \forall a \in A, \\
\sum_a z_{ab} = \omega_b, \quad \forall b \in A, b \ne \text{dawn}, \\
\sum_b z_{ab} = \omega_a, \quad \forall a \in A, a \ne \text{dusk}, \\
(z_{ab} - 1)T \le x_a + \tau_a + z_{ab}\rho_{ab} - x_b, \quad \forall a, b \in A, a \ne b, \\
(1 - z_{ab})T \ge x_a + \tau_a + z_{ab}\rho_{ab} - x_b, \quad \forall a, b \in A, a \ne b, \\
\sum_{a \in G_k} \omega_a, \le 1 \quad k = 1, \dots, K, \\
\alpha_a^m = 1 \quad \forall a \in G_{\text{home}} \\
\omega_a \le \alpha_a^m \quad \forall a \in A^m \\
\omega_a \ge \omega_b + z_{ab} - 1 \quad \forall a \in A, b \in A \setminus G_{\text{home}}
\end{cases}$$

Where, the decision variables are：

ω_a—Whether activity a is selected;

z_{ab}—Whether activity b is after a;

x_a— Start time of activity a；

τ_a—Duration of activity a；

α_a^{m}—Whether private mode m is available for activity a.

Model Estimation

MDCEV

$$\max\ U(\boldsymbol{t}) = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k}\psi_k \left\{\left(\frac{t_k}{\gamma_k}+1\right)^{\alpha_k}-1\right\}$$

$$\text{s.t.} \sum_{k=1}^{K} t_k = T$$

$$\mathcal{L} = \sum_{k} \frac{\gamma_k}{\alpha_k}[\exp(\beta' z_k + \varepsilon_k)]\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k}-1\right\} - \lambda\left[\sum_{k=1}^{K} x_k - B\right]$$

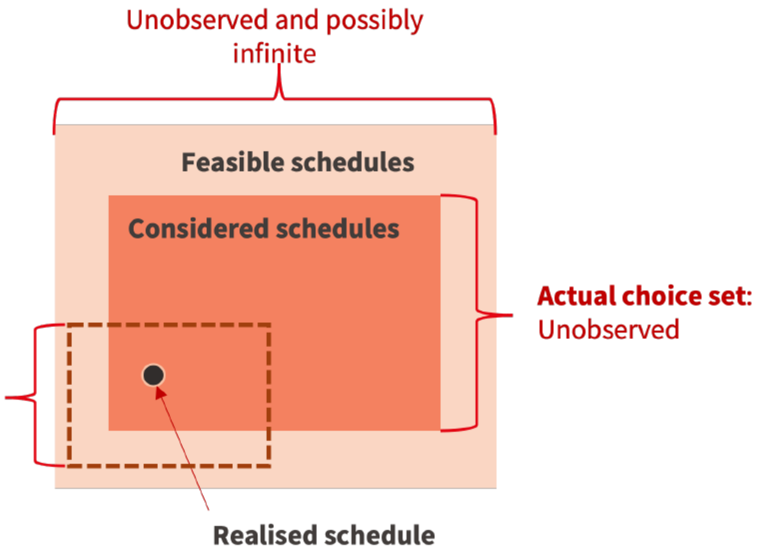
$$P(x_1^*, x_2^*, x_3^*, \ldots, x_M^*, 0,0,\ldots,0) = \left[\prod_{i=1}^{M} c_i\right]\left[\sum_{i=1}^{M}\frac{1}{c_i}\right]\left[\frac{\prod_{i=1}^{M} e^{V_i}}{\left(\sum_{j=1}^{K} e^{V_j}\right)^{M}}\right](M-1)!$$

Maximum likelihood estimation

OASIS

M-H algorithm

Estimation choice set: sample of feasible schedules generated for estimation purposes



$$P\left(i_n|\mathcal{C}_n\right) = \frac{\exp\left[V_{in} + \ln q(\mathcal{C}_n|i_n)\right]}{\sum_{j\in\mathcal{C}_n}\exp\left[V_{jn} + \ln q(\mathcal{C}_n|j)\right]}$$

Maximum likelihood estimation based on sampled choice set

Model Simulation

MDCEV

If 'i' is a chosen alternative and 'j' is not a chosen alternative.

