Introducing a Non Parametric Method for Customer Behavior Modeling

Shadi Sharif Azadeh

EPFL TRANSP-OR

19 June 2014

Workshop on Discrete Choice Models 2014

Context

Question: Why look alike products have different prices?

Context

- What is a product?





- Price
- Reservation interval
- Departure day
- Departure time
- Number of stops
- Origin-destination

- How many choices do we have?


Specialty Coffees

1.	House Coffee (Regular/Flavored)	1.50	1.75
2.	Hot Cocoa (Hot Chocolate Made w/ Milk)	1.75	2.25
3.	Hot Tea (Selection of Various Teas)	1.50	1.75
4.	Tea Au Lait (Hot Tea w/Steamed Milk)	1.75	2.25
5.	Espresso	1.50	1.75
6.	Coffee Americano	1.60	1.85
7.	Hammerhead (House Coffee Topped w/ Espresso)	1.85	2.25
8.	Cappuccino (Espresson. Steamed Milk w/ Froth)	2.25	2.75
9.	Coffee Latte (Espresso Layered w/ Steamed milk	2.50	3.00
10.	Sweet Latte (LatteInfused w/ Vanilla or Hazelnut)	2.75	3.25

- Who is 'the customer'?






Presentation Plan

Introduction		Definition of censored demand
Context 1	Part 1	Motivation of research: Theoretical, Application, Objectives
	Part 2	A Taxonomy of Demand Uncensoring Methods in Revenue Management
	Part 3	Mathematical model, Algorithm and Computational results
Context 2	Part 4	Motivation of research: Theoretical, Application, Objectives
	Part 5	Mathematical model, Algorithm and Computational results
Remarks		Conclusion and discussion

Introduction

- **RM systems:**

  The application of disciplined tactics that predict consumer behaviour at the micro-market level and optimize product availability and price to maximize revenue growth (Robert Cross, 1998)

  Recourse: Historical Data

  Main issue: Censored data and not having access to the customers' preferences

- **A Good Demand Model in RMS:**

  The optimization tool of demand model relies on the demand model.

  1) Censored demand

  2) Internal segmentation (Customer segmentation):

  "A Market segment is a subgroup of people or organizations sharing one or more characteristics that cause them to have similar product needs"

  Taking competition factor and buy-ups into account by accurately representing customer behavior (via utility estimation)

  3) External segmentation (Clustering):

  Capturing seasonal effects

Censored Demand

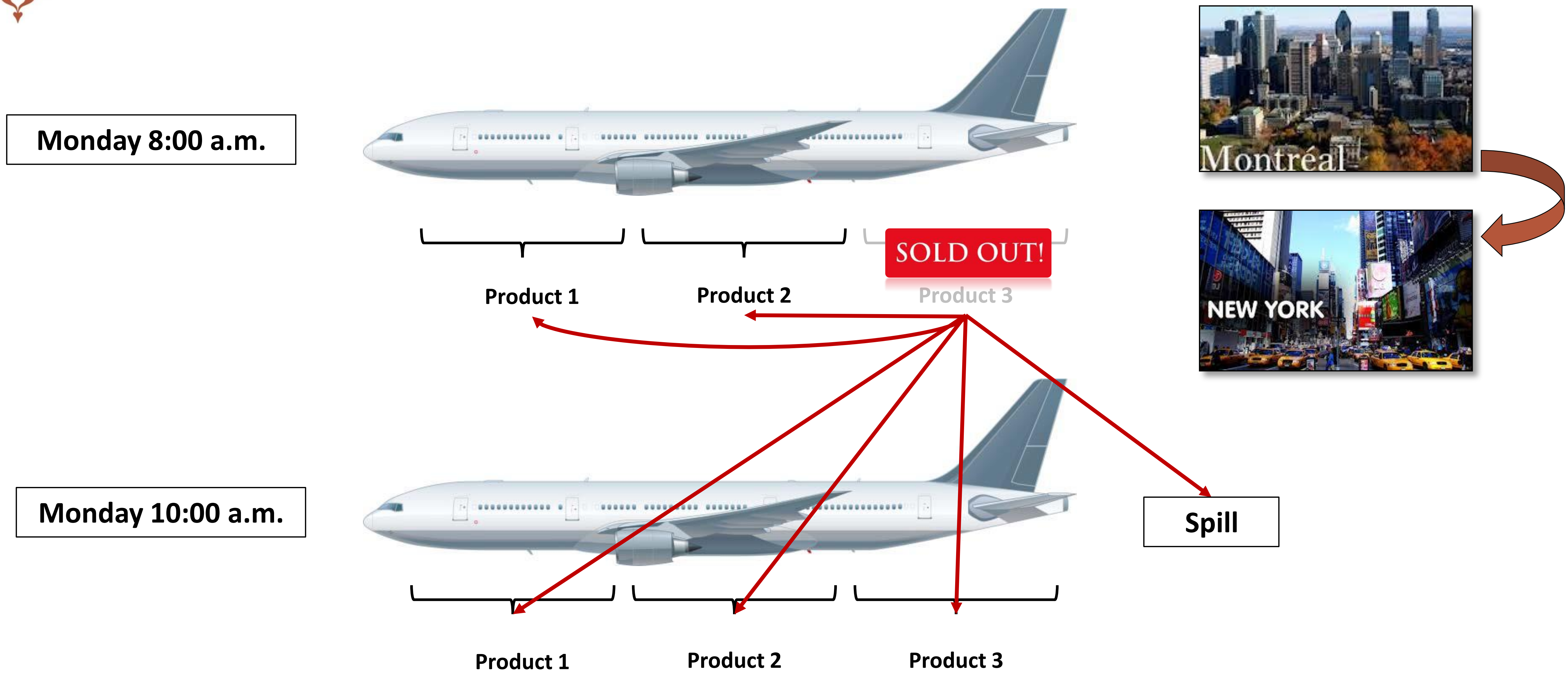
Monday 8:00 a.m.

Product 1 Product 2 Product 3

Monday 10:00 a.m.

