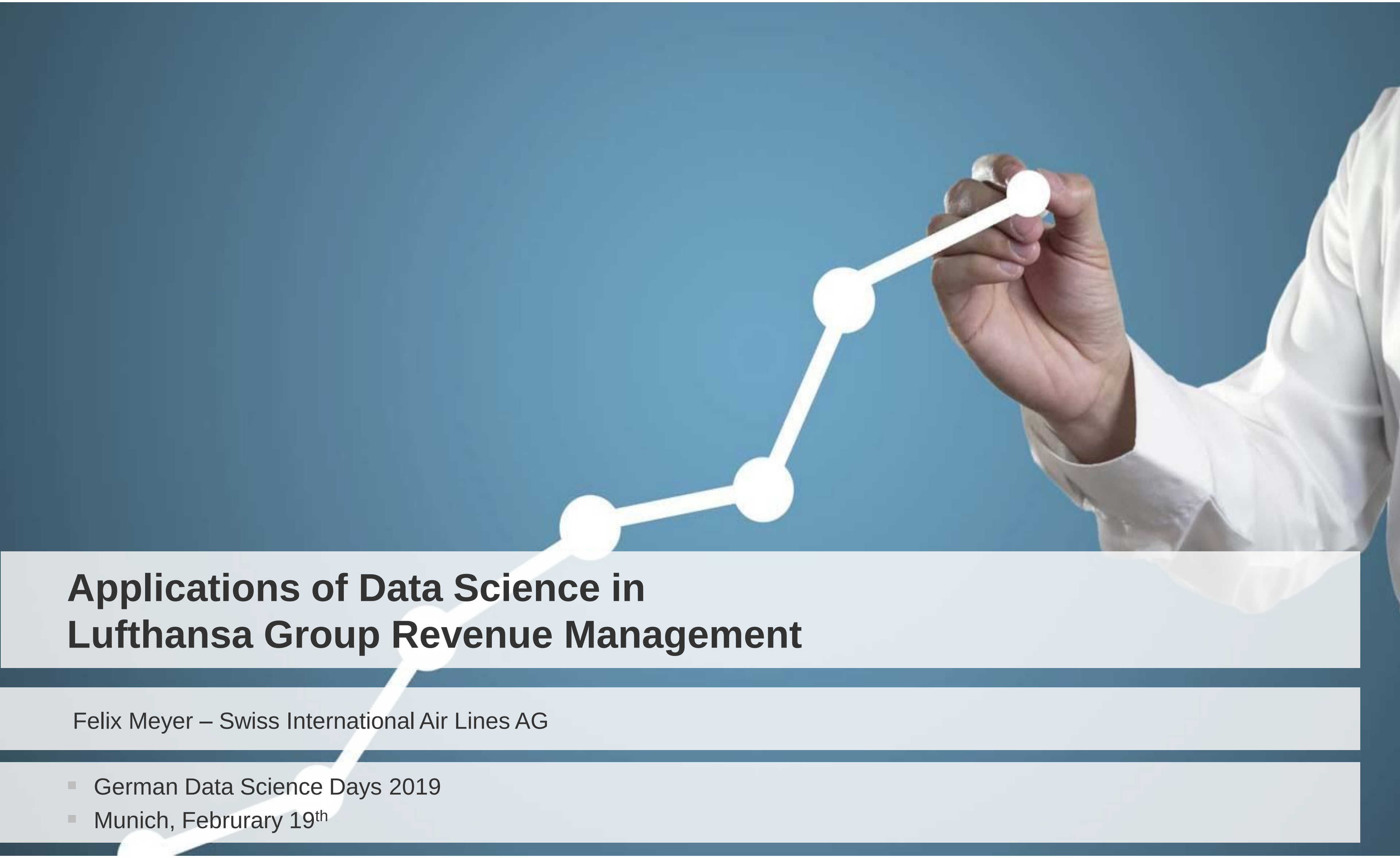


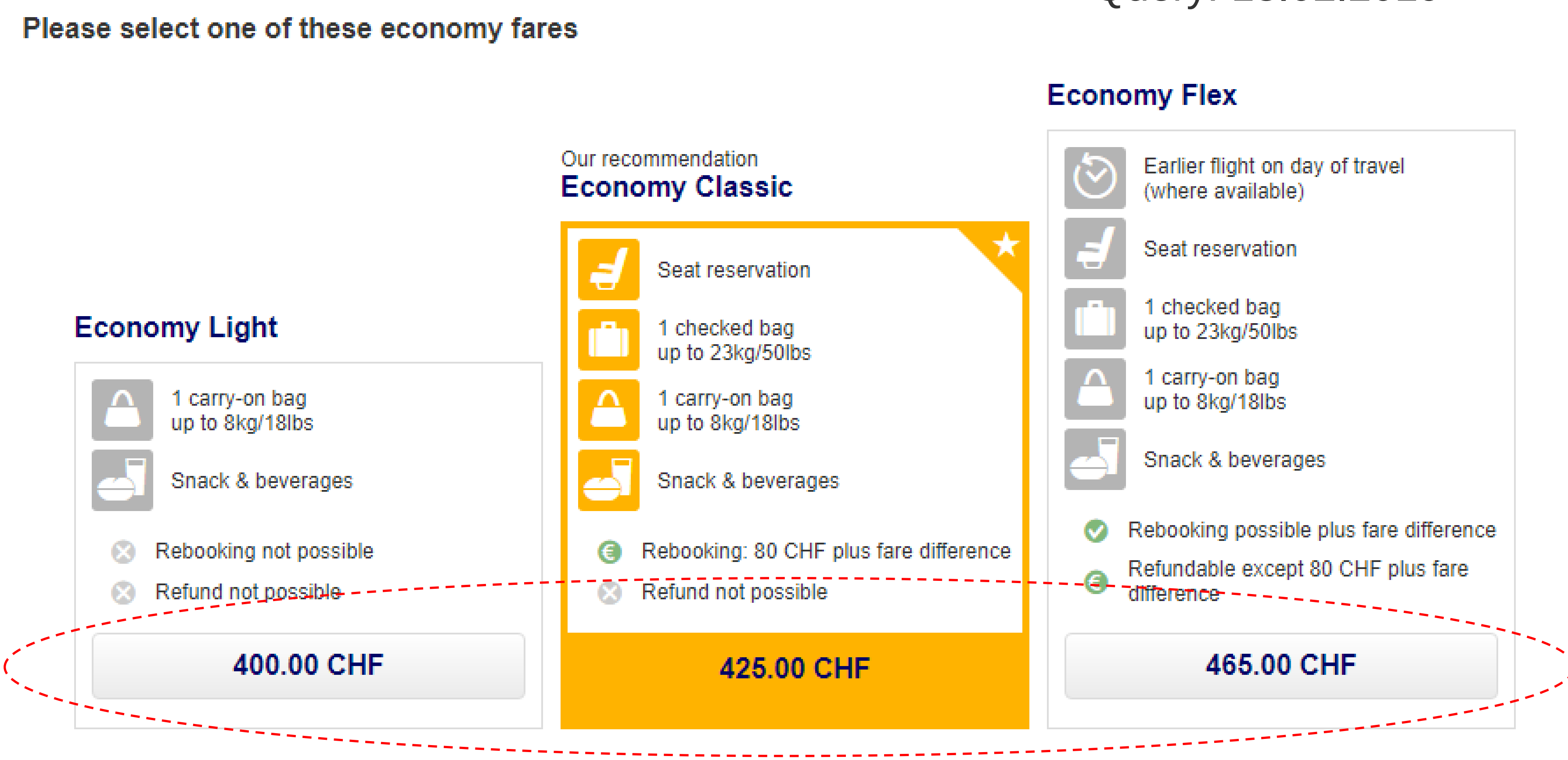
Applications of Data Science in Lufthansa Group Revenue Management

Felix Meyer – Swiss International Air Lines AG

- German Data Science Days 2019
- Munich, Februrary 19th

LHG`s pricing at their .coms

Query: 18.02.2019



LHG`s pricing at their .coms

Query: 18.01.2019

Please select one of these economy fares

Economy Light

- 1 carry-on bag up to 8kg/18lbs
- Snack & beverages
- Rebooking not possible
- Refund not possible

155.00 CHF

Our recommendation

Economy Classic

- Seat reservation
- 1 checked bag up to 23kg/50lbs
- 1 carry-on bag up to 8kg/18lbs
- Snack & beverages
- Rebooking: 80 CHF plus fare difference
- Refund not possible

180.00 CHF

Economy Flex

- Earlier flight on day of travel (where available)
- Seat reservation
- 1 checked bag up to 23kg/50lbs
- 1 carry-on bag up to 8kg/18lbs
- Snack & beverages
- Rebooking possible plus fare difference
- Refundable except 80 CHF plus fare difference

220.00 CHF

Why do the prices change?