KT conditions

$$\begin{cases} \psi_i \left(\dfrac{t_i^*}{\gamma_i} + 1 \right)^{\alpha_i - 1} = \lambda, \; t_i^* > 0 \\ \psi_j \left(\dfrac{t_j^*}{\gamma_j} + 1 \right)^{\alpha_j - 1} < \lambda, \; t_j^* = 0 \end{cases}$$

$$\psi_j < \lambda < \psi_i$$

Specific algorithm by certain rules

OASIS

$$\max_{\omega,z,x,\tau} U(\omega, z, x, \tau, \epsilon) + \sum_{a=0}^{A} \omega_a (V_a^1 + V_a^2 + V_a^3) + \sum_{a=0}^{A} \sum_{b=0}^{A} z_{ab} (V_{ab}^4 + V_{ab}^5).$$

$$\sum_a \sum_b (\omega_a \tau_a + z_{ab} \rho_{ab}) = T,$$

$$\sum_a \sum_b (\omega_a c_a + z_{ab} \kappa_{ab}) \leq B,$$

$$\omega_{\text{dawn}} = \omega_{\text{dusk}} = 1,$$

$$\tau_a \geq \omega_a \tau_a^{\min}, \quad \forall a \in A,$$

$$\tau_a \leq \omega_a T, \quad \forall a \in A,$$

$$z_{ab} + z_{ba} \leq 1, \quad \forall a, b \in A, a \neq b,$$

$$z_{a,\text{dawn}} = z_{\text{dusk},a} = 0, \quad \forall a \in A,$$

$$\sum_a z_{ab} = \omega_b, \quad \forall b \in A, b \neq \text{dawn},$$

$$\sum_b z_{ab} = \omega_a, \quad \forall a \in A, a \neq \text{dusk},$$

$$(z_{ab} - 1)T \leq x_a + \tau_a + z_{ab}\rho_{ab} - x_b, \quad \forall a, b \in A, a \neq b,$$

$$(1 - z_{ab})T \geq x_a + \tau_a + z_{ab}\rho_{ab} - x_b, \quad \forall a, b \in A, a \neq b,$$

$$\sum_{a \in G_k} \omega_a, \leq 1 \quad k = 1, \ldots, K,$$

$$\alpha_a^m = 1 \quad \forall a \in G_{\text{home}}$$

$$\omega_a \leq \alpha_a^m \quad \forall a \in A^m$$

$$\omega_a \geq \omega_b + z_{ab} - 1 \quad \forall a \in A, b \in A \setminus G_{\text{home}}$$

Where: ω_a, z_{ab}, x_a, τ_a, α_a^m are decision variables.

Standard mathematical programming algorithm

CONTENTS

Date Source and Sample Formation

- **Data Source**: 2015 Mobility and Transport Microcensus (MTMC)
- **Sample Formation** :

- Select full samples of Lausanne;
- Remove those who didn’t finish their daily schedule at home;
- Delete samples who have negative activity durations or total duration is more than 24 hours;
- Remove those who are at home all the day;
- Final sample size: 1016 (700 for estimation; 316 for simulation)

- **Activity Type**:

home, work, education, shopping, leisure, escort, errands and services, business trip

CONTENTS

MDCEV Model

- Different profiles
  - γ-profile: $max\ \ U(\mathbf{t}) = \sum_{k=1}^{K} \gamma_k \psi_k \ln\left(\frac{t_k}{\gamma_k} + 1\right)$
  - α-profile: $max\ \ U(\mathbf{t}) = \sum_{k=1}^{K} \frac{\psi_k}{\alpha_k}\{(t_k + 1)^{\alpha_k} - 1\}$