Product 1 Product 2 Product 3

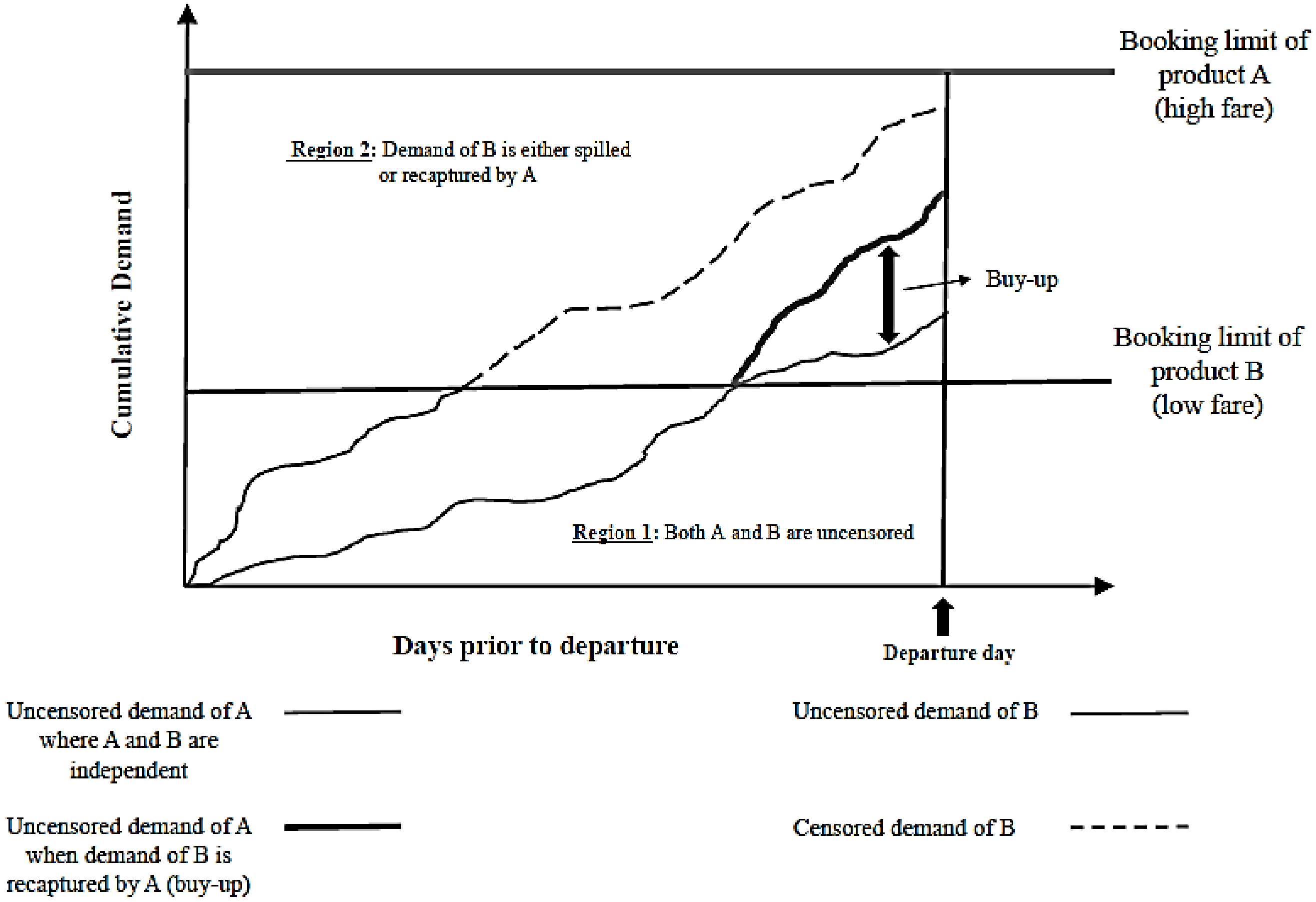
Censored Demand

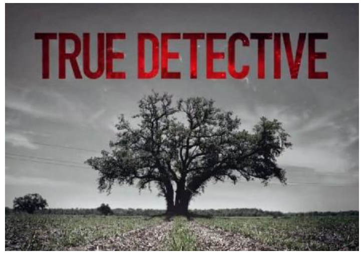


Motivation

Impact of censored demand on Revenue:

Underestimating demand by 12.5% to 25% can result in a loss of revenue from 1% to 3%, which is significant Weatherford and Belobaba (2002).

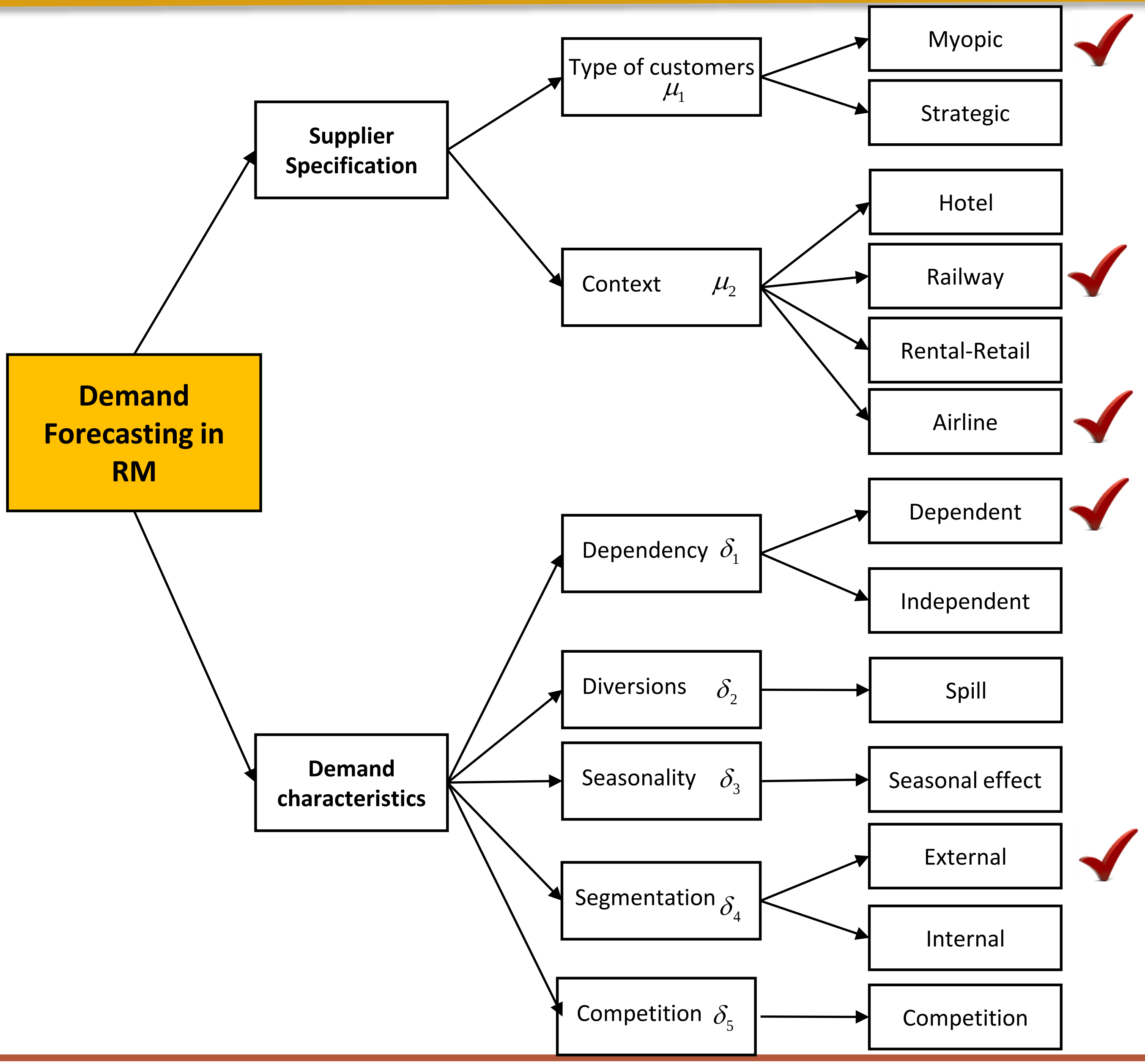




HBO GO is the successor to HBO on Broadband, originally launched in January 2008, consisting of 400 hours of movies, specials and original series (including 130 movie titles that rotate monthly) that could be downloaded to computers, at no extra charge for HBO subscribers. Meltdowns in HBO Go happen usually on Sundays and it affects the stock market. That's only because HBO is now the prime destination for some of the greatest TV shows existing today. But when you see their server crash more than once, you have to wonder whether the demand for certain TV shows is being continually underestimated.