Query: 18.12.2018

Please select one of these economy fares

Economy Light

- 1 carry-on bag up to 8kg/18lbs
- Snack & beverages
- Rebooking not possible
- Refund not possible

325.00 CHF

Our recommendation

Economy Classic

- Seat reservation
- 1 checked bag up to 23kg/50lbs
- 1 carry-on bag up to 8kg/18lbs
- Snack & beverages
- Rebooking: 80 CHF plus fare difference
- Refund not possible

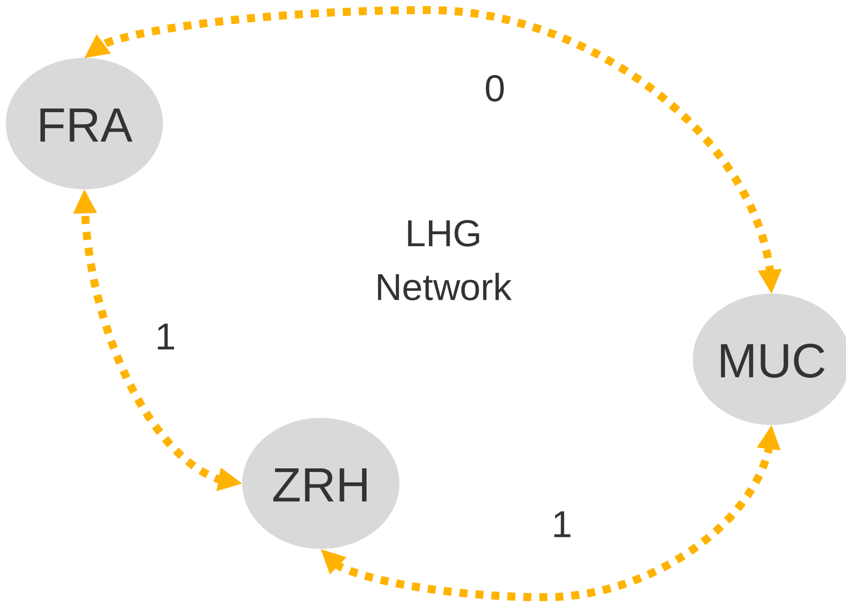
350.00 CHF

Economy Flex

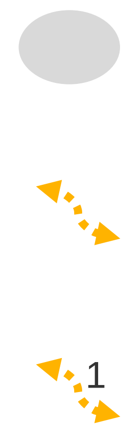
- Earlier flight on day of travel (where available)
- Seat reservation
- 1 checked bag up to 23kg/50lbs
- 1 carry-on bag up to 8kg/18lbs
- Snack & beverages
- Rebooking possible plus fare difference
- Refundable except 80 CHF plus fare difference

390.00 CHF

Why do the prices change?
The motivation of the airline revenue management problem.



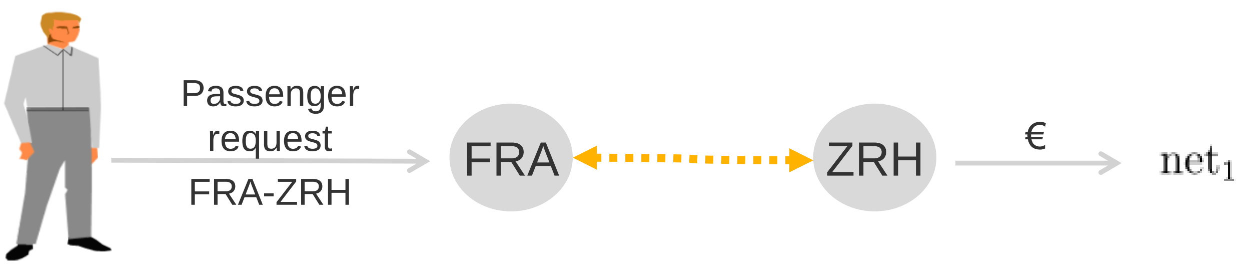
Legend:



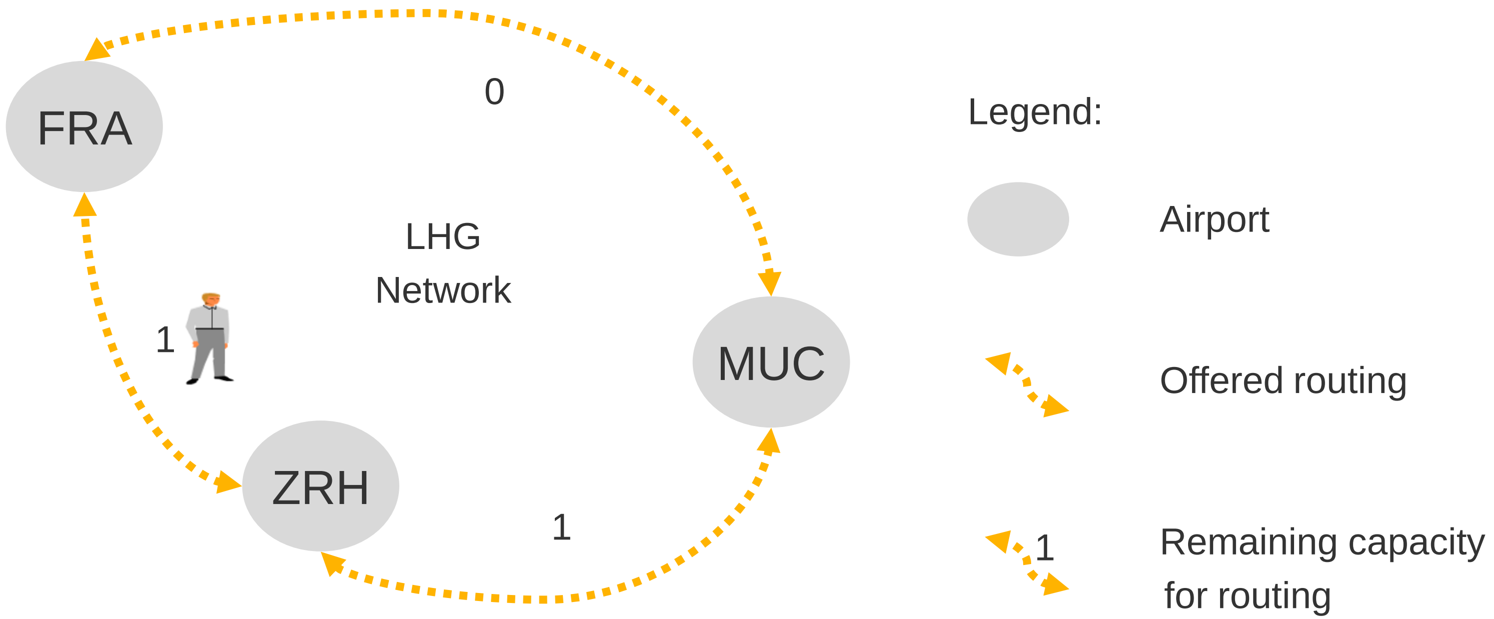
Airport

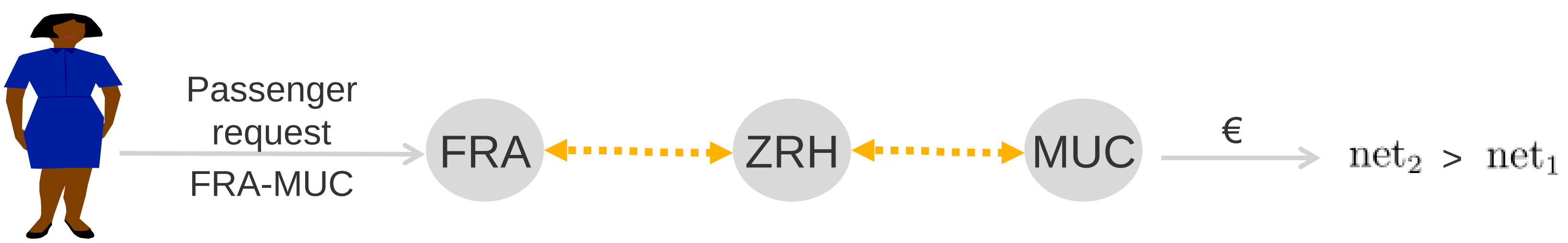
Offered routing

Remaining capacity for routing

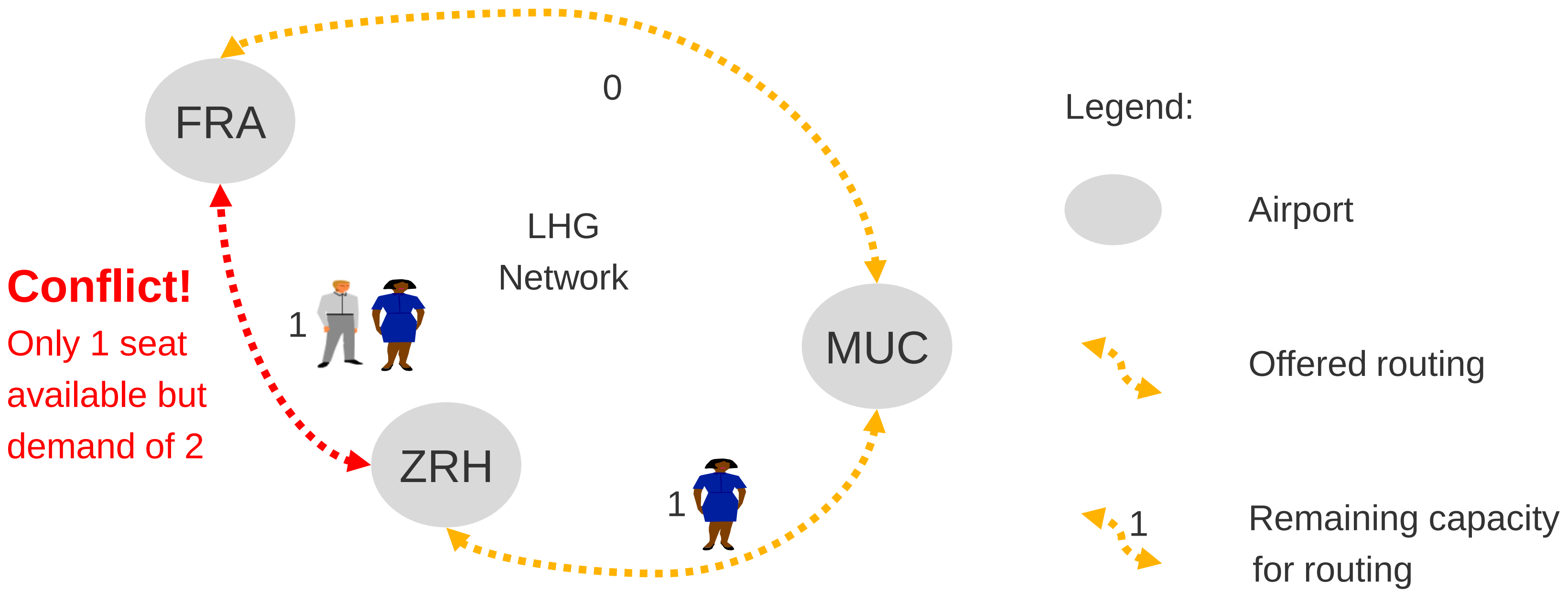


Each passenger consumes capacity for every routing that is used to build the requested journey.