  → **Traditional MDCEV (Bhat, 2008)**
  - Parabola: $max\ \ U(\mathbf{t}) = \sum_{k=1}^{K} -0.5\psi_k(t_k - m_k)^2$

  → **Parabolic MDCEV (Wang, M. and X. Ye., 2023)**
- Parameters to be estimated (activity-specific)

	γ-profile	α-profile	Parabola
parameters	Baseline marginal utility: ψ_k Satiation parameter: γ_k	Baseline marginal utility: ψ_k Satiation parameter:α_k	Opening size : ψ_k Extreme point: m_k

MDCEV Model

Traditional MDCEV models

Parameters	γ-Profile		α-Profile	
	Value	t-Statistic	Value	t-Statistic
ψ_k component				
Constants				
Working	0.6141	6.043	0.5892	5.903
Education	- 0.7964	- 6.087	- 0.7504	- 5.770
Shopping	0.5426	5.500	0.4390	4.601
Leisure	0.9144	9.558	1.0627	11.186
Others	—	—	—	—
Satiation parameters				
Working	1.0000	(fixed)	1.0000	(fixed)
Education	1.0000	(fixed)	1.0000	(fixed)
Shopping	0.6557	17.771	1.0000	(fixed)
Leisure	1.8053	14.018	0.4257	1.971
Others	0.8302	12.758	1.0000	(fixed)
Summary statistics				
Number of cases	700		700	
Final Log-likelihood	- 2775.5		- 2703.6	

Parabolic MDCEV model

Parameters	Value	t-Statistic
m_k component		
Constants		
Working	- 1.9531	- 14.027
Education	- 1.1320	- 21.744
Shopping	- 0.7094	- 33.436
Leisure	0.0000	(fixed)
Others	- 1.1403	- 15.374
ψ_k component		
Constants		
Working	- 1.0348	- 48.763
Education	1.6480	8.197
Shopping	2.2063	12.400
Leisure	—	—
Others	2.2156	5.337
Summary statistics		
Number of cases	700	
Final Log-likelihood	- 4571.119	

OASIS Model

- Parameters to be estimated (activity-specific)

Parameter	Notation	Activities
Activity-specific constant	$ASC_{a,n}$	
Penalty start time (early)	θ^{e}_{F}	
Penalty start time (late)	$\theta^{\ell}_{\mathrm{F}}$	All (*reference: home*)
Penalty duration (short)	β^{s}_{F}	
Penalty duration (long)	$\beta^{\ell}_{\mathrm{F}}$	

- Choice set size: 25 per observation
- Travel parameters are not included
- Desired start times and durations:
  drawn from pre-determined distributions

$$U_s = U + \sum_{a=0}^{A-1}\left(U_a^1 + U_a^2 + U_a^3 + \sum_{b=0}^{A-1}\left(U_{a,b}^4 + U_{a,b}^5\right)\right)$$

$$U_a^1 = \beta_{\text{cost}} * c_a + \varepsilon_1$$

$$U_a^2 = \theta_a^e max(0, x_a^- - x_a) + \theta_a^\ell max(0, x_a - x_a^+)$$

$$U_a^3 = \beta_{a_a}^e max(0, \tau_a^- - \tau_a) + \beta_a^\ell max(0, \tau_a - \tau_a^+)$$

$$U_{a,b}^4 = \beta_{t,\text{ cost}} * c_t + \varepsilon_4$$

$$U_{a,b}^5 = \theta_t \rho_{ab}$$

OASIS Model

Parameters	Value	t-Statistic
Constants		
Working	12.800	8.54
Education	9.160	8.69
Shopping	10.100	13.20
Leisure	10.400	11.00
Early		
Working	- 1.380	- 4.95
Education	- 1.490	- 7.85
Shopping	- 1.060	- 12.00
Leisure	- 0.254	- 1.46
Late		
Working	- 0.532	- 3.01
Education	- 0.575	- 3.42
Shopping	- 0.685	- 4.46
Leisure	- 1.030	- 8.75
Long		
Working	- 0.746	- 3.95
Education	- 5.290	- 4.37
Shopping	- 1.150	- 4.56
Leisure	- 0.434	- 5.13
Short		
Working	- 1.240	- 4.85
Education	- 0.336	- 2.10
Shopping	- 7.910	- 2.67
Leisure	- 0.566	- 2.32