Ref: Greg Brian Apr 10, 2014

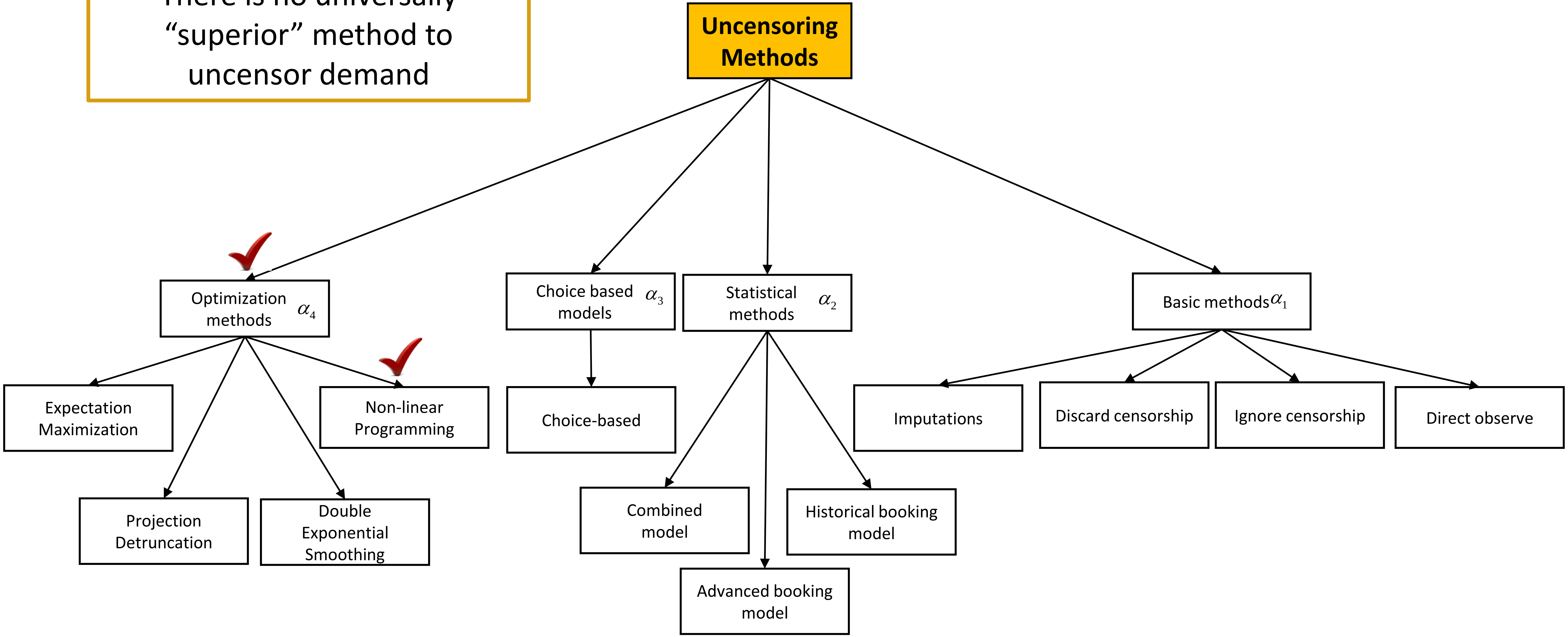
Elements of Demand Forecasting in RMS



Ref: Sh. Sharif, P. Marcotte, G. Savard. A taxonomy of demand uncensoring methods in RMS. *Journal of Revenue Management and Pricing* (2014).

Methods of Uncensoring Demand

There is no universally “superior” method to uncensor demand



Context 1

Universal Utility Assumption and External Segmentation

Sh. Sharif Azadeh, P. Marcotte, G. Savard. *A Non-Parametric Method to Demand Forecasting in Revenue Management Systems. Under Review. C&OR (2014)*

Objectives

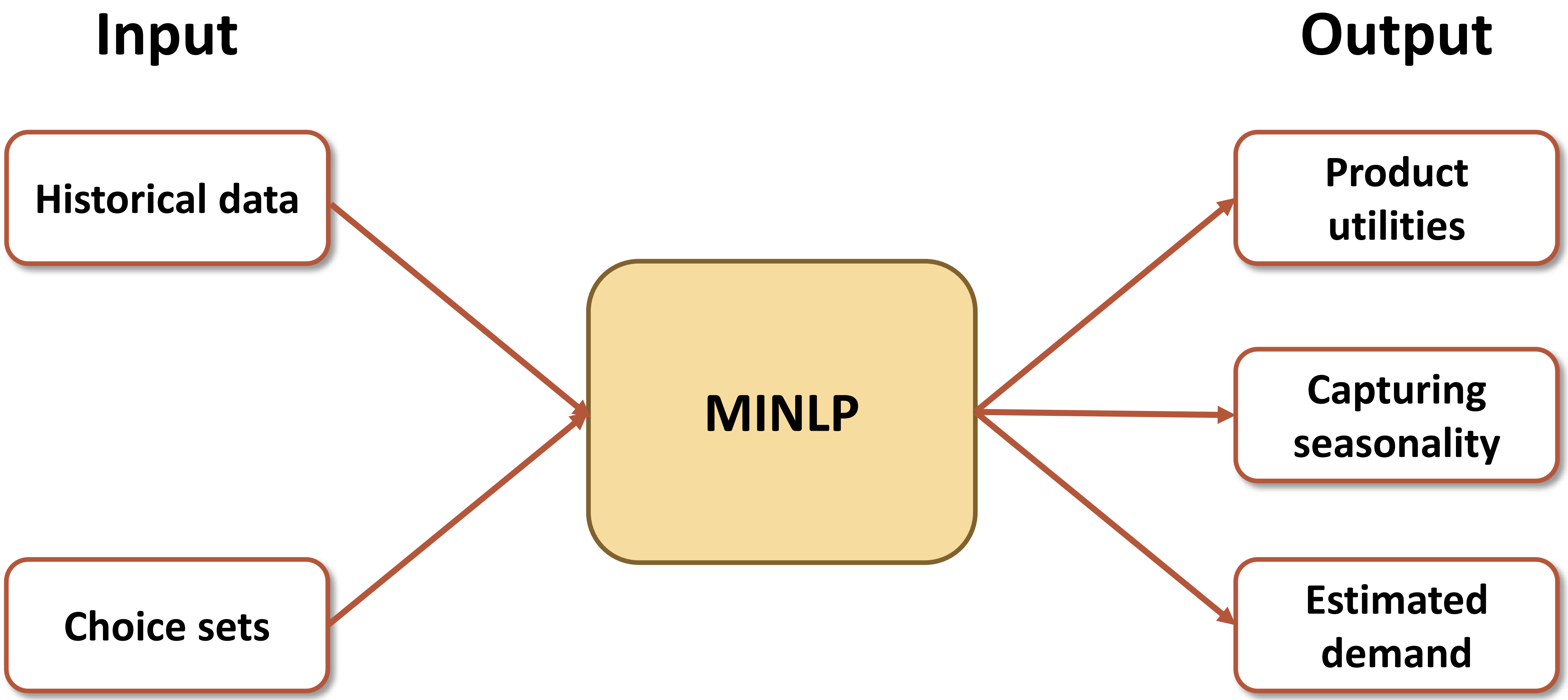
- **The main objectives:**

1. To avoid primary demand and stochastic process assumption (Poisson).
2. To estimate daily potential demand for a given O-D .
3. To obtain choice probabilities based on the estimated utilities for all products.
4. To capture seasonal effects by clustering departure days.

- **The forecasting aspect:**

By modifying the choice set, we can run the simulation again and observe how demand is expected to react to the modifications, for instance the closing of a booking class.

