Activity-based models: an optimization perspective

Janody Pougala  Tim Hillel  Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

November 23, 2020
Outline

1 Introduction

2 Model

3 Mixed integer optimization problem

4 Example

5 Parameter estimation

Pougala, Hillel, Bierla (EPFL)
Travel demand is derived from activity demand.

Activity demand is influenced by socio-economic characteristics, social interactions, cultural norms, basic needs, etc. [Chapin, 1974]

Activity demand is constrained in space and time [Hägerstraand, 1970].
Econometric models

Rule-based models
State of the art: econometric approach

[Pinjari et al., 2011]

- ... individuals make their activity-travel decisions to maximize the utility derived from the choices they make.
- These model systems usually consist of a series of ... discrete choice models ... that are used to predict ... individuals’ activity-travel decisions.
- these model systems employ econometric systems of equations ... to capture relationships between ... socio-demographics and ... attributes on the one hand and the observed activity-travel decision outcomes on the other.
State of the art: econometric approach

[Pinjari et al., 2011]: main criticisms

- individuals are not necessarily fully rational utility maximizers
- the approach does not explicitly model the underlying decision processes and behavioral mechanisms that lead to observed activity-travel decisions.
State of the art: econometric approach

[Bhat, 2005]

- Multiple Discrete Continuous Extreme Value
- Based on first principles.
- Decision-maker solves an optimization problem, with a time budget.
- Several alternatives may be chosen.
- Model derived from KKT conditions.
Introduction

Research question

Relax the *series of discrete choice models* approach

- The interactions of all decisions is complex.
- Sequence of models is most of the time arbitrary.

Integrated approach

Develop a model involving many decisions:

- activity participation,
- activity location,
- activity duration,
- activity scheduling,
- travel mode,
- travel route.

Pougala, Hillel, Bierlaire (EPFL)

November 23, 2020 8 / 45
Research objectives

- Integrated approach based on first principles.
- Theoretical framework: utility maximization.
- Individuals solve a scheduling problem.
- Important aspects: trade-offs on activity duration.
Outline

1 Introduction
2 Model
3 Mixed integer optimization problem
4 Example
5 Parameter estimation
First principles

- Each individual $n$ has a time-budget (a day).
- Each activity $a$ considered by $n$ is associated with a utility $U_{an}$.
- Individuals schedule their activities as to maximize the total utility, subject to their time-budget constraint.
Further assumptions

Individuals are **time sensitive**

- Have a desired *start time, duration* and/or *end time* for each activity
- Deviations from their desired times in the scheduling process decrease the utility function
Time horizon: 24 hours.
Discretization: $T$ time intervals.
Trade-off between model accuracy and computational time.
Space

- Discrete and finite set $S$ of locations, indexed by $s$.
- For each individual, each activity is associated with a list of potential locations.
Travel

- For each pair OD, list of possible modes.
- For each mode, list of possible routes.
- For each \((O, D, m, r)\), \(\rho(O, D, m, r)\) is the travel time.
- Exogenously given.
Activities

Definition: Activity
An activity is associated with a location and a trip.
Activities

Location, mode and route choices

- Lunch at location A, followed by trip by bus on route A.
- Lunch at location A, followed by trip by bus on route B.
- Lunch at location A, followed by trip by car on route A.
- Lunch at location B, followed by trip by car on route B.

Constraint

Only one of the “duplicates” can be chosen.
Activities

**Given**
- Set $A$ of activities.
- Location $s_a$.
- Feasible time interval: $[\gamma_a^-, \gamma_a^+]$ (e.g. opening hours).

**Decisions**
- Participation: $w_a \in \{0, 1\}$.
- Starting time $x_a$, $0 \leq x_a \leq T$.
- Schedule: $z_{ab} \in \{0, 1\}$.
- Duration: $\tau_a \geq 0$. 


---

Pougala, Hillel, Bierlaire (EPFL)
Scheduling

![Diagram of daily activities schedule]

- Home
- Work
- Lunch
- Work
- Leisure
- Home

Pougala, Hillel, Bierlaire (EPFL)
Categories

- [Castiglione et al., 2014]: mandatory, maintenance, discretionary.
- Flexible, somewhat flexible, not flexible.

Category

Activities that share the same preference profile.
Preferences

- desired starting time $x_a^*$,
- desired duration $\tau_a^*$.

Penalties

- Starting early [Small, 1982]: $\theta_e \max(x_a^* - x_a, 0)$.
- Starting late [Small, 1982]: $\theta_\ell \max(x_a - x_a^*, 0)$.
- Shorter activity: $\theta_{ds} \max(\tau_a^* - \tau_a, 0)$.
- Longer activity: $\theta_{d\ell} \max(\tau_a - \tau_a^*, 0)$. 
Preferences

Parameters depend on the category type

Utility

Time

Early

Late

Flexible

Somewhat flexible

Not flexible

x^*

Pougala, Hillel, Bierlaire (EPFL)
Disutility of travel

Each activity is followed by a trip

- Travel time from $a$ to $a^+$: $t_a$.
- Depends on the next activity.

$$t_a = \sum_{b} z_{ab} \rho(s_a, s_b, m_a, r_a).$$

- Over variables can be included (cost, etc.)
- Note: If $s_a = s_b$, $\rho(s_a, s_a, m_a, r_a) = 0$
- Exception: last activity of the day (home).
Utility function