Summary statistics
Number of observations = 700
$L(0) = -1875.052$
$L(\beta) = -218.583$
$\bar{\rho}^2 = 0.873$

CONTENTS

Statistical Analysis

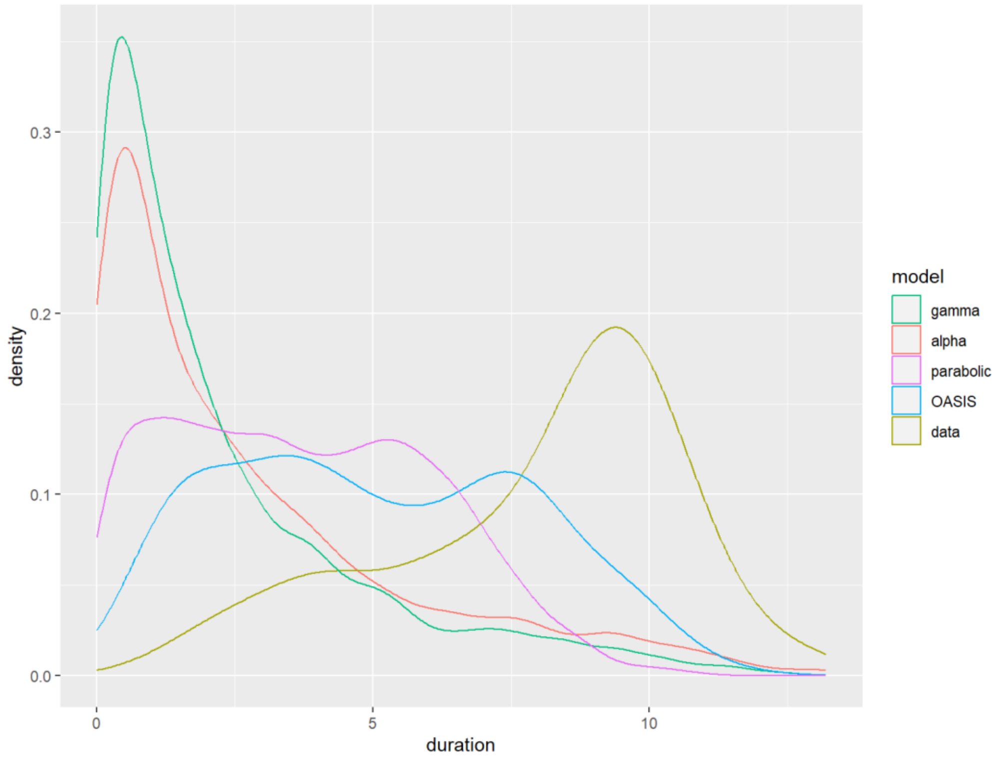
- Average time spent out-of-home, in hh:min (20 schedules for each sample)