Objectives


Input
Output
Historical data
Choice sets
MINLP
Product utilities
Capturing seasonality
Estimated demand

Original Mathematical Model

$$\text{MINLP:} \quad \min_{\delta_c, u, z} \sum_{j \in J} \sum_{i \in S_j} (p_{ij}(S_j, u_i) d_j A_{ij} - O_{ij})^2$$

$$\text{subject to} \quad p_{ij}(S_j, u_i) = \frac{\exp(u_i)}{\sum_{k \in S_j} \exp(u_k) + \exp(u_0)} \quad i \in S_j$$

$$d_j = \sum_c \delta_c z_{jc} \quad j \in J$$

$$\sum_{c \in C} z_{jc} = 1 \quad j \in J$$

$$z_{jc} \in \{0, 1\} \quad j \in J, c \in C.$$

- **The variables:**
  1. External segmentation (Class membership variable) z_{jc}
  2. Product utilities u_i
  3. Potential demand of each cluster δ_c

- **Challenges of solving MINLP**

  Highly non-convex even its continuous relaxation. Global and nonlinear optimizers fail to precisely solve the problem.

General Algorithm

Input: Registered bookings O_{ij}, set of available products, S_j

Output: Daily demand flows d_j, cluster memberships z_{jc}, utilities u_i

(1) Transformation into a MIP

- i: Linearization
- ii: Relaxation
- ii: Convexification

(2) Preprocessing at root node

- iii: Valid inequalities
- iv: Initial solution
- v: Domain reduction

(3) Branch-and-bound

- vi: Branching strategy
- vii: Adjustment of bounds at branching nodes

Transformation to MIP- Notations

sets

Product	$i \in I = \{1, \dots, \lvert I \rvert\}$
Day	$j \in J = \{1, \dots, \lvert J \rvert\}$
Cluster	$c \in C = \{1, \dots, \lvert C \rvert\}$
Choice set	S_j, set of products available on day j

parameters

O_{ij}	observed bookings for product $i \in I$ on day $j \in J$
A_{ij}	availability status of product $i \in I$ on day $j \in J$
R_c^U	upper bound on potential demand for cluster $c \in C$
R_c^L	lower bound on potential demand for cluster $c \in C$
D_j^U	upper bound on potential demand on day $j \in J$
D_j^L	lower bound on potential demand on day $j \in J$
P_{ij}^U	upper bound on choice probability for product $i \in I$ on day $j \in J$
P_{ij}^U	lower bound on choice probability of product $i \in I$ on day $j \in J$

variables

e_{ij}	difference between estimated demand w_{ij} and observed bookings O_{ij}
w_{ij}	expected demand for product $i \in I$ on day $j \in J$
d_j	daily potential demand (integer)
d_{jc}^N	normalized daily potential demand $\in [0, 1]$
p_{ij}	probability of selecting product $i \in I$ on day $j \in J$
z_{jc}	cluster membership variable (binary)
r_c^N	normalized potential demand for each cluster $\in [0, 1]$
u_i	utility of product i
δ_c	potential demand of cluster c

Transformation to MIP-Linearization,Relaxation,Convexification

$$\text{RELAX:} \quad \min_{p,d,z} \sum_{i \in I} \sum_{j \in J} e_{ij}^2$$

$$w_{ij} A_{ij} - O_{ij} = e_{ij} \qquad i \in I, j \in J$$

Estimation

$$P_{ij}^U d_j + D_j^U p_{ij} - P_{ij}^U D_j^U \leq w_{ij} \qquad i \in I, j \in J$$

$$P_{ij}^L d_j + D_j^L p_{ij} - P_{ij}^L D_j^L \leq w_{ij} \qquad i \in I, j \in J$$

$$P_{ij}^U d_j + D_j^L p_{ij} - P_{ij}^U D_j^L \geq w_{ij} \qquad i \in I, j \in J$$

$$P_{ij}^L d_j + D_j^U p_{ij} - P_{ij}^L D_j^U \geq w_{ij} \qquad i \in I, j \in J$$

Classification

$$d_{jc}^N \leq z_{jc} \qquad j \in J, c \in C$$

$$z_{jc} + r_c^N \leq d_{jc}^N + 1 \qquad j \in J, c \in C$$

$$z_{jc} + d_{jc}^N \leq r_c^N + 1 \qquad j \in J, c \in C$$

$$d_j = \sum_{c \in C} d_{jc}^N R_c^U \qquad j \in J$$

$$R_c^L / R_c^U \leq r_c^N \qquad c \in C$$

$$\sum_{c \in C} z_{jc} = 1 \qquad j \in J$$

$$0 \leq z_{jc} \leq 1 \qquad j \in J, c \in C$$

General Algorithm

Input: Registered bookings O_{ij}, set of available products, S_j

Output: Daily demand flows d_j, cluster memberships z_{jc}, utilities u_i

(1) Transformation into a MIP

- i: Linearization
- ii: Relaxation
- ii: Convexification

(2) Preprocessing at root node

- iii: Valid inequalities
- iv: Initial solution
- v: Domain reduction

(3) Branch-and-bound

- vi: Branching strategy
- vii: Adjustment of bounds at branching nodes

Branch and Bound-Feasibility Conditions at a Branching Node

- For the RELAX model solved by linear solver (CPLEX):

$$r_1^N R_1^U \leq r_2^N R_2^U \leq \ldots \leq r_k^N R_k^U \leq \ldots \leq r_{|C|}^N R_{|C|}^U.$$

- For the original MINLP model solved by nonlinear solver (IPOPT):

$$P_{ij}^L \leq \frac{\exp(u_i)}{\sum_{k \in S_j} \exp(u_k) + \exp(u_0)} \leq P_{ij}^U \quad \forall i \in I$$

Branch and Bound-Branching Strategy

1- Potential demand of each cluster δ_c , stopping criterion: disjoint cluster ranges

Variable selection:

Interval length

$$I_c(\hat{d}_j(n)) = \begin{cases} 1 & \text{if} \quad \hat{d}_j(n) \in [R_c^L(n), R_c^U(n)] \qquad c \in C \\ 0 & \text{otherwise} \end{cases}$$

$$I(n) = \sum_{j \in J} I_c(\hat{d}_j(n)) \qquad c \in C.$$

2- Daily potential demand d_j , stopping criterion: assigning all days to clusters

Variable selection:

