Formulating discrete choice models as Mixed Integer Linear Problems (MILP)

Michel Bierlaire

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

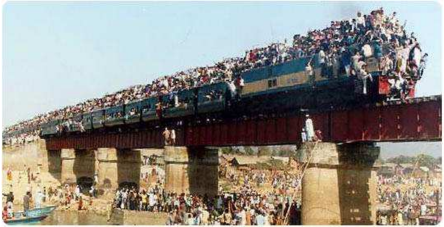
October 10, 2017

Outline

Demand models

- Supply = infrastructure
- Demand = behavior, choices
- Congestion = mismatch

Demand models



- Usually in OR:
- optimization of the supply
- for a given (fixed) demand

Aggregate demand

- Homogeneous population
- Identical behavior
- Price (P) and quantity (Q)
- Demand functions: $P = f(Q)$
- Inverse demand: $Q = f^{-1}(P)$

Disaggregate demand

- Heterogeneous population
- Different behaviors
- Many variables:
  - Attributes: price, travel time, reliability, frequency, etc.
  - Characteristics: age, income, education, etc.
- Complex demand/inverse demand functions.

Demand-supply interactions

Operations Research

- Given the demand...
- configure the system

Behavioral models

- Given the configuration of the system...
- predict the demand

Johnson City Enterprise.

Published Every Saturday,

$1. per year—Advance Payment.

SATURDAY, APRIL 7, 1883.

TIME TABLE

E. T., V. & G. R. R.

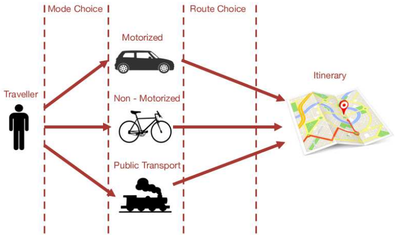
PASSENGER,	ARRIVES,
No. 1, West,	6:37, a. m.
No. 2, East,	9:45, p. m.
No. 3, West,	11:51, p.m.
No. 4, East,	3:56, a. m.
LOCAL FREIGHT,	ARRIVES,
No. 5,	7:20, a. m.
No. 8,	6:20, p. m.

JNO. W. EAKIN, Agent.

E. T. & W. N. C. R. R.

Passenger, leaves,	7, a. m.
" arrives,	6, p. m.

J. C. HARDIN, Agent.



Demand-supply interactions

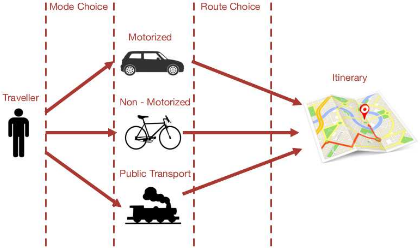
Multi-objective optimization

Minimize costs

Maximize satisfaction

Outline

Choice models



Behavioral models

- Demand = sequence of choices
- Choosing means trade-offs
- In practice: derive trade-offs from choice models

Choice models

Theoretical foundations

- Random utility theory
- Choice set: $\mathcal{C}_n$
- $y_{in} = 1$ if $i \in \mathcal{C}_n$, 0 if not
- Logit model:

$$P(i|\mathcal{C}_n) = \frac{y_{in} e^{V_{in}}}{\sum_{j \in \mathcal{C}} y_{jn} e^{V_{jn}}}$$

2000

Logit model

Utility

$$U_{in} = V_{in} + \varepsilon_{in}$$

Choice probability

$$P_n(i|\mathcal{C}_n) = \frac{y_{in} e^{V_{in}}}{\sum_{j \in \mathcal{C}} y_{jn} e^{V_{jn}}}.$$

- Decision-maker n
- Alternative $i \in \mathcal{C}_n$

Variables: $x_{in} = (p_{in}, z_{in}, s_n)$

Attributes of alternative i: z_{in}

- Cost / price (p_{in})
- Travel time
- Waiting time
- Level of comfort
- Number of transfers
- Late/early arrival
- etc.

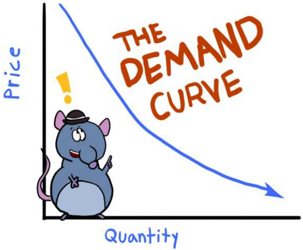
Characteristics of decision-maker n: s_n

- Income
- Age
- Sex
- Trip purpose
- Car ownership
- Education
- Profession
- etc.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Demand curve