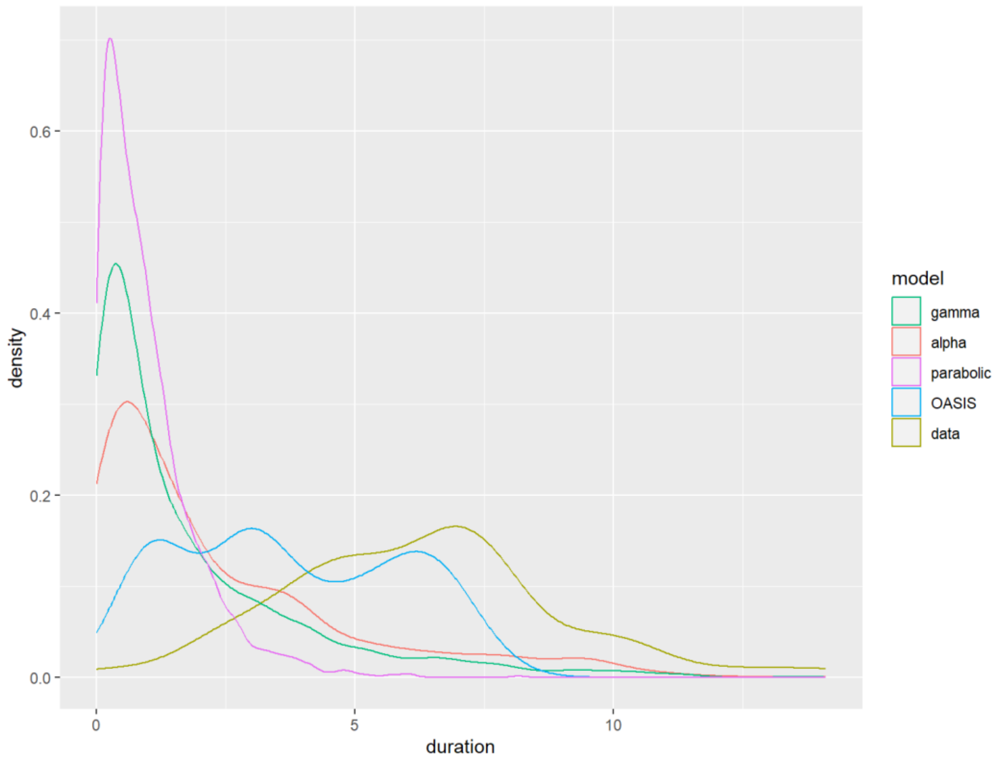
Activity	Data	MDCEV			OASIS
		γ-Profile	α-Profile	Parabolic	
Working	02:36	01:15	01:15	02:53	01:03
Education	01:07	00:21	00:22	00:20	00:31
Shopping	00:13	00:42	01:07	00:27	00:14
Leisure	01:25	02:55	02:19	01:50	01:30

- Proportion of schedules containing each activity (20 schedules for each sample)

Activity	Data	MDCEV			OASIS
		γ-Profile	α-Profile	Parabolic	
Working	0.33	0.54	0.45	0.77	0.21
Education	0.18	0.19	0.16	0.33	0.14
Shopping	0.33	0.49	0.42	0.43	0.17
Leisure	0.58	0.73	0.64	0.86	0.75