Interval length

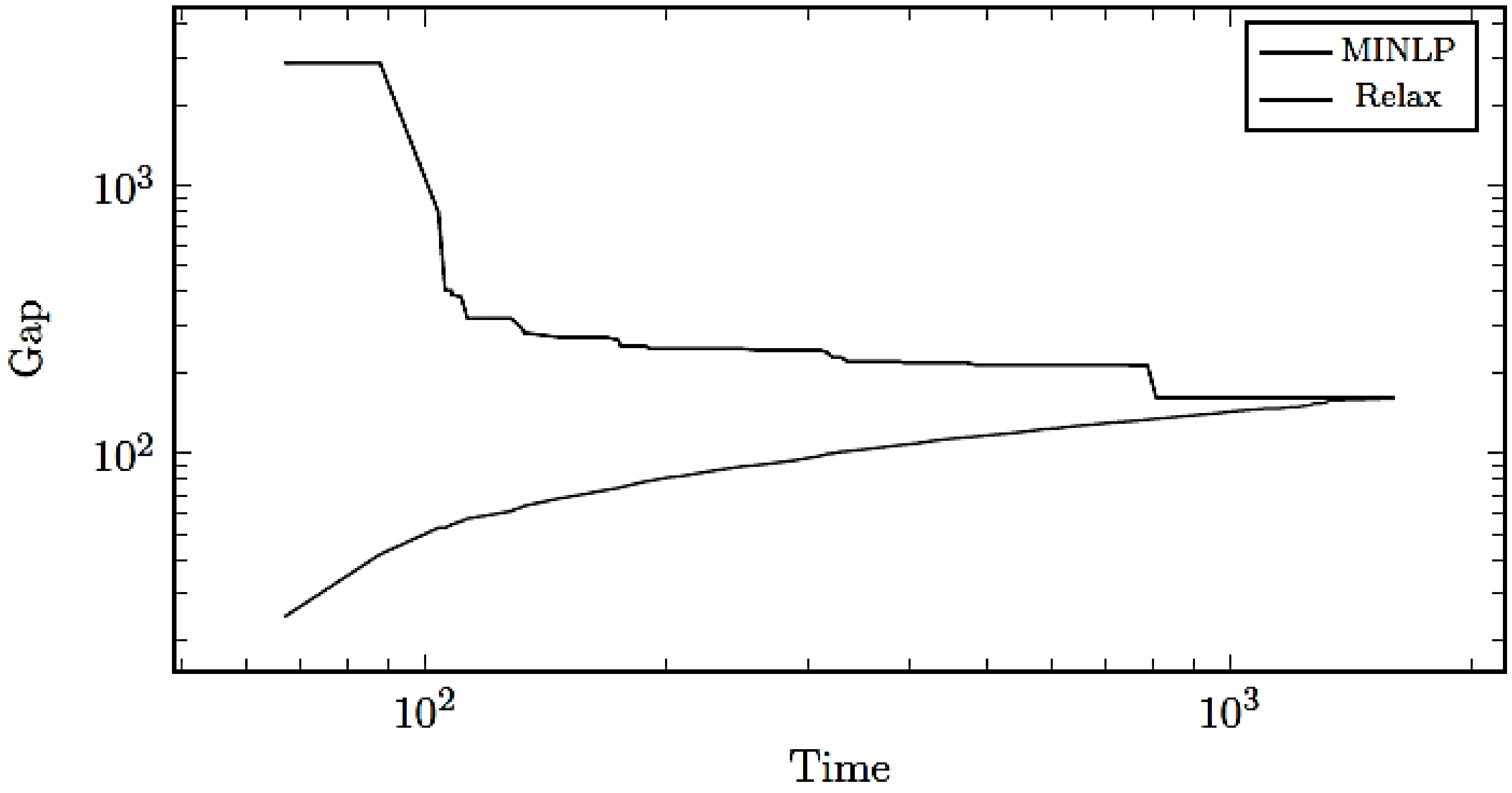
Choosing a d_j with more available products

3- Potential demand of each cluster δ_c , stopping criterion: cluster ranges reaching a singleton

4- Solving MINLP model to find product utilities u_i

5- Branching on choice probabilities p_{ij} , in case of a large gap between RELAX and MINLP solutions

Computational Results-Algorithm Effectiveness



Computational Results- Perturbed Data

			Time			Node										
Class	Days	Product	Total	Relax	NLP	Pre-Proc.	Gen.	Br.	Dis.	Dom.	#NLP	VI#1	Ini Sol.	Best	Bound	Gap(%)
		4	0.70	0.48	0.00	1.14	125	54	26	24	6	32	154.55	5.70	5.68	0.02
	7	6	4.06	2.31	0.11	2.64	539	438	30	34	228	38	761.24	7.53	6.96	0.57
		8	4.23	2.59	0.17	4.19	468	362	36	51	173	44	561.02	6.66	6.30	0.36
		4	4.33	2.30	0.13	4.59	436	339	40	55	168	94	536.44	30.75	30.01	0.74
	14	6	15.69	7.53	0.36	8.77	1089	902	40	137	467	115	1589.76	28.70	27.90	0.80
2		8	9.89	4.80	0.14	70.59	590	476	61	46	234	109	1212.62	16.56	15.72	0.84
		4	4238.61	1243.47	0.55	3.94	121336	71764	5178	40387	916	90	1361.45	11.87	11.03	0.84
	14	6	4247.70	1500.89	0.84	7.80	110504	62968	3582	31329	1104	78	1688.36	12.60	9.43	3.17
		8	7302.53	1864.11	11.11	12.41	155860	93126	4153	41149	13749	50	859.01	12.13	7.45	4.68
		4	8083.58	2306.41	11.50	11.52	180764	102470	5470	60386	15246	169	248.52	39.63	37.23	2.40
4	21	6	6189.22	1774.72	9.89	17.52	118253	72650	4220	23518	13046	165	1328.48	34.23	29.72	4.51
		8	5645.56	2245.20	10.13	29.92	140178	94594	4552	27345	14177	102	1696.79	38.47	33.85	4.62
		4	4064.64	1112.59	15.48	23.00	94901	60890	2288	13886	19503	386	416.61	53.68	49.64	4.04
	28	6	9252.73	3292.19	7.67	43.23	132019	87375	5832	29557	13077	343	953.01	56.95	52.10	4.85
		8	11689.00	3746.00	9.83	296.54	147664	90428	6373	32346	15746	368	1234.02	45.63	40.93	4.70

$$d_j = d_j(1+\gamma(2\epsilon_j - 1))$$

Comparing with KNITRO and BARON

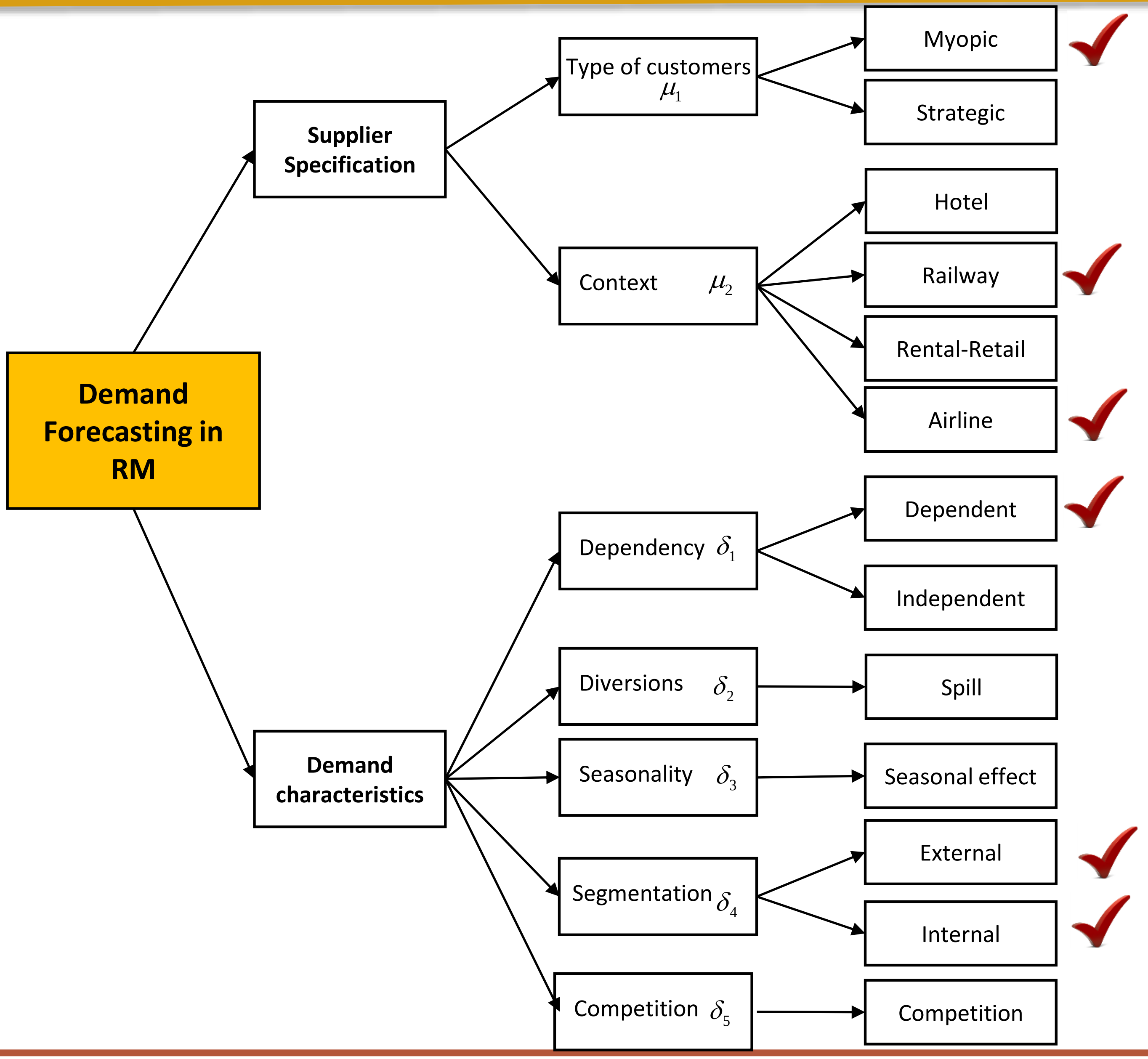
			BB Sol			Baron		Knitro		Generalization
Class	Days	Product	Ini Sol	MSE	Class.(%)	MSE	Class.(%)	MSE	Class.(%)	MSE
2	7	4	154.55	5.70	0.00	5.70	0.00	5.63	0.00	32.95
		6	761.24	7.53	0.00	7.53	0.00	7.53	0.00	22.17
		8	561.02	6.66	0.00	8.08	14.29	8.12	14.29	14.01
	14	4	536.44	30.75	0.00	3682.29	28.57	157.58	21.43	54.58
		6	1589.76	28.70	0.00	3441.98	50.00	147.56	42.86	37.94
4	14	4	1361.45	11.87	0.00	796.87	64.29	456.36	42.86	23.67
		6	1688.36	12.60	0.00	5620.43	57.14	2345.00	57.14	69.58
		8	859.01	12.13	0.00	4235.06	71.43	4235.06	57.14	14.44
	21	4	248.52	39.63	0.00	n\a	n\a	n\a	n\a	41.12
		6	1328.48	34.23	0.00	n\a	n\a	n\a	n\a	44.48
		8	1696.79	38.47	0.00	n\a	n\a	n\a	n\a	33.67
	28	4	416.60	53.68	0.00	n\a	n\a	n\a	n\a	58.51
		6	953.01	56.95	0.00	n\a	n\a	n\a	n\a	69.58
		8	1234.02	45.63	0.00	n\a	n\a	n\a	n\a	58.63