Disaggregate model

$$P_n(i|p_{in}, z_{in}, s_n)$$

Total demand

$$D(i) = \sum_n P_n(i|p_{in}, z_{in}, s_n)$$

Difficulty

Non linear and non convex in p_{in} and z_{in}

Outline

Choice-Based Optimization Models

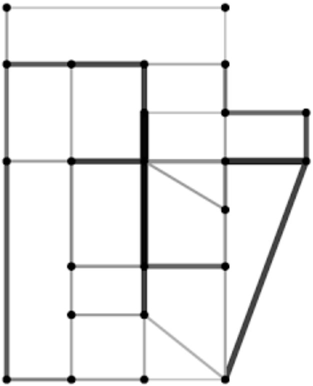
Benefits

- Merging supply and demand aspect of planning
- Accounting for the heterogeneity of demand
- Dealing with complex substitution patterns
- Investigation of demand elasticity against its main driver (e.g. price)

Challenges

- Nonlinearity and nonconvexity
- Assumptions for simple models (logit) may be inappropriate
- Advanced demand models have no closed-form
- Endogeneity: same variable(s) both in the demand function and the cost function

Stochastic traffic assignment



Features

- Nash equilibrium
- Flow problem
- Demand: path choice
- Supply: capacity

Selected literature

- [Dial, 1971]: logit
- [Daganzo and Sheffi, 1977]: probit
- [Fisk, 1980]: logit
- [Bekhor and Prashker, 2001]: cross-nested logit
- and many others...

Revenue management

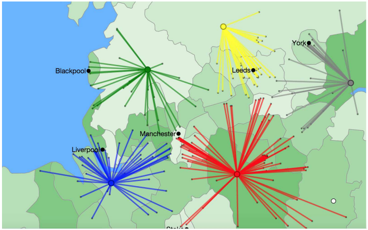
Features

- Stackelberg game
- Bi-level optimization
- Demand: purchase
- Supply: price and capacity

Selected literature

- [Labbé et al., 1998]: bi-level programming
- [Andersson, 1998]: choice-based RM
- [Talluri and Van Ryzin, 2004]: choice-based RM
- [Gilbert et al., 2014a]: logit
- [Gilbert et al., 2014b]: mixed logit
- [Azadeh et al., 2015]: global optimization
- and many others...

Facility location problem



Features

- Competitive market
- Opening a facility impact the costs
- Opening a facility impact the demand
- Decision variables: availability of the alternatives

$$P_n(i|\mathcal{C}_n) = \frac{y_{in} e^{V_{in}}}{\sum_{j \in \mathcal{C}} y_{jn} e^{V_{jn}}}.$$

Selected literature

- [Hakimi, 1990]: competitive location (heuristics)
- [Benati, 1999]: competitive location (B & B, Lagrangian relaxation, submodularity)
- [Serra and Colomé, 2001]: competitive location (heuristics)
- [Marianov et al., 2008]: competitive location (heuristic)
- [Haase and Müller, 2013]: school location (simulation-based)

Outline

A linear formulation

Utility function

$$U_{in} = V_{in} + \varepsilon_{in} = \sum_k \beta_k x_{ink} + f(z_{in}) + \varepsilon_{in}.$$

Simulation

- Assume a distribution for ε_{in}
- E.g. logit: i.i.d. extreme value
- Draw R realizations ξ_{inr}, $r = 1, \ldots, R$
- The choice problem becomes deterministic

Scenarios

Draws

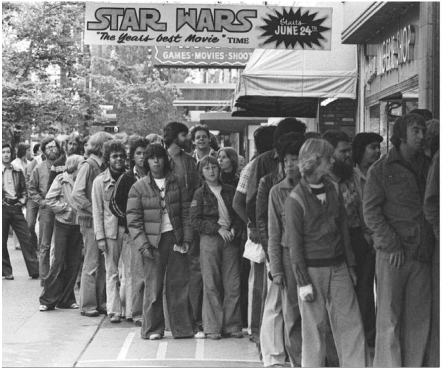
- Draw R realizations ξ_{inr}, $r = 1, \ldots, R$
- We obtain R scenarios

$$U_{inr} = \sum_k \beta_k x_{ink} + f(z_{in}) + \xi_{inr}.$$

- For each scenario r, we can identify the largest utility.
- It corresponds to the chosen alternative.

Capacities

- Demand may exceed supply
- Each alternative i can be chosen by maximum c_i individuals.
- An exogenous priority list is available.
- The numbering of individuals is consistent with their priority.



Priority list

Application dependent

- First in, first out
- Frequent travelers
- Subscribers
- ...