To optimize revenue in situations where the capacity is scarce, the airline decides which passenger to accept/reject.

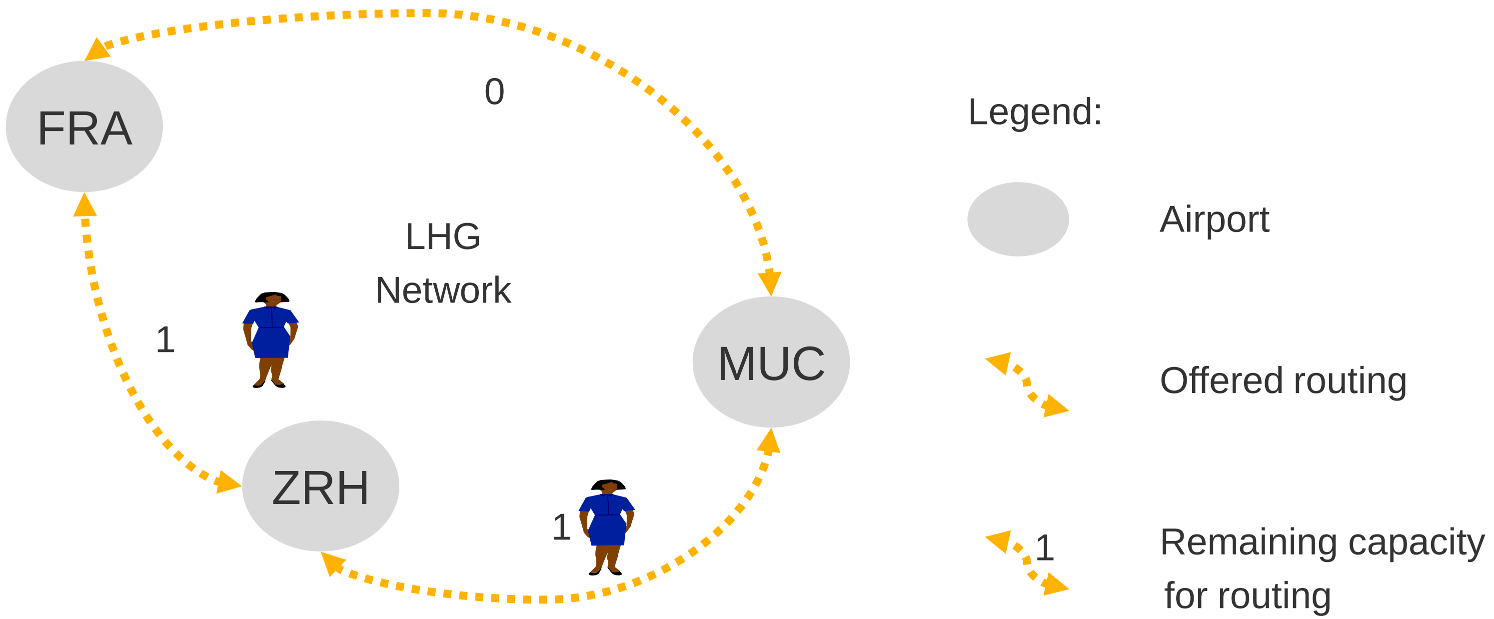


Goal: only accept that request which maximizes the revenue contribution to the network**:**
Introduction of opportunity costs (bid-price) for each segment (FRA-ZRH, ZRH-MUC, FRA-MUC).

$\text{net}_1 < \text{bid-price}_{\text{FRA-ZRH}}$ → Rejected

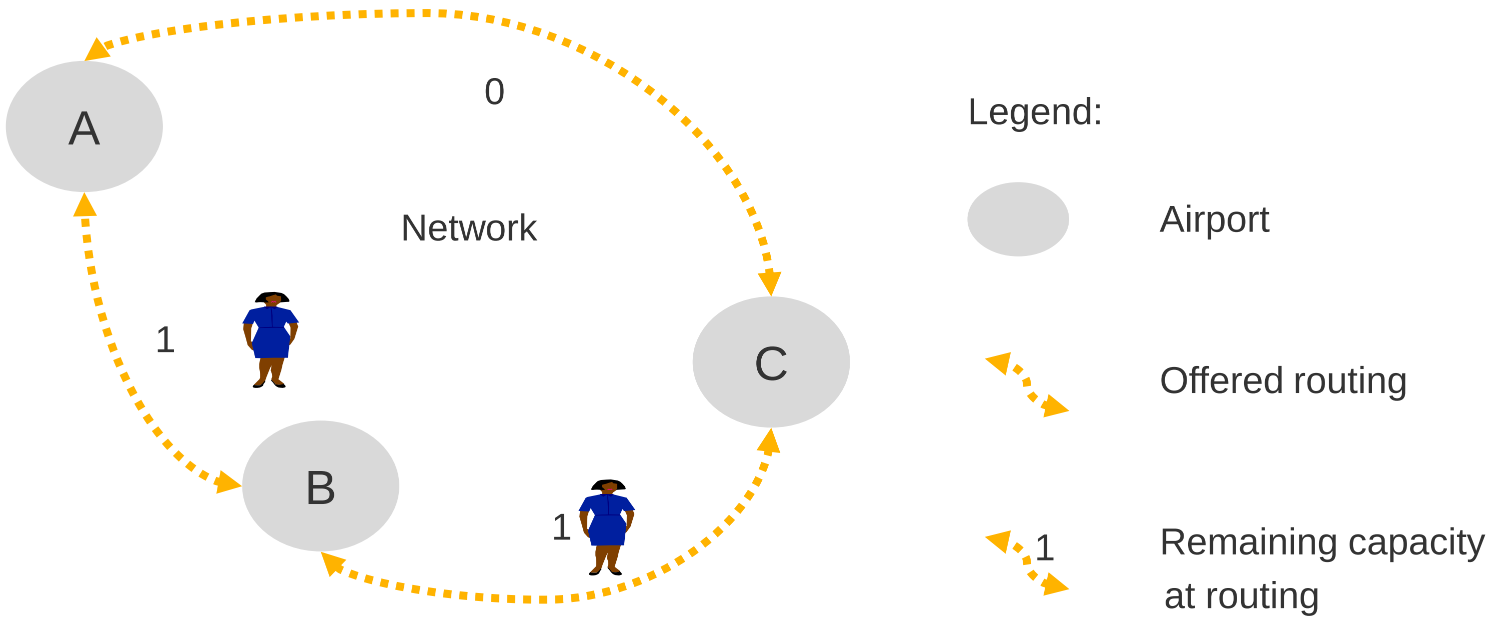
$\text{net}_2 > \text{bid-price}_{\text{FRA-ZRH}} + \text{bid-price}_{\text{ZRH-MUC}}$ → Accepted

Besides the evaluation of opportunity costs, the willingness to pay has to be evaluated to define the (revenue) optimal price.



- Suppose the airline offers two products (classes) with $\text{net}_{\text{class}_1} < \text{net}_{\text{class}_2}$.
- Let the willingness to pay of  be equal to $\text{net}_{\text{class}_2}$.
- The airline accepts if $\text{net}_2 > \text{bid-price}_{\text{FRA-ZRH}} + \text{bid-price}_{\text{ZRH-MUC}}$.
- Accepting  within class 1 results in (price-elasticity) costs of $\text{net}_{\text{class}_2} - \text{net}_{\text{class}_1}$.

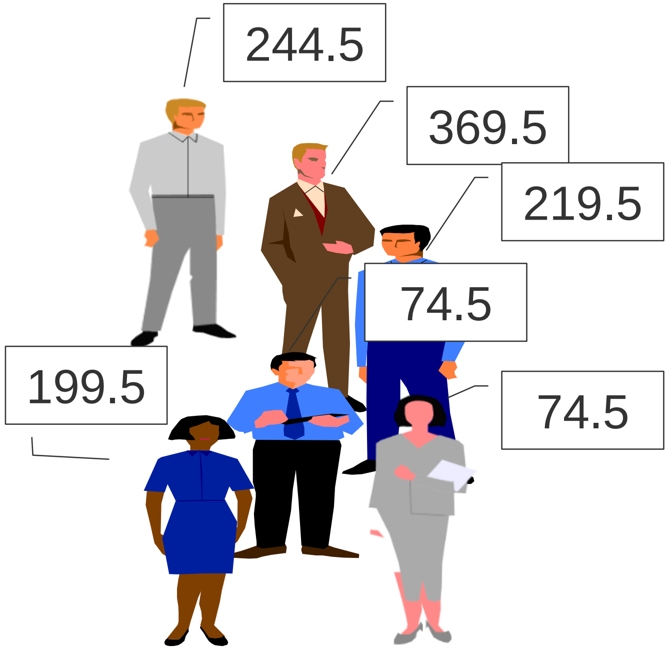
For each request, the optimal price is set to be as close as possible to the customer`s willingness to pay.