Context 2

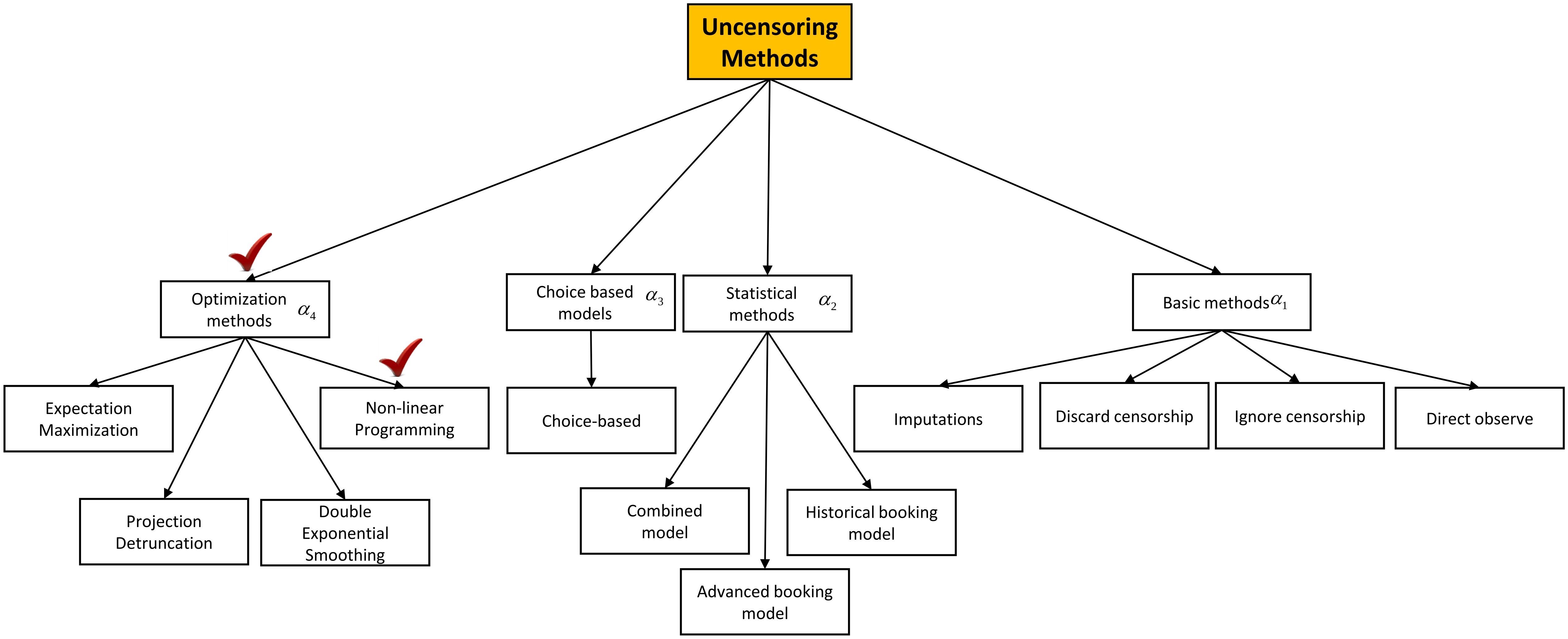
Internal and External Segmentation

Sh. Sharif Azadeh, M.Y. Maknoon, P. Marcotte. Working Paper. *Demand Modeling with Customer Segmentation Assumption (2014)*

Elements of Demand Forecasting in RMS



Methods of Uncensoring Demand



Internal vs. External Segmentation

<u>Internal Segmentation:</u>

Divides a market into distinct subsets (segments) that have similar needs (behavior). The reason behind is that customers behavior within each segment are fairly homogeneous; therefore, they are likely to respond similarly to a given marketing strategy (pricing, inventory control). Identifying these segments results in choosing appropriate marketing policies to simultaneously respond to segment's needs and to improve revenue.

- Product-based utility: u_i
  - Price, Brand loyalty, Quality-Comfort, Physical characteristics …

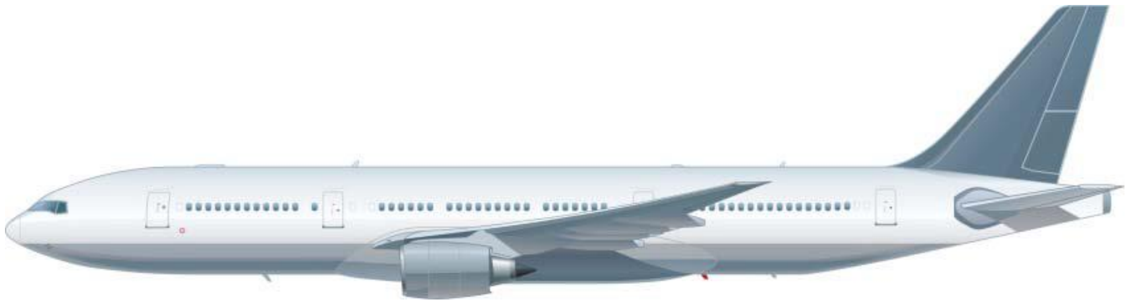
- Customer and Product-based utility: u_i^n
  - Demographic (age, marital status, …), Socioeconomic (income, education, religion, …), Life style.