Duration Distributions

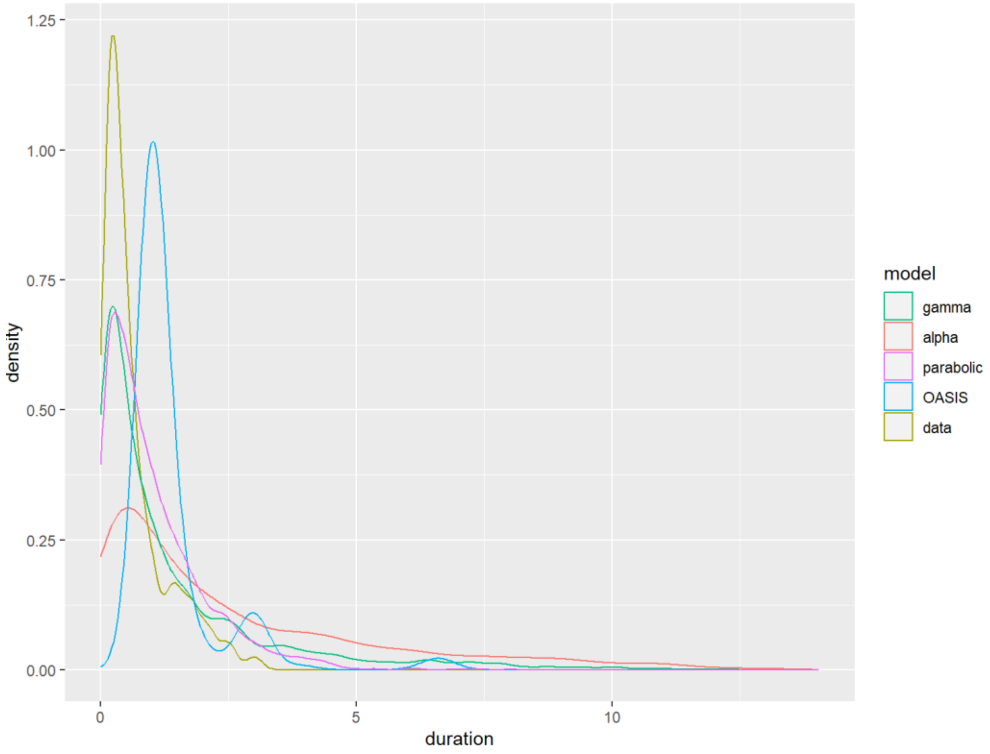




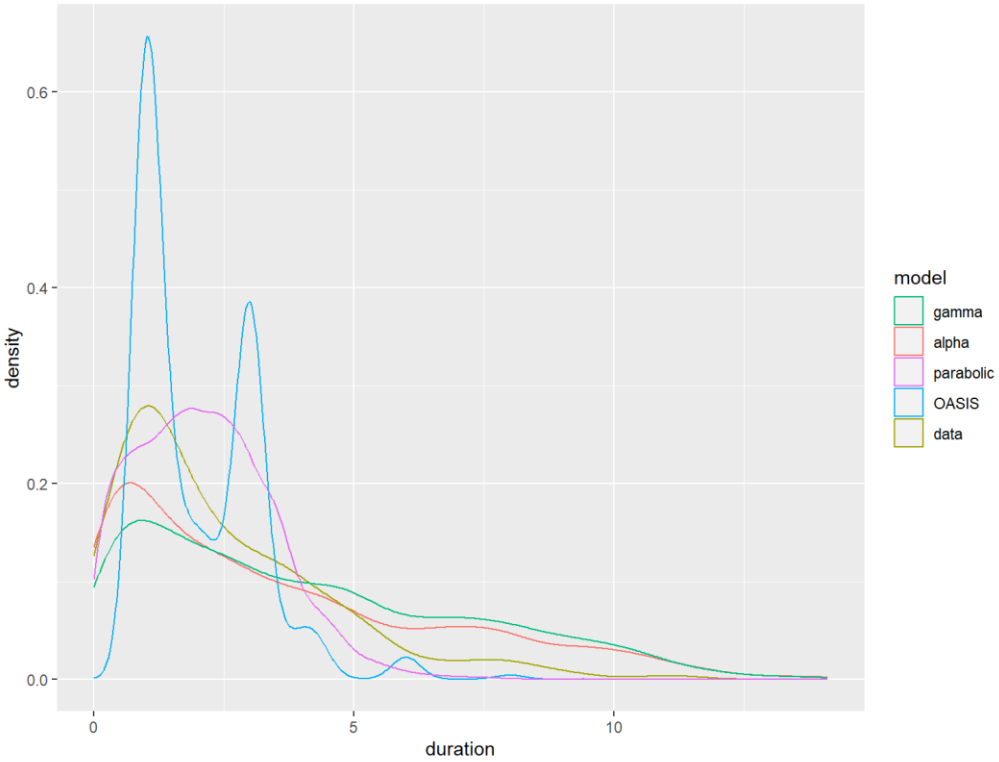
Working

Education

Duration Distributions



Shopping



Leisure

CONTENTS

Limitations and Future Directions

- Socio-demographics and more sophisticated utility functions are to be considered;
- Further comparative study is to be carried out when the MDCEV model is combined with extra scheduling algorithm;
- More other datasets are to be tested.

Thanks for listening!

Mengyi Wang Janody Pougala Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering Ecole Polytechnique F´ed´erale de Lausanne

23rd, STRC, Ascona, 2023