In this framework

The list of customers must be sorted

References

- Technical report: [Bierlaire and Azadeh, 2016]
- TRISTAN presentation: [Pacheco et al., 2016]
- STRC proceeeding: [Pacheco et al., 2017]

Demand model

- Population of N customers (n)
- Choice set $\mathcal{C}$ (i)
- $\mathcal{C}_n \subseteq \mathcal{C}$: alternatives considered by customer n

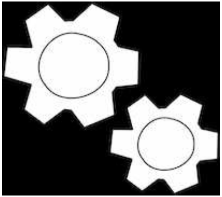
Behavioral assumption

- $U_{in} = V_{in} + \varepsilon_{in}$
- $V_{in} = \sum_k \beta_{ink} x^e_{ink} + q^d(x^d)$
- $P_n(i|\mathcal{C}_n) = \Pr(U_{in} \geq U_{jn}, \forall j \in \mathcal{C}_n)$

Simulation

- Distribution ε_{in}
- R draws $\xi_{in1}, \ldots, \xi_{inR}$
- $U_{inr} = V_{in} + \xi_{inr}$

Supply model



- Operator selling services to a market
  - Price p_{in} (to be decided)
  - Capacity c_i
- Benefit (revenue – cost) to be maximized
- Opt-out option ($i = 0$)

Price characterization

- Lower and upper bound
- Discretization: price levels
- Binary representation ($\lambda_{in\ell}$)

Capacity allocation

- Exogenous priority list of customers
- Here it is assumed as given
- Capacity as decision variable

MILP (in words)

MILP

max	benefit
subject to	utility definition
	availability
	discounted utility
	choice
	capacity allocation
	price selection

Variables

Availability

$y_i \in \{0,1\}$	services proposed by the operator
$y_{in} \in \{0,1\}$	$y_i = 1$ and services considered by customers
$y_{inr} \in \{0,1\}$	capacity restrictions

Utility and choice

U_{inr}	utility
z_{inr}	discounted utility
U_{nr}	maximum discounted utility
$w_{inr} \in \{0,1\}$	choice

Pricing

$\lambda_{in\ell} \in \{0,1\}$	binary representation of the price
$\alpha_{inr\ell} \in \{0,1\}$	linearization of the product $w_{inr}\lambda_{in\ell}$

MILP

MILP

max	benefit
subject to	**utility definition**
	availability
	discounted utility
	choice
	capacity allocation
	price selection

Utility

$$U_{inr} = \overbrace{\beta_{in} p_{in} + q_d(x_d)}^{V_{in}} + \xi_{inr} \quad \forall i, n, r \quad (1)$$

p_{in}	endogenous variable
β_{in}	associated parameter ($\beta_{0n} = 0$)
$q_d(x_d)$	exogenous demand variables

MILP

MILP

max	benefit
subject to	utility definition
	availability
	discounted utility
	choice
	capacity allocation
	price selection

$$y_{in}^d = \begin{cases} 1 & \text{if } i \in \mathcal{C}_n \\ 0 & \text{otherwise} \end{cases} \quad \forall i, n$$

Product of decisions

$$y_{in} = y_{in}^d y_i \quad \forall i, n \tag{2}$$

Availability at operator and scenario level

$$y_{inr} \leq y_{in} \quad \forall i, n, r \tag{3}$$

MILP

MILP

max	benefit
subject to	utility definition
	availability
	discounted utility
	choice
	capacity allocation
	price selection

$$z_{inr} = \begin{cases} U_{inr} & \text{if } y_{inr} = 1 \\ \ell_{nr} & \text{if } y_{inr} = 0 \end{cases} \qquad \forall i, n, r$$

(ℓ_{nr} smallest lower bound)

Discounted utility

$$\ell_{nr} \leq z_{inr} \qquad \forall i, n, r \quad (4)$$

$$z_{inr} \leq \ell_{nr} + M_{inr} y_{inr} \qquad \forall i, n, r \quad (5)$$

$$U_{inr} - M_{inr}(1 - y_{inr}) \leq z_{inr} \qquad \forall i, n, r \quad (6)$$

$$z_{inr} \leq U_{inr} \qquad \forall i, n, r \quad (7)$$

MILP

MILP

max	benefit
subject to	utility definition
	availability
	discounted utility
	choice
	capacity allocation
	price selection

$$U_{nr} = \max_{i \in \mathcal{C}} z_{inr} \qquad \forall n, r$$

$$w_{inr} = \begin{cases} 1 & \text{if } i = \operatorname{argmax}\{U_{nr}\} \\ 0 & \text{otherwise} \end{cases} \qquad \forall i, n, r$$

Choice

$$z_{inr} \leq U_{nr} \qquad \forall i, n, r \qquad (8)$$

$$U_{nr} \leq z_{inr} + M_{nr}(1 - w_{inr}) \qquad \forall i, n, r \qquad (9)$$

$$\sum_i w_{inr} = 1 \qquad \forall n, r \qquad (10)$$

$$w_{inr} \leq y_{inr} \qquad \forall i, n, r \qquad (11)$$

MILP

MILP

max	benefit
subject to	utility definition
	availability
	discounted utility
	choice
	capacity allocation
	price selection