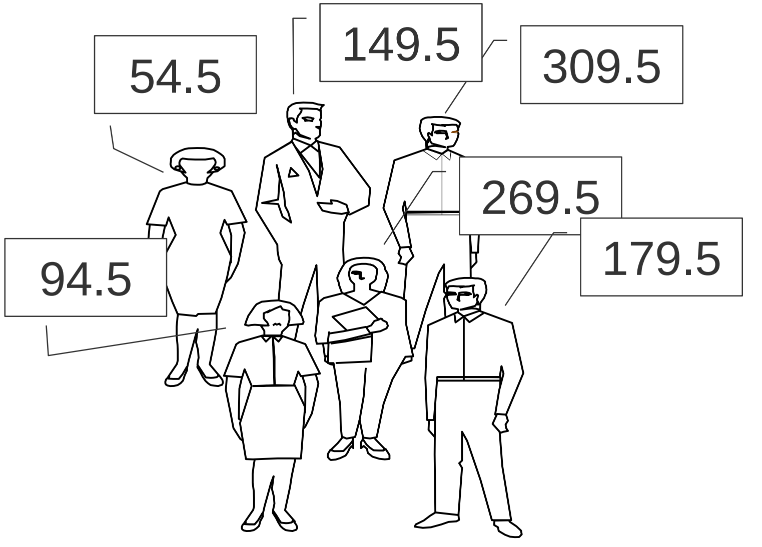


To minimise price-elasticity cost, i.e., minimize the risk of  buying down into class 1, **the price-sensitivity of demand needs to be evaluated (the topic of this presentation).**

Focus topic: what are price-elasticity costs and how does LHG make use of them. A small experiment

Suppose we offer the same product, such as a seat on a plane, under the same conditions, repeatedly for a different price to N = 12 people.





Bookings = 6 x 1 = 6
accepted as
willingness to pay >= price

Non-Bookings = 6 x 0 = 0
rejected as
willingness to pay < price

Using the collected data we are able to device (optimal) pricing strategies (for simplicity assume price = revenue, i.e., no cost)

k	$PRICE_k$	$\sum_{i=1}^{k}$ bookings$_i$	bookings$_k$
11	54.5	12	1
10	74.5	11	2
9	94.5	9	1
8	149.5	8	1
7	179.5	7	1
6	199.5	6	1
5	219.5	5	1
4	244.5	4	1
3	269.5	3	1
2	309.5	2	1
1	369.5	1	1

Option 1 (maximize bookings): 54.5
Every passenger buys (down) for 54.5.

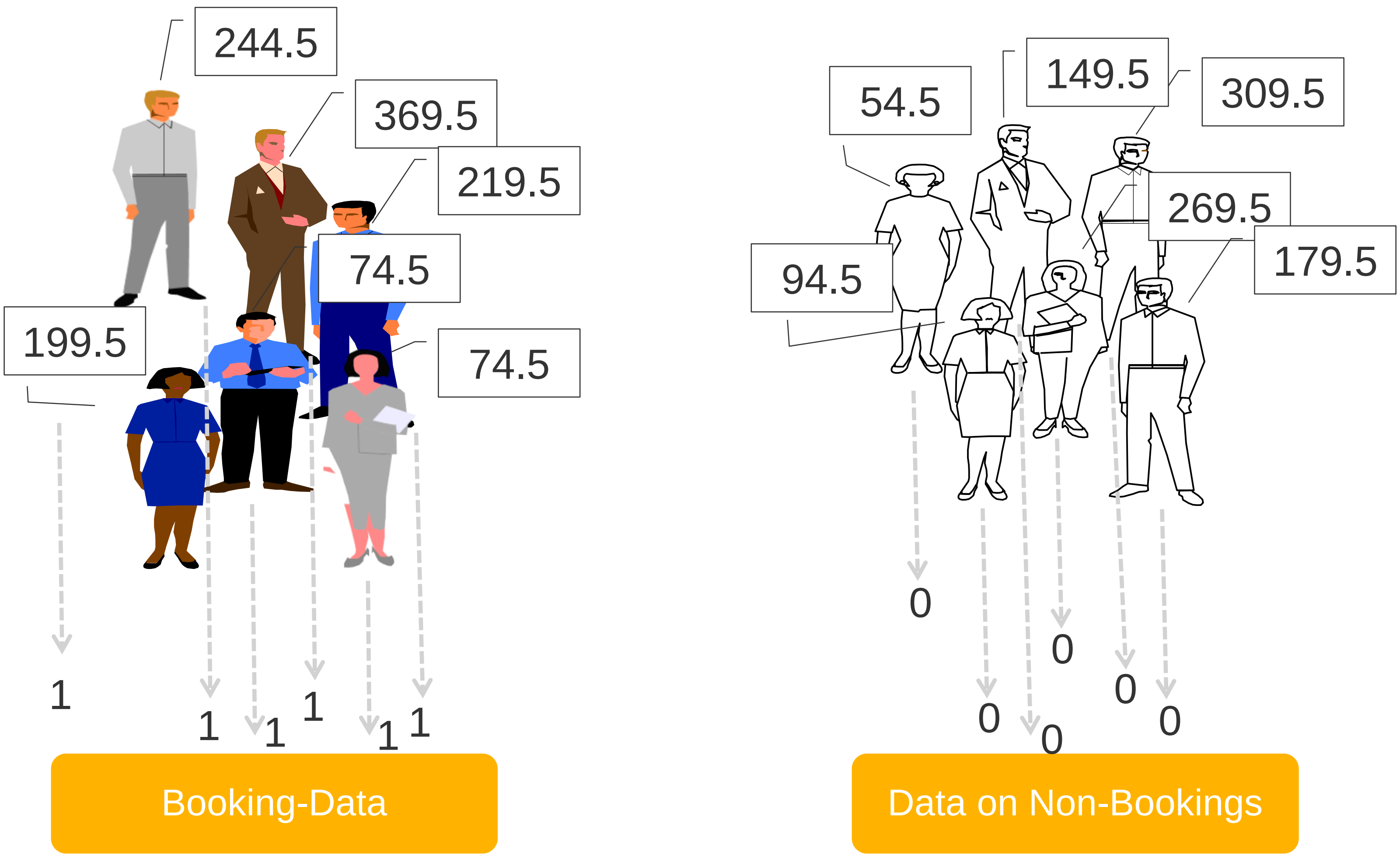
Option 2 (between): 149.5.
Every passenger willing to pay more then 149.5 buys down. Demand below 149.5 is lost.

Option 3 (maximizing revenue): 179.5

Option	k	$PRICE_k$	y_k	revenue
1	12	54.5	12	654.0
2	8	149.5	8	1196.0
3	7	179.5	7	1256.5

This approach is not feasible in practice as N as and the number of non-bookings for each price is generally unknown.

In practice, LHG observes bookings (=1) and collects non-bookings (=0) for days (single snapshot each day) when nothing is sold



From experiment to practice: suppose one repetition of the experiment is the observed booking process for a flight i = 1,…,M.

A booking (=1) with arrival t_{i1} days before departure

Non-bookings (=0) are recognised for flight *i* at days *t* if no booking was observed.

$t_{i1}\ t_{i2}$

$t = t_i^{\text{open}}$

(first day of sale of flight i)

$t = t_i^{\text{close}}$

(last day of sale of flight *i*)

Let $N_i(t)$ denote the cumulated number of bookings for flight *i* at time *t* with $N_i(t_i^{\text{open}}) \equiv 0$.

This booking process is observed several times with different information on the booking- and flight-level

On each flight $i \in I_{Flight}$, we observe the number of cumulated bookings at a time to departure $t = \tilde{t}$ as $N_i(\tilde{t}) \in \mathbb{N}_0$

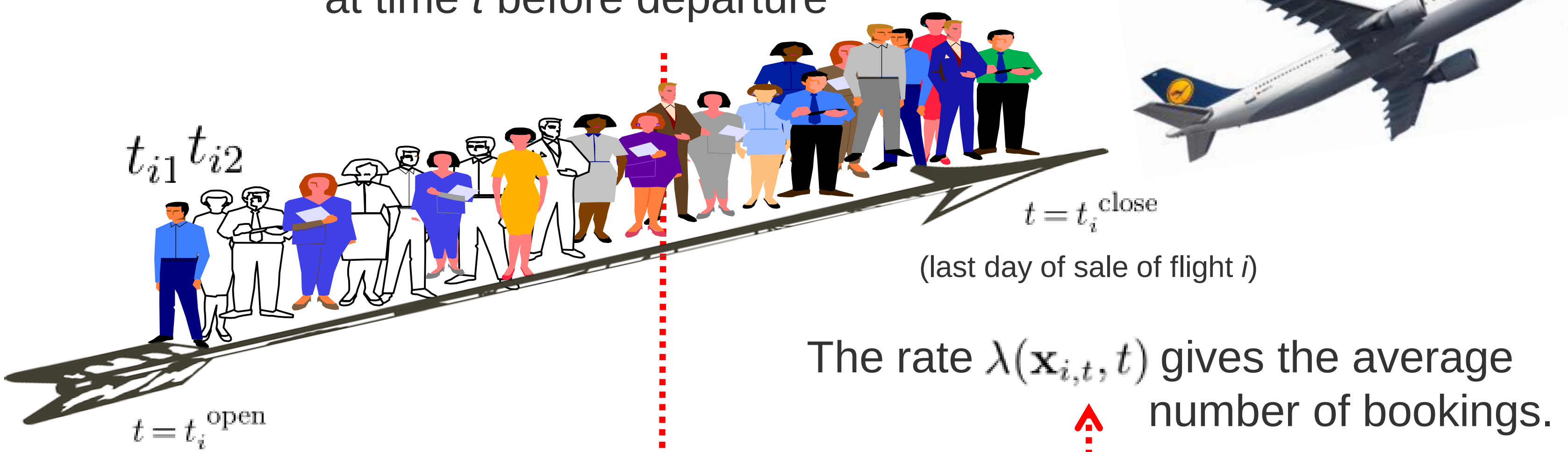
$N_1(\tilde{t})$ $N_2(\tilde{t})$ $N_3(\tilde{t})$ . . . $N_n(\tilde{t})$

LHGs statistical model assumes that each increment, i.e., the number of bookings during day t, is Poisson distributed.

For flight $i \in I_{Flight}$ we look at the incremental booking process

$$N_i(t) - N_i(t-1)$$

at time *t* before departure



(last day of sale of flight *i*)

(first day of sale of flight i)

The rate $\lambda(\mathbf{x}_{i,t}, t)$ gives the average number of bookings.