<u>External Segmentation:</u>

The demand of products within a given nest on a given day belongs to one of the predefined number of clusters. The reason behind defining this type of segmentation is to capture seasonal effects based on daily demand flow for separate customer segments.

Censored Demand, Segmentation

Monday 8:00 a.m.



Product 1 Product 2 Product 3 Product 4 Product 5 Product 6

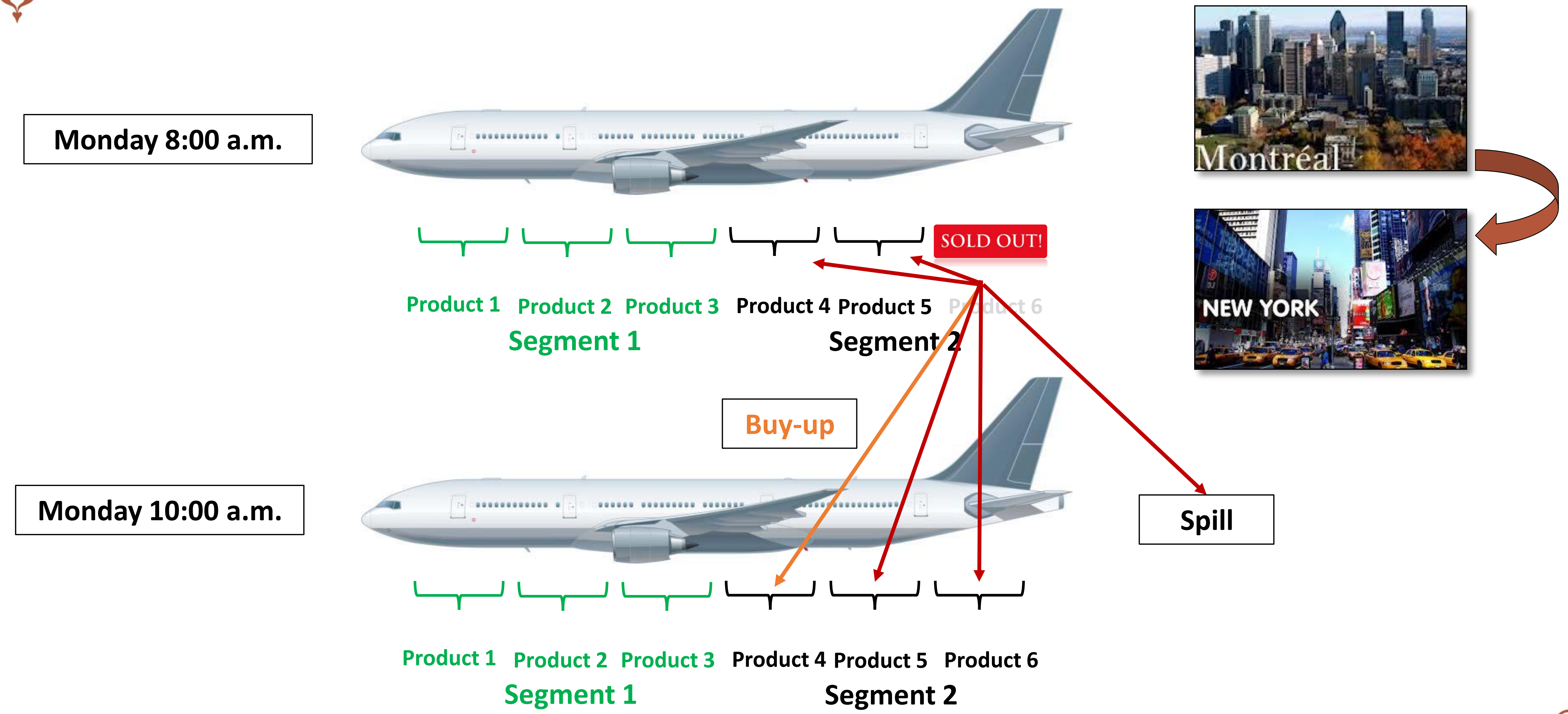
Segment 1 Segment 2

Monday 10:00 a.m.

Product 1 Product 2 Product 3 Product 4 Product 5 Product 6

Segment 1 Segment 2

Censored Demand, Segmentation



Objectives

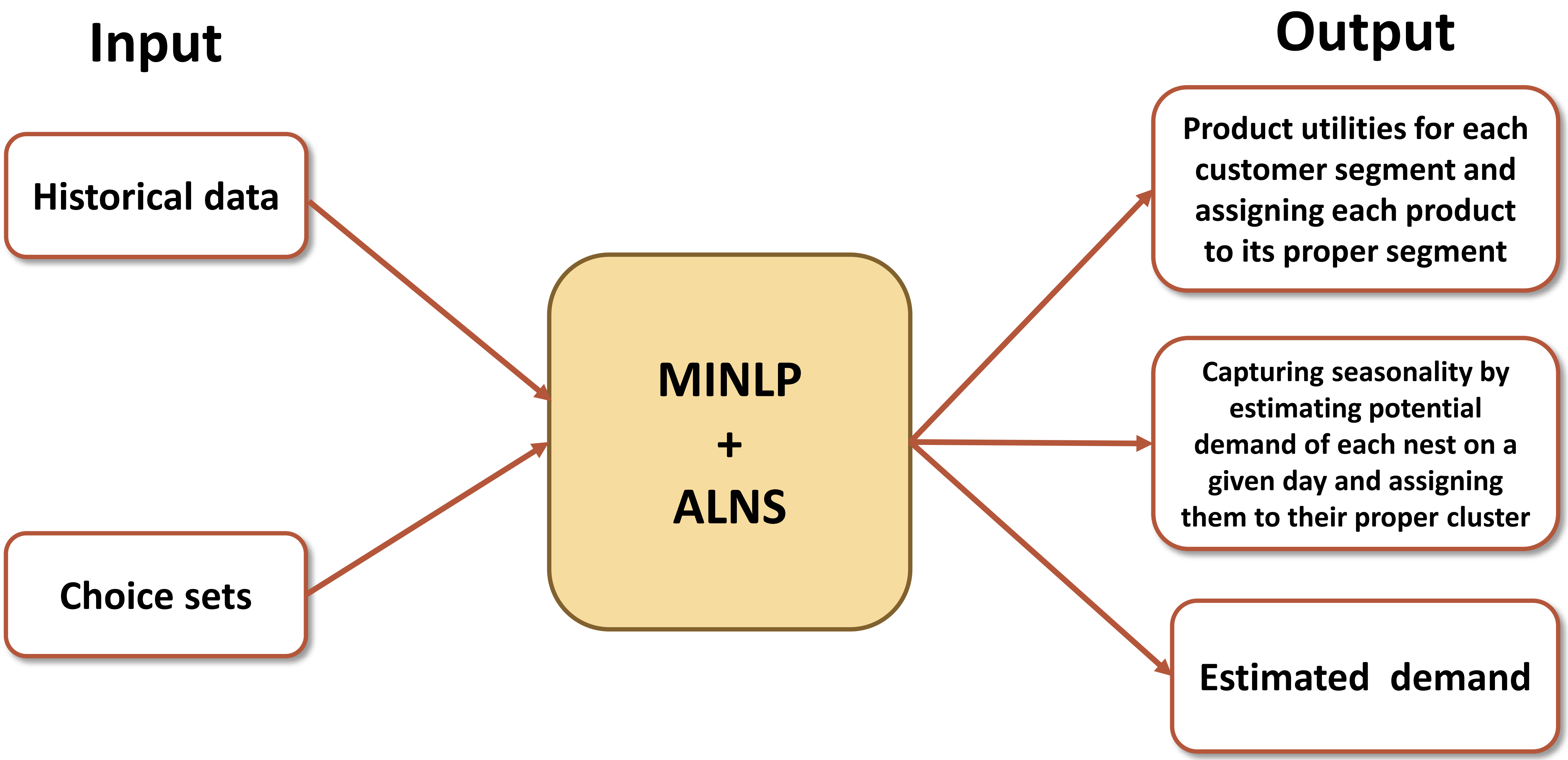
- **The main objectives:**

  1. To avoid primary demand and stochastic process assumption (Poisson) (Vulcano, van Ryzin)
  2. To estimate daily **potential demand for a given O-D and for a given customer segment**.
  3. To obtain **product utilities for each customer segment**.
  4. To capture seasonal effects.