Capacity allocation

- Priority list
- Two sets of constraints $\forall i > 0$
  - Capacity cannot be exceeded ($\Rightarrow y_{inr} = 1$)
  - Capacity has been reached ($\Rightarrow y_{inr} = 0$)

Price: linearization

$$\mathrm{LP}_{in} w_{inr} \leq \alpha_{inr} \leq \mathrm{UP}_{in} w_{inr}$$

$$p_{in} - (1 - w_{inr})\mathrm{UP}_{in} \leq \alpha_{inr} \leq p_{in} - (1 - w_{inr})\mathrm{LP}_{in}$$

MILP

MILP

max	**benefit**
subject to	utility definition
	availability
	discounted utility
	choice
	capacity allocation
	price selection

$$\max \sum_{i>0} (R_i - C_i)$$

Revenue

$$R_i = \frac{1}{R} \sum_n \sum_r p_{in} w_{inr},$$

Cost

$$C_i = (f_i + v_i c_i) y_i$$

Outline

A simple example

Context

- $\mathcal{C}$: set of movies
- Population of N individuals
- Competition: stying home watching TV

One theater – homogenous population

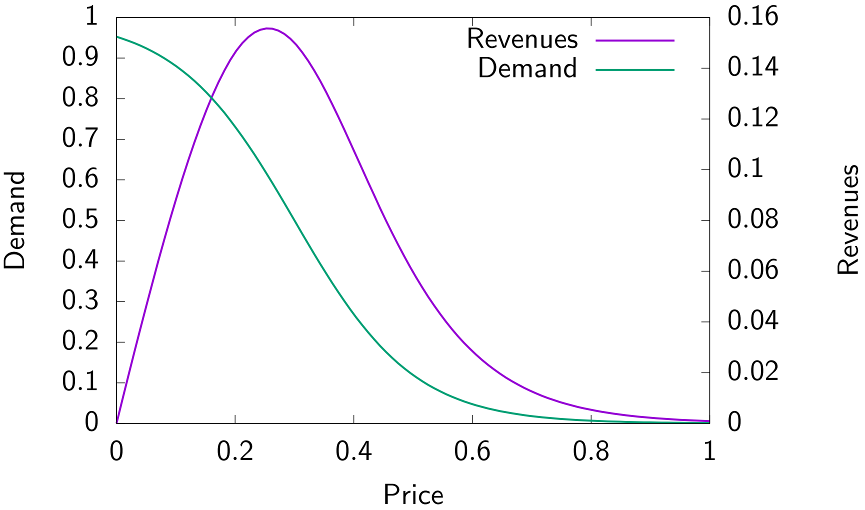
Alternatives

- Staying home: $U_{cn} = 0 + \varepsilon_{cn}$
- My theater: $U_{mn} = -10.0 p_m + 3 + \varepsilon_{mn}$

Logit model

ε_m i.i.d. EV(0,1)