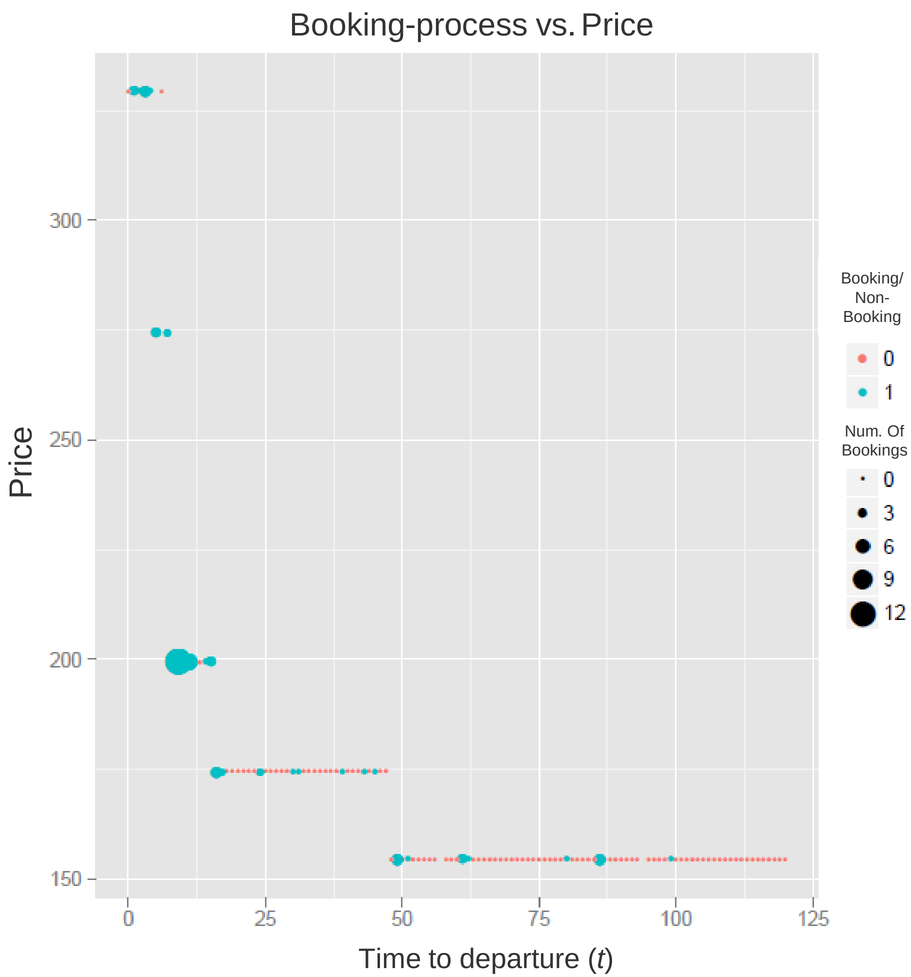
$$y_{it} = N_i(t) - N_i(t-1) \sim \text{Poisson}\left(\lambda(\mathbf{x}_{i,t}, t) \right)$$

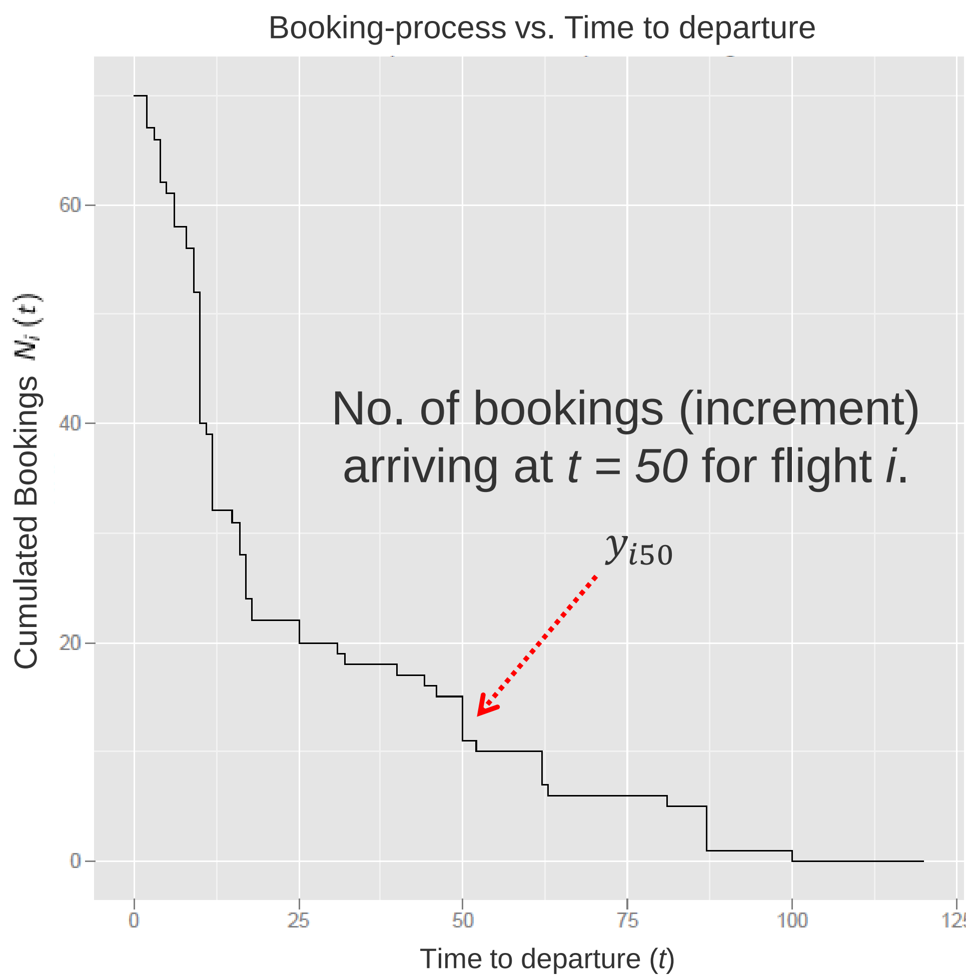
No. of bookings for flight *i* at time *t*.

$$y_{it} \in \{0, 1, 2, \ldots\} \text{ as } N_i(t) \geq N_i(t-1)$$

The rate is influenced by time-depending covariates, such as price, subsumed in $x_i(t) \in \mathbb{R}^p$.

An example of how the booking-process looks like (real data) for one particular flight (flight number, departure date, routing)





The Poisson intensity *λ(t)* accounts for changes in booking intensity and depends on price and additional covariates.

- Covariates for flight *i* at time *t* with price = $\mathrm{PRICE}_{i,t}$ are given by a covariate vector $\mathbf{x}_{i,t} = (x_{1,i,t}, \ldots, x_{K_2,i,t}, \mathrm{PRICE}_{i,t}, z_{1,i,t}, \ldots, z_{K_1,i,t}, t)$.
- Index-sets $I_1 = \{1, \ldots, K_1\}$ and $I_2 = \{1, \ldots, K_2\}$ give the positions of categorical $(x_{k,i,t}, k \in I_1)$ and continuous $(z_{k,i,t}, k \in I_2)$ covariates.
- For covariates that belong to I_1, the *k-th* categorical covariate takes values from the set $J_k = \{1, \ldots, G_k\}$.

This leads to the model:

$$\lambda(\mathbf{x}_{i,t}, t) = \lambda\left(x_{1,i,t}, \ldots, x_{K_2,i,t}, \mathrm{PRICE}_{i,t}, z_{1,i,t}, \ldots, z_{K_1,i,t}, t\right)$$

Quantifying the price effect on the booking intensity given the factors $(x_{k,i,t}, k \in I_1)$ and $(z_{k,i,t}, k \in I_2)$.

$$\lambda(\mathbf{x}_{i,t}, t) = \lambda\left(x_{1,i,t}, \ldots, x_{K_2,i,t}, \mathrm{PRICE}_{i,t}, z_{1,i,t}, \ldots, z_{K_1,i,t}, t\right)$$

To specify how the covariates influence the booking intensity, a model that captures all interaction effects of the continuous covariates is set:

$$\log\Big(\lambda(\mathbf{x}_{i,t}, t)\Big) = \beta_0 + \sum_{k \in I_1} \mathbf{1}_{\{x_{k,i,t}=j\}} \beta_{k,j}$$

$$+ f_p\left(\mathrm{PRICE}_{i,t}\right) + f_{p,t}\left(\mathrm{PRICE}_{i,t}, t\right) + \sum_{k \in I_2} f_{p,k}\left(\mathrm{PRICE}_{i,t}, z_{k,i,t}\right)$$

$$+ f_t(t) + \sum_{k \in I_2} f_k\left(z_{k,i,t}\right) + \sum_{k \in I_2} f_{t,k}\left(t, z_{k,i,t}\right) + \sum_{\substack{k_1 < k_2 \\ k_1, k_2 \in I_2}} f_{k_1,k_2}\left(z_{k_1,i,t}, z_{k_2,i,t}\right)$$

The booking intensity is modelled in a full factorial design, i.e., all main effects f(.), as well as all interaction effects f(.,.), are captured.

Classification of the model components:

- Describing the volume of demand $f_t(.), f_{k \in I_2}(.), f_{t,k \in I_2}(.,.), f_{\substack{k_1 < k_2 \\ k_1, k_2 \in I_2}}(.,.)$
- Influenced the slope of PRICE representing price-sensitivity
  $f_p(.), f_{p,t}(.,.)$, $f_{p,k \in I_2}(.,.)$ (for these functions we impose monotonicity within PRICE)

Example:

- $f_p(\mathrm{PRICE}_{i,t})$ determines the general level of price-sensitivity,
- $f_t(t)$ describes the general booking intensity along t,
- $f_{p,t}(\mathrm{PRICE}_{i,t}, t)$ changes the price-sensitivity within t.