- **The forecasting aspect:**

  By modifying the choice set, we can run the simulation again and observe how demand is expected to react to the modifications, for instance the closing of a booking class.

Objectives


Input
Output
Historical data
Choice sets
MINLP
+
ALNS
Product utilities for each customer segment and assigning each product to its proper segment
Capturing seasonality by estimating potential demand of each nest on a given day and assigning them to their proper cluster
Estimated demand

Notations

sets

Product	$i \in I = \{1, \ldots, \vert I\vert\}$
Day	$j \in J = \{1, \ldots, \vert J\vert\}$
Cluster	$c \in C = \{1, \ldots, \vert C\vert\}$
Nest	$n \in N = \{1, \ldots, \vert N\vert\}$

parameters

O_{ij}	observed bookings for product $i \in I$ on day $j \in J$
A_{ij}	availability status of product $i \in I$ on day $j \in J$

variables

d_j^n	potential demand of nest n on day j
p_{ij}	probability of selecting product $i \in I$ on day $j \in J$
z_{jc}^n	binary assignment variable to show if potential demand of nest $n \in N$ on day $j \in J$ belongs to cluster $c \in C$
u_i^n	utility of product $i \in I$ in nest $n \in N$
y_i^n	binary variable to assign product utility $i \in I$ to relevant nest $n \in N$
δ_c^n	potential demand of cluster c in nest n

Mathematical Model

$$\min_{\delta,u,z,d,y} \quad \sum_{j\in J}\sum_{i\in I}\left(p_{ij}d_j^n A_{ij} - O_{ij}\right)^2$$

$$\text{subject to} \quad p_{ij} = \sum_{n\in N} y_i^n \frac{\exp(u_i^n)}{\sum_{k\in I}\exp(u_k^n)A_{kj} + \exp(u_0^n)} \qquad i\in I, j\in J$$

$$\sum_{i\in I}\sum_{n\in N} y_i^n \frac{\exp(u_i^n)}{\sum_{k\in I}\exp(u_k^n)y_k^n A_{kj} + \exp(u_0^n)} + \sum_{n\in N}\frac{\exp(u_0^n)}{\sum_{k\in I}\exp(u_k^n)y_k^n A_{kj} + \exp(u_0^n)} = 1 \qquad j\in J$$

$$d_j^n = \sum_c \delta_c^n z_{jc}^n \qquad j\in J, n\in N$$

$$\sum_{c\in C} z_{jc}^n = 1 \qquad j\in J, n\in N$$

$$\sum_{n\in N} y_i^n = 1 \qquad i\in I$$

$$z_{jc}^n \in \{0,1\} \qquad j\in J, c\in C, n\in N.$$

- **Challenges of solving MINLP**

  Highly non-convex even its continuous relaxation. Global and nonlinear optimizers fail to precisely solve the problem. Heuristic method of ALNS is used to obtain assignment variables (z, y) and IPOPT is used to solve the nonlinear problem (u, δ).

ALNS-Algorithm

Output: Best Solution S^{Best}

$S^{Current}$ ←**Find** an initial solution

While Stopping Criterion is not met **do**

Randomly choose destroy and repair operators (H^{-}, H^{+})

S ←apply (H^{-}, H^{+}) on $S^{Current}$

if $F(S) \leq F(S^{Best})$

then $(S^{Best}, S^{Current}) \leftarrow S$

else

$S^{Current}$ ←**Return the accepted solution** $(S \quad , S^{Current})$

ALNS-Algorithm

Destroy Operators $(\boldsymbol{H}^{-})$:

1. Randomly remove y_i^n
2. Randomly remove z_{jc}^n (among available choices)
3. Remove worst y_i^n
   (column with the highest error)
4. Remove worst z_{jc}^n
   (element with the highest error)
5. History on nests y_i^n
6. History of z_{jc}^n

Insertion Operators $(\boldsymbol{H}^{+})$:

First y_i^n are repaired then z_{jc}^n

1. Greedy insertion
   Variable Selection: Random
2. Best insertion
   Variable Selection: Product with highest availability based on historical data
3. Regret insertion
   Variable Selection:
   Max Regret value:
   max $(MSE_{(minimum)} - MSE_{(minimum+1)})$

ALNS-Algorithm

Operator selection criteria: The probability of selecting operators varies according to their historical performance, we increase the weight of better performed operators.

Acceptance criteria: Simulated Annealing

Stopping criteria: θ<0.1 or # of iterations >10000

Initial solution: randomly assign y_i^n, then use regret insertion operator for z_{jc}^n

Initial temperature: initial solution, cooling rate (α) 0.99

Computational Results

Number of nests: 3

Class	Days	Alt	MSE	Time	Class	Nest	Simulation
2	7	8	8.36	986	0.0%	4.2%	49.31
		12	11.05	324	0.0%	0.0%	33.18
		16	9.78	962	0.0%	0.0%	20.96
	14	8	45.15	590	0.0%	0.0%	81.67
		12	42.14	1109	0.0%	2.8%	56.77
		16	24.31	364	0.0%	2.1%	79.52
	21	8	56.96	571	0.0%	0.0%	71.05
		12	47.66	975	0.0%	0.0%	60.44
		16	71.69	426	0.0%	4.2%	124.90
	28	8	71.49	302	0.0%	0.0%	64.76
		12	69.04	456	0.0%	0.0%	58.12
		16	40.09	867	0.0%	2.1%	54.63
3	7	8	16.55	1598	4.8%	0.0%	34.91
		12	7.32	1477	9.5%	0.0%	17.54
		16	5.97	1718	4.8%	2.1%	25.92
	14	8	25.42	1458	2.4%	0.0%	73.46
		12	22.12	663	0.0%	0.0%	50.90
		16	30.04	1616	2.4%	2.1%	27.31
	21	8	56.96	1433	0.0%	0.0%	55.02
		12	36.85	1643	0.0%	0.0%	50.56
		16	41.85	841	1.6%	0.0%	62.43
	28	8	111.04	1798	1.2%	0.0%	139.58
		12	53.97	1584	0.0%	0.0%	95.60
		16	43.88	1676	0.0%	2.1%	68.47
4	14	8	17.43	1330	1.8%	0.0%	35.42
		12	18.49	477	3.6%	0.0%	104.12
		16	17.80	1454	0.0%	6.3%	21.61
	21	8	58.19	605	0.0%	0.0%	61.53
		12	50.26	552	3.6%	2.8%	66.56
		16	56.48	1396	1.2%	4.2%	50.38
	28	8	78.82	988	0.0%	0.0%	87.55
		12	83.61	760	1.8%	2.8%	104.12
		16	66.99	855	0.9%	6.3%	87.73

Challenges and Perspectives

- Large scale problems with real data
- Buy-up, buy down probabilities (spill and recapture prediction)
- Ordered preference lists

Thank You