Demand and revenues



Optimization

Solver

GLPK v4.61 under PyMathProg

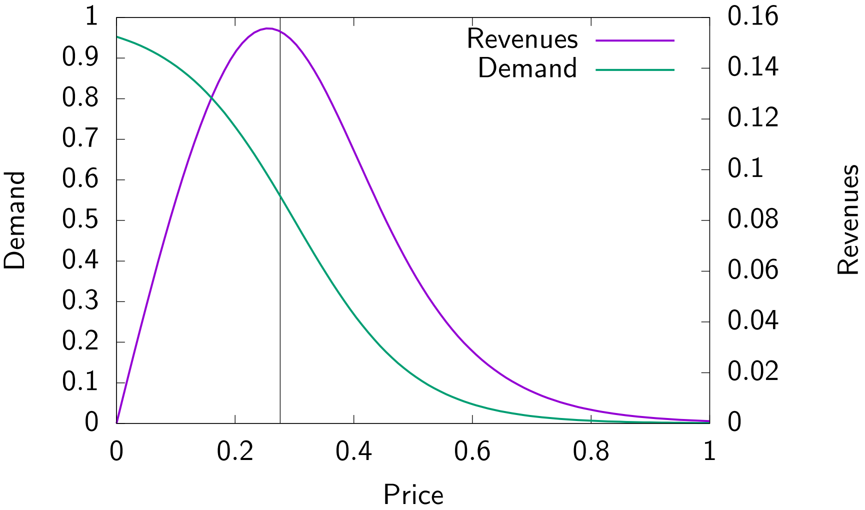
Data

- $N = 1$
- $R = 1000$

Results

- Optimum price: 0.276
- Demand: 57.4%
- Revenues: 0.159

Demand and revenues



Heterogeneous population

Two groups in the population

$$U_{mn} = -\beta_n p_m + c_n$$

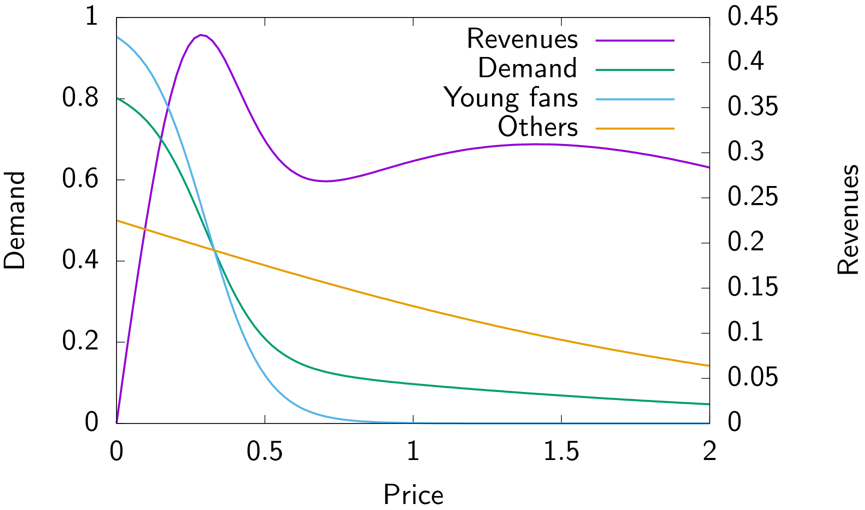
Young fans: 2/3

$\beta_1 = -10,\ c_1 = 3$

Others: 1/3

$\beta_2 = -0.9,\ c_2 = 0$

Demand and revenues



Optimization

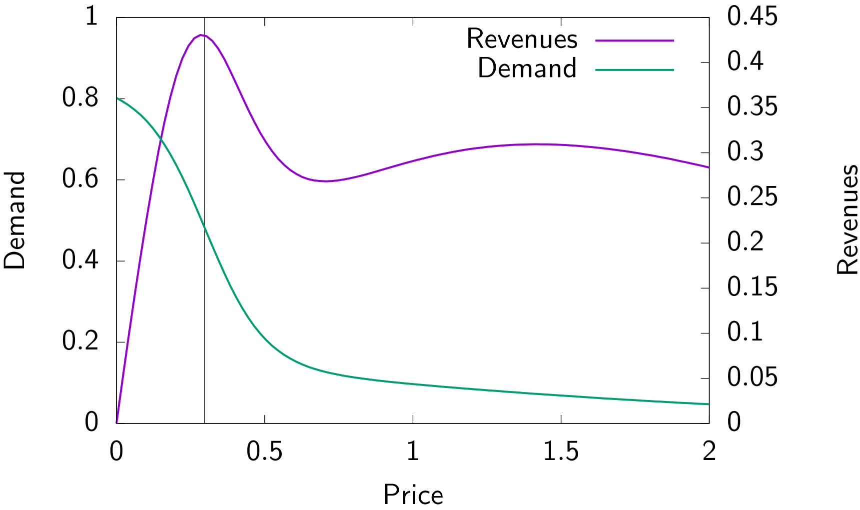
Data

- $N = 3$
- $R = 500$

Results

- Optimum price: 0.297
- Customer 1 (fan): 52.4% [theory: 50.8 %]
- Customer 2 (fan) : 49% [theory: 50.8 %]
- Customer 3 (other) : 45.8% [theory: 43.4 %]
- Demand: 1.472 (49%)
- Revenues: 0.437

Demand and revenues



Two theaters, different types of films

Two theaters, different types of films

Theater m

- Attractive for young people
- Star Wars Episode VII

Theater k

- Not particularly attractive for young people
- Tinker Tailor Soldier Spy

Heterogeneous demand

- Two third of the population is young (price sensitive)
- One third of the population is not (less price sensitive)

Two theaters, different types of films

Data

- Theaters m and k
- $N = 9$
- $R = 50$
- $U_{mn} = -10p_m + 4$, n =young
- $U_{mn} = -0.9p_m$, n =others
- $U_{kn} = -10p_k + 0$, n =young
- $U_{kn} = -0.9p_k$, n =others

Theater m

- Optimum price m: 0.390
- Young customers: 3.48 / 6
- Other customers: 1.08 / 3
- Demand: 4.56 (50.7%)
- Revenues: 1.779

Theater k

- Optimum price k: 1.728
- Young customers: 0.0 / 6
- Other customers: 0.38 / 3
- Demand: 0.38 (4.2%)
- Revenues: 0.581

Two theaters, same type of films

Theater m

- Expensive
- Star Wars Episode VII

Theater k

- Cheap (half price)
- Star Wars Episode VIII

Heterogeneous demand

- Two third of the population is young (price sensitive)
- One third of the population is not (less price sensitive)