Optimal (continuous) pricing: to optimize the price, the marginal revenue over opportunity cost π is maximized.

$$\max_{\text{PRICE}} \left\{ \lambda(\boldsymbol{x}_{i,t}, t) \times \text{NET} - \lambda(\boldsymbol{x}_{i,t}, t) \times \pi \right\}$$

$$\Leftrightarrow \frac{\partial \Big(\lambda(\boldsymbol{x}_{i,t}, t) \times \text{NET} - \lambda(\boldsymbol{x}_{i,t}, t) \times \pi \Big)}{\partial \text{PRICE}} \stackrel{!}{=} 0$$

where

- the total revenue gain is defined by $\lambda(\boldsymbol{x}_{i,t}, t) \times \text{NET}$
- the total opportunity costs of capacity are $\lambda(\boldsymbol{x}_{i,t}, t) \times \pi$
  (is zero if capacity is not a constraint/scarce)

To calculate the derivative of $\lambda(\boldsymbol{x}_{i,t}, t)$ with respect to PRICE, we use the fact that the derivative of a B-Spline is a linear combination of lower order B-Splines.

Mapping from optimal NET- to optimal PRICE-values.

$$\max_{\text{PRICE}} \left\{ \lambda(\boldsymbol{x}_{i,t}, t) \times \text{NET} - \lambda(\boldsymbol{x}_{i,t}, t) \times \pi \right\}$$

$$\Leftrightarrow \frac{\partial \Big(\lambda(\boldsymbol{x}_{i,t}, t) \times \text{NET} - \lambda(\boldsymbol{x}_{i,t}, t) \times \pi \Big)}{\partial \text{PRICE}} \stackrel{!}{=} 0$$

Solving the maximization problem gives the optimal NET-value, which is mapped to the optimal PRICE-value by:

$$\text{PRICE} - \text{NET} = \alpha_0 + \alpha_1 \times \text{PRICE}$$

$$\Leftrightarrow \text{DIFF} = \alpha_0 + \alpha_1 \times \text{PRICE}$$

The difference between NET and PRICE is described by a fix-amount α_0 and a variable factor α_1 (VAT) describing how DIFF depends upon PRICE (for non-domestic flights there is no VAT).

The optimal closed form solution results as the sum of marginal revenue and marginal costs.

$$\max_{\text{PRICE}} \left\{ \lambda(\boldsymbol{x}_{i,t}, t) \times \text{NET} - \lambda(\boldsymbol{x}_{i,t}, t) \times \pi \right\}$$

$$\Leftrightarrow \frac{\partial\Big(\lambda(\boldsymbol{x}_{i,t}, t) \times \text{NET} - \lambda(\boldsymbol{x}_{i,t}, t) \times \pi\Big)}{\partial \text{PRICE}} \overset{!}{=} 0$$

If $\lambda(\boldsymbol{x}_{i,t}, t)$ is taken to be linear in PRICE the maximization-problem has the closed-form solution:

$$\text{PRICE}_{\text{optimal}} = \underbrace{-\frac{1}{f'_p + f'_{p,t} + \sum_{k \in I_2} f'_{p,k}}}_{\text{marginal revenue}} + \underbrace{\frac{\alpha_0}{1-\alpha_1} + \frac{\pi}{1-\alpha_1}}_{\text{marginal costs}}$$

where $f'_p, f'_{p,t}$ and $f'_{p,k}, k \in I_2$ correspond to the first derivative with respect to PRICE.

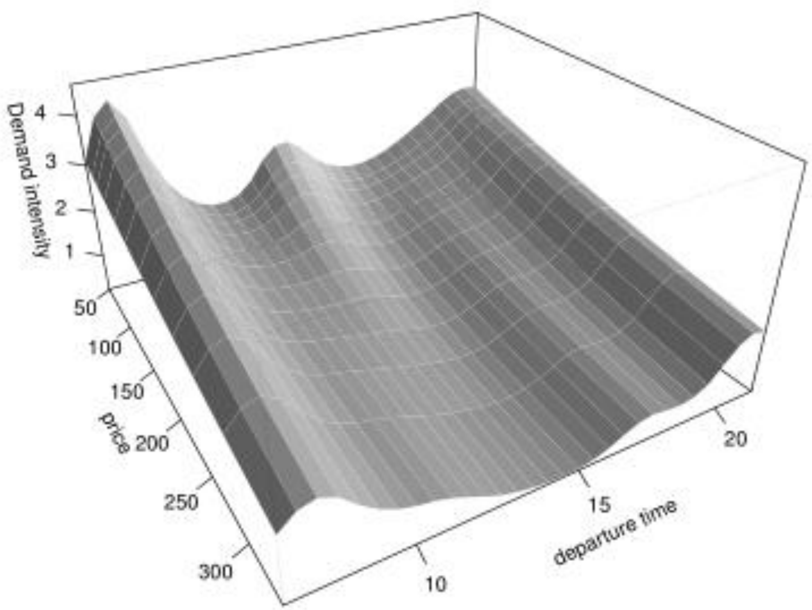
Results: conditional demand estimates.

where:

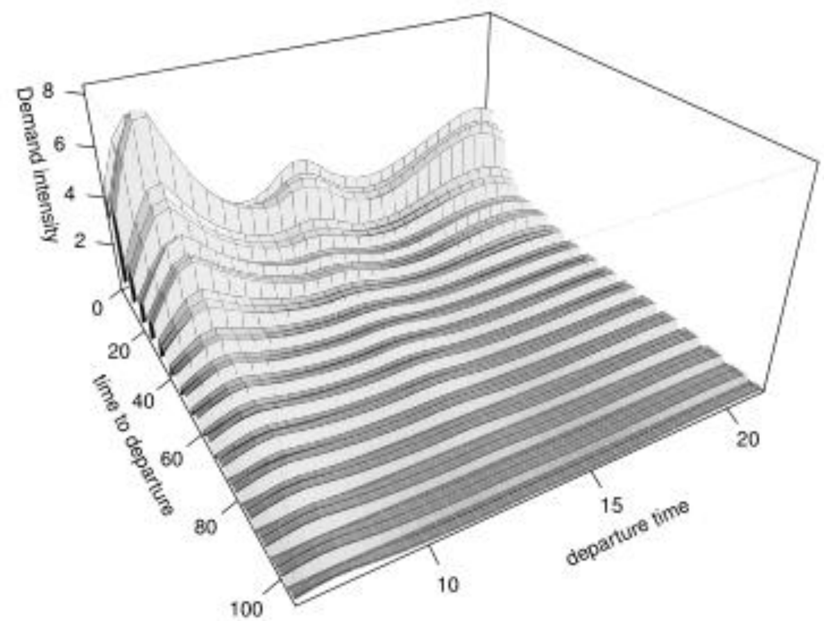
- *YDAY* gives the day of the year, taking values from 1,….,365,
- *DTIME* is the departure time (local) of a flight,
- *BDAY* is the booking day of the week, taking values Monday,…,Sunday,
- *t* indicates the number of days before departure.

Conditional estimates of smooth effects for:

(a) t = 0 & YDAY = 7

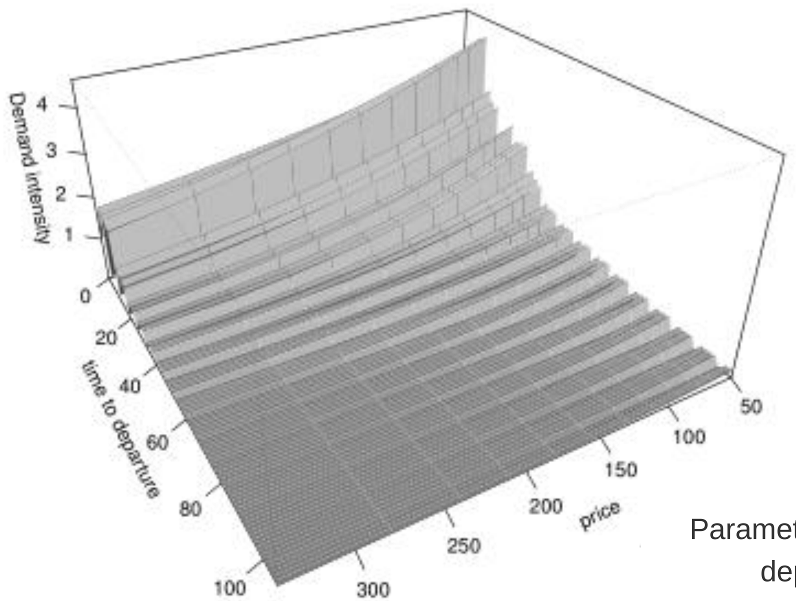


(b) PRICE = 59,5 & YDAY = 7

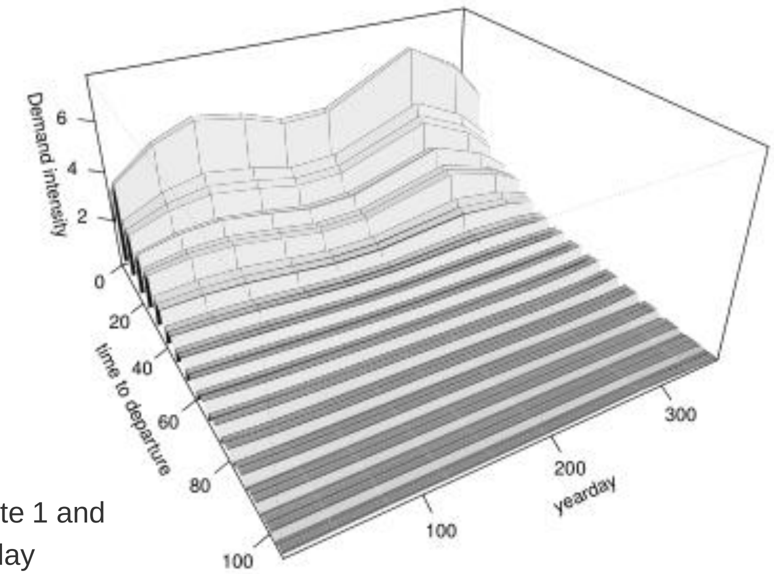


Results: conditional demand estimates.