An individual $n$ derives the following utility from performing activity $a$, with a schedule flexibility $k$:

$$U_{an} = \theta_e \max(x_a^* - x_a, 0) + \theta_\ell \max(x_a - x_a^*, 0) + \theta_{ds} \max(\tau_a^* - \tau_a, 0) + \theta_{d\ell} \max(\tau_a - \tau_a^*, 0) + c_{an} + \varepsilon_{an},$$

where $\varepsilon_{an}$ are error components.
Utility function

Utility of a schedule

\[ U_{sn} = \sum_a w_a U_{an} + \theta_t \sum_a \sum_b z_{ab} \rho(a, b, m_a, r_a) \]

Error components

\[ \sum_a w_a \varepsilon_{an} \]

where \( \varepsilon_{an} \) normally distributed.
Utility function

- Rely on simulation.
- Draw $\varepsilon_{anr}, r = 1, \ldots, R$.
- Optimization problem for each $r$.
- Utility: $U_{anr}$. 
Outline

1. Introduction
2. Model
3. Mixed integer optimization problem
4. Example
5. Parameter estimation
Decision variables for individual $n$ and draw $r$

For each (potential) activity $a$:

- Activity participation: $w_{anr} \in \{0, 1\}$.
- Starting time: $x_{anr} \in \{0, \ldots, T\}$.
- Duration: $\tau_{anr} \in \{0, \ldots, T\}$.
- Scheduling: $z_{abnr} \in \{0, 1\}$: 1 if activity $b$ immediately follows $a$. 
Objective function

Additive utility

\[
\max \sum_{a \in A} w_{anr} U_{anr} + \theta_t \sum_{a \in A} \sum_{b \in A} z_{abnr} \rho(a, b, m_a, r_a).
\]
Constraints

Time budget

$$\sum_{a \in A} w_{anr} \tau_{anr} + \sum_{a \in A} \sum_{b \in A} z_{abnr} \rho(a, b, m_a, r_a) = T, \forall n, r.$$ 

Time windows

$$0 \leq \gamma_a^- \leq x_{anr} \leq x_{anr} + \tau_{anr} \leq \gamma_a^+ \leq T, \forall a, n, r.$$
Mixed integer optimization problem

Constraints

Precedence constraints

\[ z_{abnr} + z_{banr} \leq 1, \ \forall a, b, n, r. \]

Single successor/predecessor

\[
\sum_{b \in A \setminus \{a\}} z_{abnr} = w_{anr}, \ \forall a, n, r, \\
\sum_{b \in A \setminus \{a\}} z_{banr} = w_{anr}, \ \forall a, n, r.
\]
Mixed integer optimization problem

Constraints

Consistent timing

\[(z_{abnr} - 1)T \leq x_{anr} + \tau_{anr} + t_{anr} - x_b \leq (1 - z_{abnr})T, \ \forall a, b, n, r.\]

where

\[t_{anr} = \sum_{b \in A} z_{abnr} \rho(s_a, s_b).\]

Mutually exclusive duplicates

\[\sum_{a \in B_k} w_{anr} = 1, \ \forall k, n, r.\]
Simulation-based optimization

- For each realization of the error terms, we have an optimal schedule.
- It includes all the choice dimensions (activity participation, location, duration, scheduling, and mode and route).
- We can generate an empirical distribution of chosen schedules.
Outline

1. Introduction
2. Model
3. Mixed integer optimization problem
4. Example
5. Parameter estimation
Real data

Dataset

- 2015 Swiss Mobility and Transport Microcensus.
- Daily trip diaries for 57’000 individuals.
- Records of activities and visited location.

Challenges: classical RP issues

- No information about unchosen alternatives.
- Latent preferences.
Real data

Assumptions

- Desired start times and durations are the recorded ones.

- Feasible time windows: average start and end times from out of sample distribution.

- Only the recorded locations are considered.

- Uniform flexibility profile across population.
Individual 1 (weekday)

Optimal schedules generated for random draws of $\varepsilon_{an}$
Individual 2 (weekday)

Optimal schedules generated for random draws of $\varepsilon_{an}$
Individual 3 (weekday)

Optimal schedules generated for random draws of $\varepsilon_{an}$
Outline

1. Introduction
2. Model
3. Mixed integer optimization problem
4. Example
5. Parameter estimation
Parameter estimation

Choice set generation

- Full set of schedules $C_n$ is combinatorial, approximated with a sample of alternatives $\tilde{C}_n$
- Sampling protocol using Metropolis-Hastings algorithm [Flötteröd and Bierlaire, 2013]

Choice model estimation

- Include an EV error term to obtain a mixture of logit.
- Probability of choosing a schedule $y$ for individual $n$ is conditional on the parameters $\beta_n$, the variables $x_n$ and the sampled choice set $\tilde{C}_n$ [Guevara and Ben-Akiva, 2013]
- Maximum likelihood estimators of the parameters:

$$\max_{\hat{\beta}} L(y|\hat{\beta}, X) = \prod_n P(y|x_n, \hat{\beta}_n, \tilde{C}_n)$$
Conclusions

Achievements so far

- Formulation of the model.
- Applied on real data.
- The results make sense.
- We are able to draw from a distribution of activity schedules.

Ongoing work

- Parameter estimation.


Bibliography II

Metropolis–hastings sampling of paths.

Sampling of alternatives in logit mixture models.

What about people in regional science?
*Papers in Regional Science*.

Activity-based travel demand analysis.
The scheduling of consumer activities: work trips.
American Economic Review.