Two theaters, same type of films

Data

- Theaters m and k
- $N = 9$
- $R = 50$
- $U_{mn} = -10p + 4$, $n =$young
- $U_{mn} = -0.9p$, $n =$others
- $U_{kn} = -10p/2 + 4$, $n =$young
- $U_{kn} = -0.9p/2$, $n =$others

Theater m

- Optimum price m: 3.582
- Young customers: 0
- Other customers: 1.9
- Demand: 1.9 (31.7%)
- Revenues: 3.42

Theater k

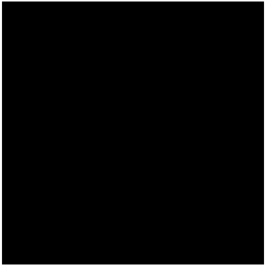
Closed

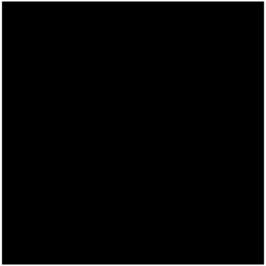
Outline

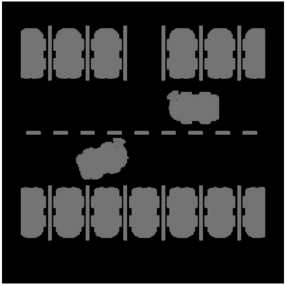
Challenge

- Select a real choice model from the literature
- Integrate it in an optimization problem.

Parking choices







- $N = 50$ customers
- $\mathcal{C} = \{\text{PSP}, \text{PUP}, \text{FSP}\}$
- $\mathcal{C}_n = \mathcal{C} \quad \forall n$

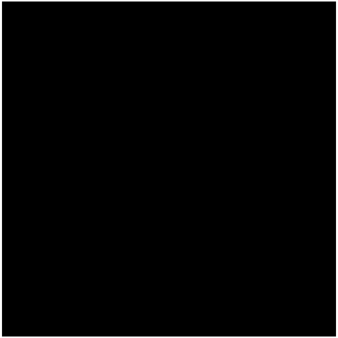
- PSP: $0.50, 0.51, \ldots, 0.65$ (16 price levels)
- PUP: $0.70, 0.71, \ldots, 0.85$ (16 price levels)
- Capacity of 20 spots

Choice model: mixtures of logit model [Ibeas et al., 2014]

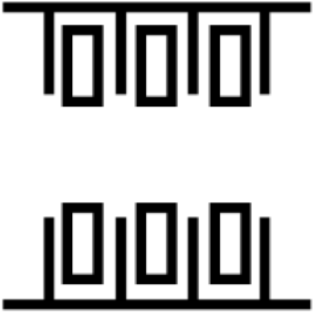
$$
\begin{aligned}
V_{FSP} &= \beta_{AT} AT_{FSP} + \boxed{\beta_{TD}} TD_{FSP} + \boxed{\beta_{Origin_{INT\_FSP}}} Origin_{INT\_FSP} \\
V_{PSP} &= \boxed{ASC_{PSP}} + \beta_{AT} AT_{PSP} + \boxed{\beta_{TD}} TD_{PSP} + \beta_{FEE} \mathbf{FEE_{PSP}} \\
&\quad + \boxed{\beta_{FEE_{PSP(LowInc)}}} \mathbf{FEE_{PSP}} LowInc + \boxed{\beta_{FEE_{PSP(Res)}}} \mathbf{FEE_{PSP}} Res \\
V_{PUP} &= \boxed{ASC_{PUP}} + \beta_{AT} AT_{PUP} + \boxed{\beta_{TD}} TD_{PUP} + \beta_{FEE} \mathbf{FEE_{PUP}} \\
&\quad + \boxed{\beta_{FEE_{PUP(LowInc)}}} \mathbf{FEE_{PUP}} LowInc + \boxed{\beta_{FEE_{PUP(Res)}}} \mathbf{FEE_{PUP}} Res \\
&\quad + \boxed{\beta_{AgeVeh_{\leq 3}}} AgeVeh_{\leq 3}
\end{aligned}
$$

- Parameters
  - Circle: distributed parameters
  - Rectangle: constant parameters
- Variables: all given but FEE (in bold)

Experiment 1: uncapacitated vs capacitated case (1)



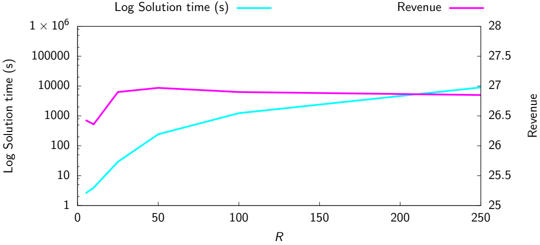
- Capacity constraints are ignored
- Unlimited capacity is assumed