(c) DTIME = 9:50 am & YDAY = 7



(d) PRICE = 59,5 & DTIME = 9:50 am

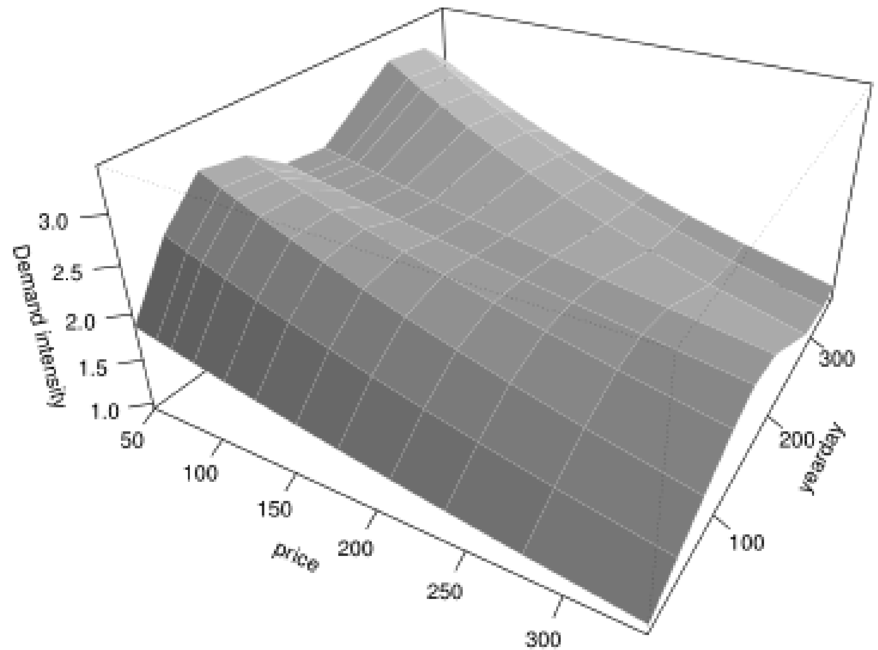


Parameter estimates for route 1 and departure day = Tuesday

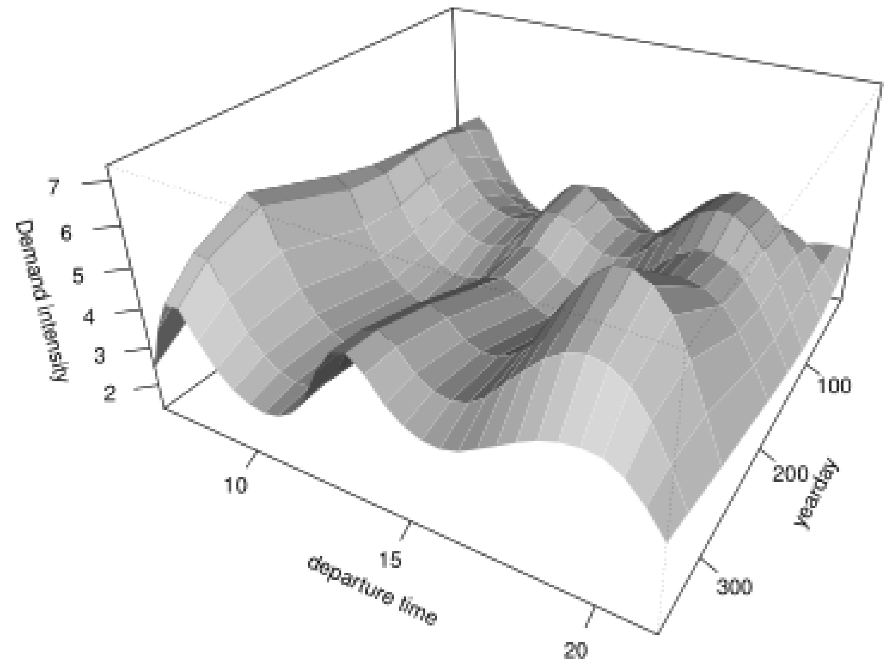
Parameter	Estimate	Std. Error
Intercept	-2.0880	0.1867
BDAY = Monday	1.7700	0.0333
BDAY = Tuesday	1.6740	0.0334
BDAY = Wednesday	1.8282	0.0335
BDAY = Thursday	1.7695	0.0340
BDAY = Friday	1.8981	0.0331
BDAY = Saturday	-0.1878	0.0460

Results: conditional demand estimates.