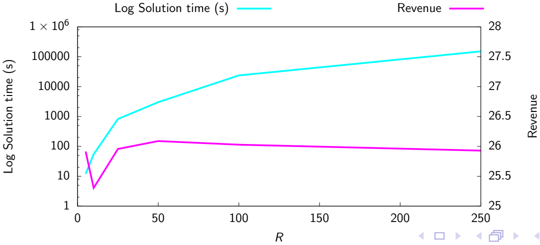
- 20 spots for PSP and PUP
- Opt-out has unlimited capacity

Experiment 1: uncapacitated vs capacitated case (2)

Uncapacitated

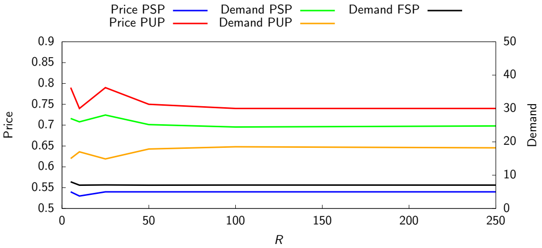


Capacitated

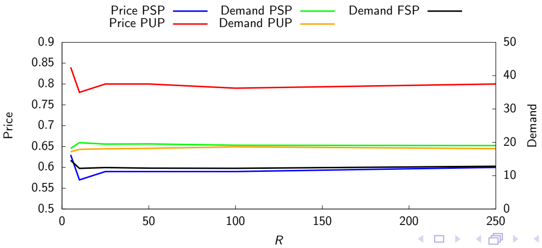


Experiment 1: uncapacitated vs capacitated case (3)

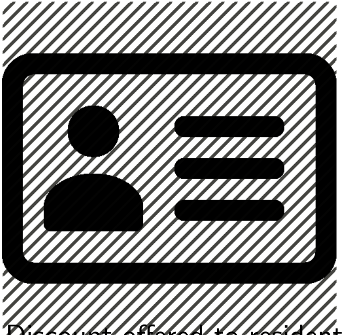
Uncapacitated

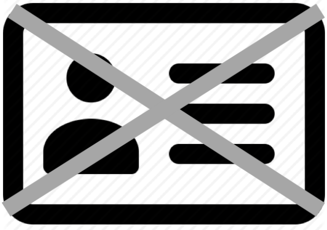


Capacitated



Experiment 2: price differentiation by segmentation (1)

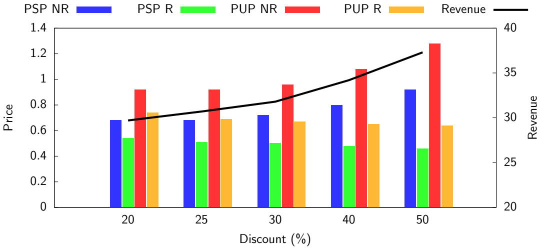




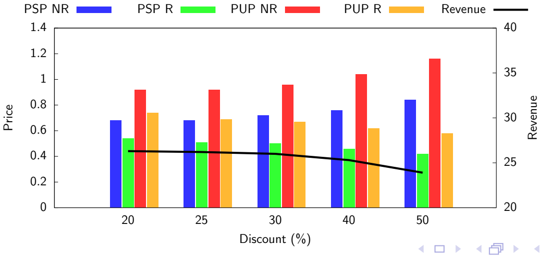
- Discount offered to residents
- Two scenarios (municipality)
  1. Subsidy offered by the municipality
  2. Operator obliged to offer reduced fees
- We expect the price to increase
  - PSP: $\{0.60, 0.64, \ldots, 1.20\}$
  - PUP: $\{0.80, 0.84, \ldots, 1.40\}$

Experiment 2: price differentiation by segmentation (2)

Scenario 1

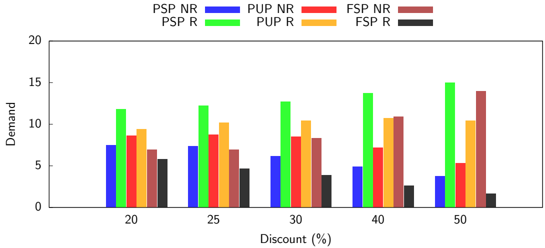


Scenario 2

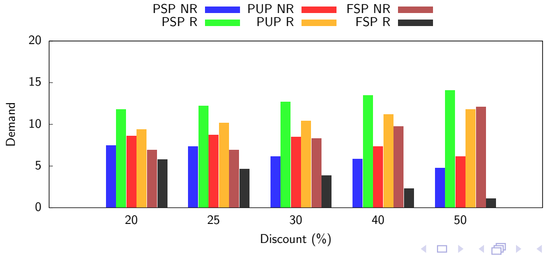


Experiment 2: price differentiation by segmentation (3)

Scenario 1



Scenario 2



Other experiments

Impact of the priority list

- Priority list = order of the individuals in the data (i.e., random arrival)
- 100 different priority lists
- Aggregate indicators remain stable across random priority lists

Benefit maximization through capacity allocation

- 4 different capacity levels for both PSP and PUP: 5, 10, 15 and 20
- Optimal solution: PSP with 20 spots and PUP is not offered
- Both services have to be offered: PSP with 15 and PUP with 5

Outline

Summary

Demand and supply

- Supply: prices and capacity
- Demand: choice of customers
- Interaction between the two

Discrete choice models

- Rich family of behavioral models
- Strong theoretical foundations
- Great deal of concrete applications
- Capture the heterogeneity of behavior
- Probabilistic models

Optimization

Discrete choice models

- Non linear and non convex
- Idea: use utility instead of probability
- Rely on simulation to capture stochasticity

Proposed formulation

- Linear in the decision variables
- Large scale
- Fairly general

Ongoing research

- Decomposition methods
- Scenarios are (almost) independent from each other (except objective function)
- Individuals are also loosely coupled (except for capacity constraints)

Bibliography I

Andersson, S.-E. (1998).
Passenger choice analysis for seat capacity control: A pilot project in scandinavian airlines.
International Transactions in Operational Research, 5(6):471–486.

Azadeh, S. S., Marcotte, P., and Savard, G. (2015).
A non-parametric approach to demand forecasting in revenue management.
Computers & Operations Research, 63:23–31.

Bekhor, S. and Prashker, J. (2001).
Stochastic user equilibrium formulation for generalized nested logit model.
Transportation Research Record: Journal of the Transportation Research Board, (1752):84–90.

Bibliography II

Benati, S. (1999).
The maximum capture problem with heterogeneous customers.
Computers & operations research, 26(14):1351–1367.

Bierlaire, M. and Azadeh, S. S. (2016).
Demand-based discrete optimization.
Technical Report 160209, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

Daganzo, C. F. and Sheffi, Y. (1977).
On stochastic models of traffic assignment.
Transportation science, 11(3):253–274.

Bibliography III

Dial, R. B. (1971).
A probabilistic multipath traffic assignment model which obviates path enumeration.
Transportation research, 5(2):83–111.

Fisk, C. (1980).
Some developments in equilibrium traffic assignment.
Transportation Research Part B: Methodological, 14(3):243–255.

Gilbert, F., Marcotte, P., and Savard, G. (2014a).
Logit network pricing.
Computers & Operations Research, 41:291–298.

Gilbert, F., Marcotte, P., and Savard, G. (2014b).
Mixed-logit network pricing.
Computational Optimization and Applications, 57(1):105–127.

Bibliography IV

Haase, K. and Müller, S. (2013).
Management of school locations allowing for free school choice.
Omega, 41(5):847–855.

Hakimi, S. L. (1990).
Locations with spatial interactions: competitive locations and games.
Discrete location theory, pages 439–478.

Ibeas, A., dell'Olio, L., Bordagaray, M., and de D. OrtÃºzar, J. (2014).
Modelling parking choices considering user heterogeneity.
Transportation Research Part A: Policy and Practice, 70:41 – 49.

Bibliography V

Labbé, M., Marcotte, P., and Savard, G. (1998).
A bilevel model of taxation and its application to optimal highway pricing.
Management science, 44(12-part-1):1608–1622.

Marianov, V., Ríos, M., and Icaza, M. J. (2008).
Facility location for market capture when users rank facilities by shorter travel and waiting times.
European Journal of Operational Research, 191(1):32–44.

Pacheco, M., Azadeh, S. S., Bierlaire, M., and Gendron, B. (2017).
Integrating advanced demand models within the framework of mixed integer linear problems: A lagrangian relaxation method for the uncapacitated case.
In *Proceedings of the 17th Swiss Transport Research Conference*, Ascona, Switzerland.

Bibliography VI

Pacheco, M., Bierlaire, M., and Azadeh, S. S. (2016).
Incorporating advanced behavioral models in mixed linear optimization.
Presented at TRISTAN IX, Oranjestad, Aruba.

Serra, D. and Colomé, R. (2001).
Consumer choice and optimal locations models: formulations and heuristics.
Papers in Regional Science, 80(4):439–464.

Talluri, K. and Van Ryzin, G. (2004).
Revenue management under a general discrete choice model of consumer behavior.
Management Science, 50(1):15–33.