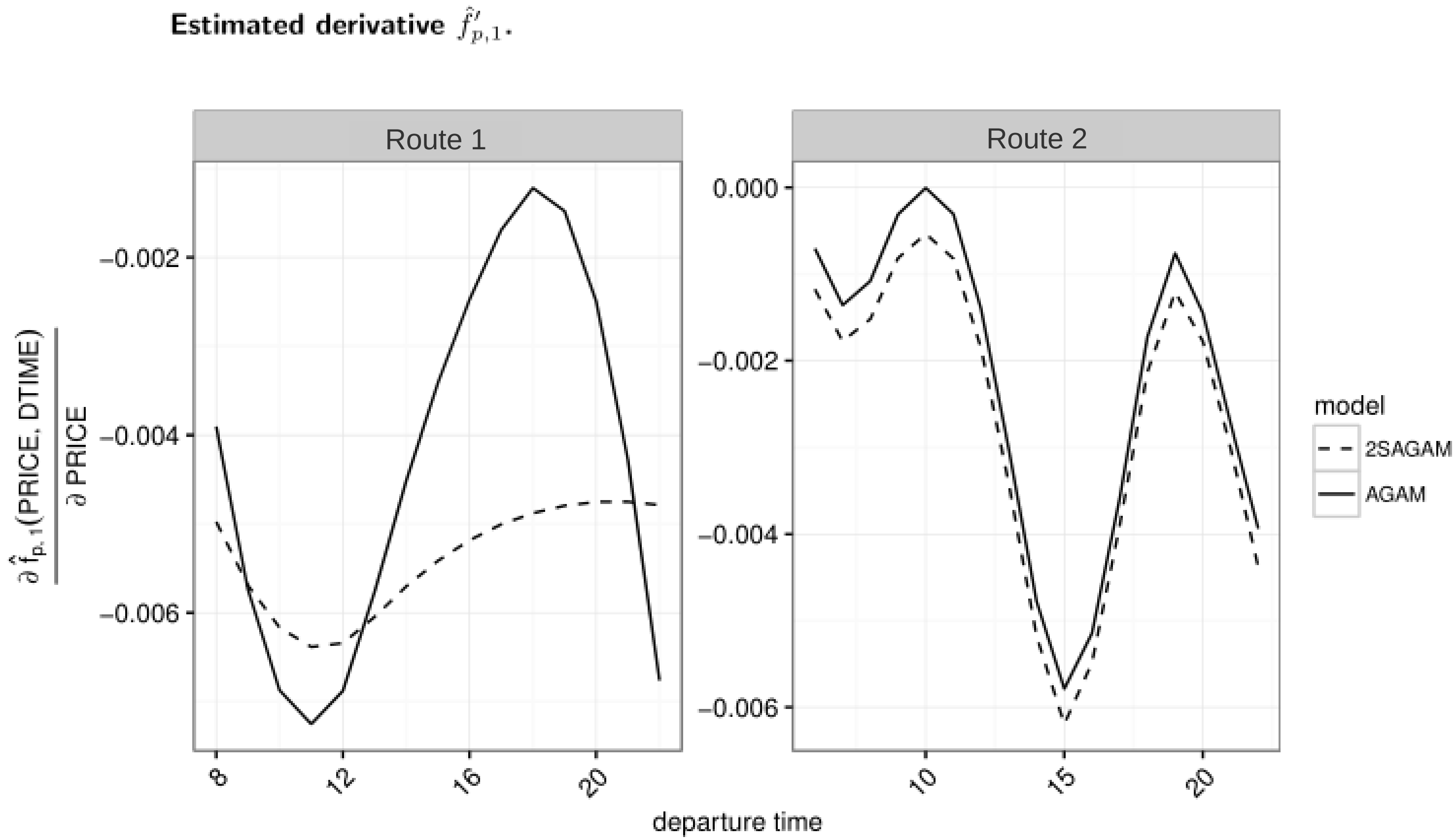
(e) $t = 0$ & DTIME = 9:50 am



(f) PRICE = 59,5 & $t = 0$

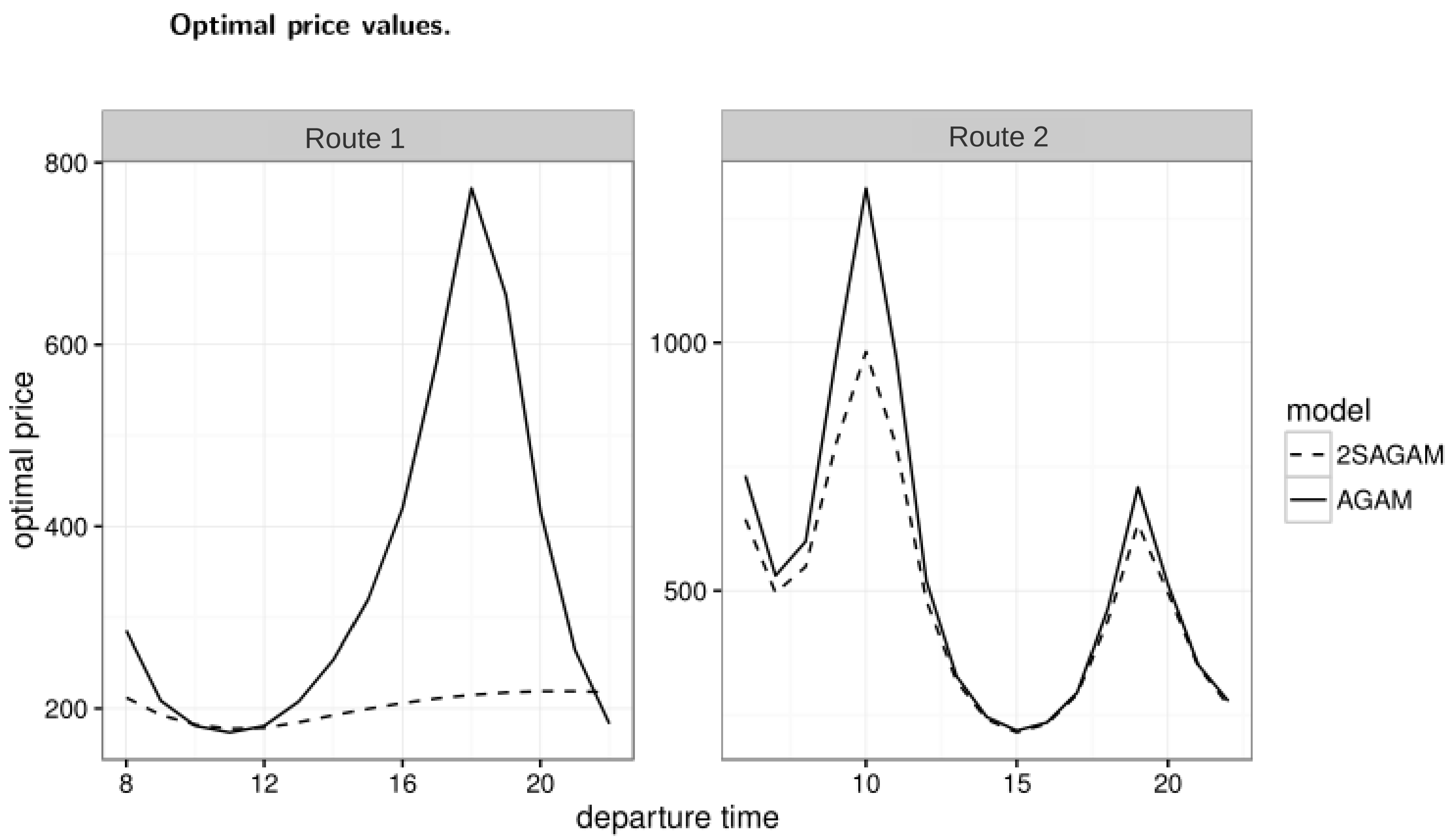


Results: estimated price derivatives.

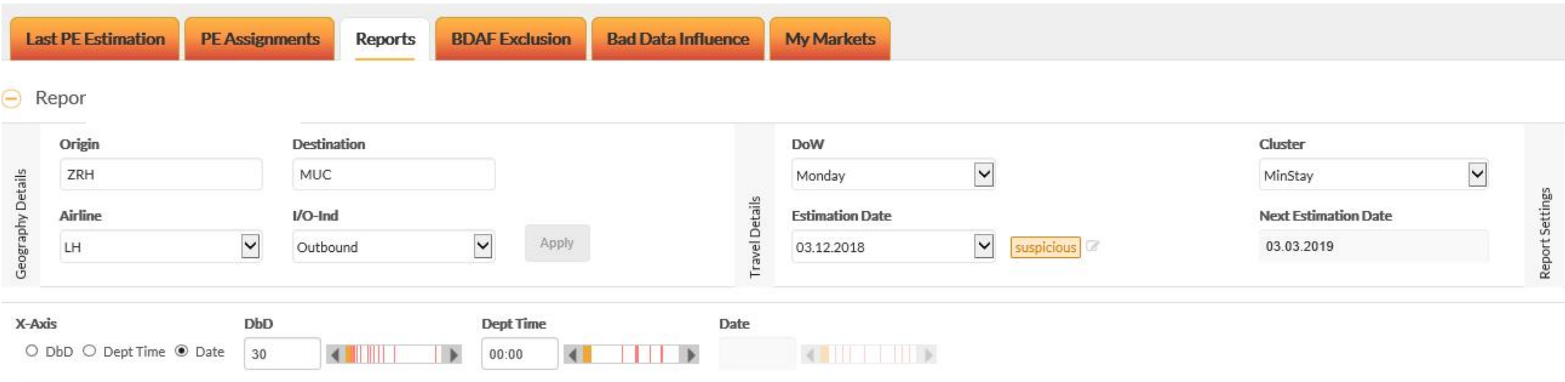


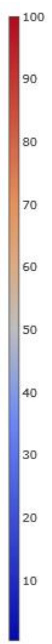
(AGAM corresponds to the reference model whereas 2SAGAM to the reference model where the potential endogeneity of the price variable has been accounted for.)

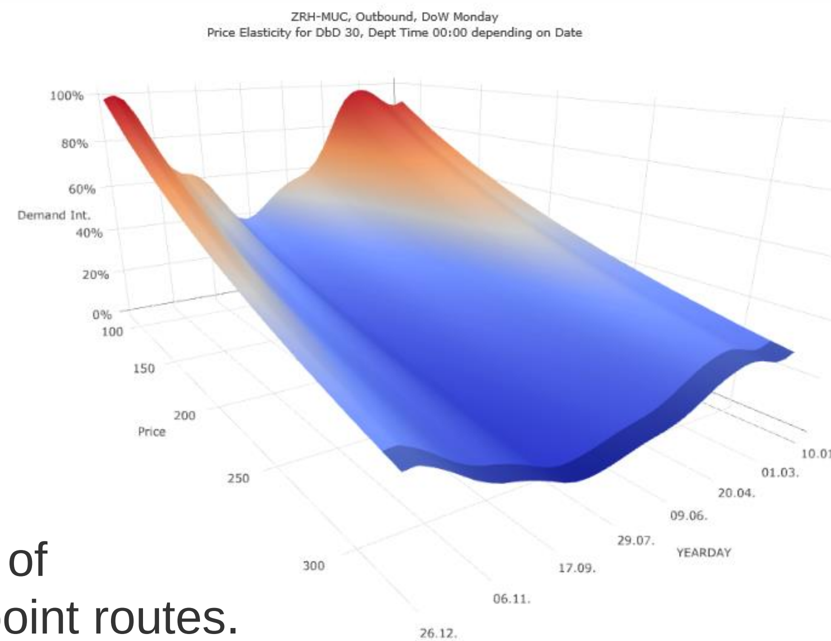
Results: prediction of optimal price values.



The price elasticity model in action (user screen)

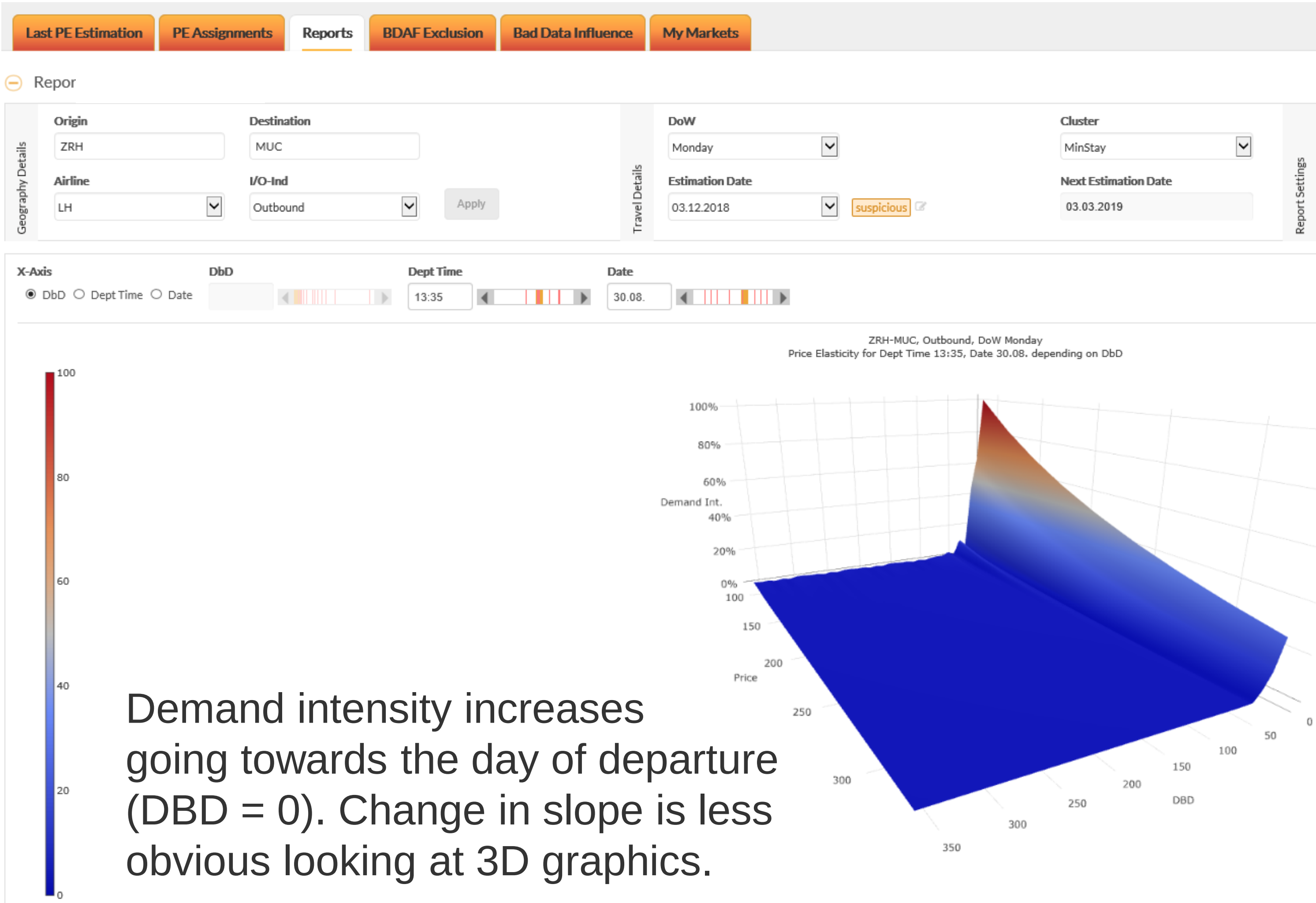






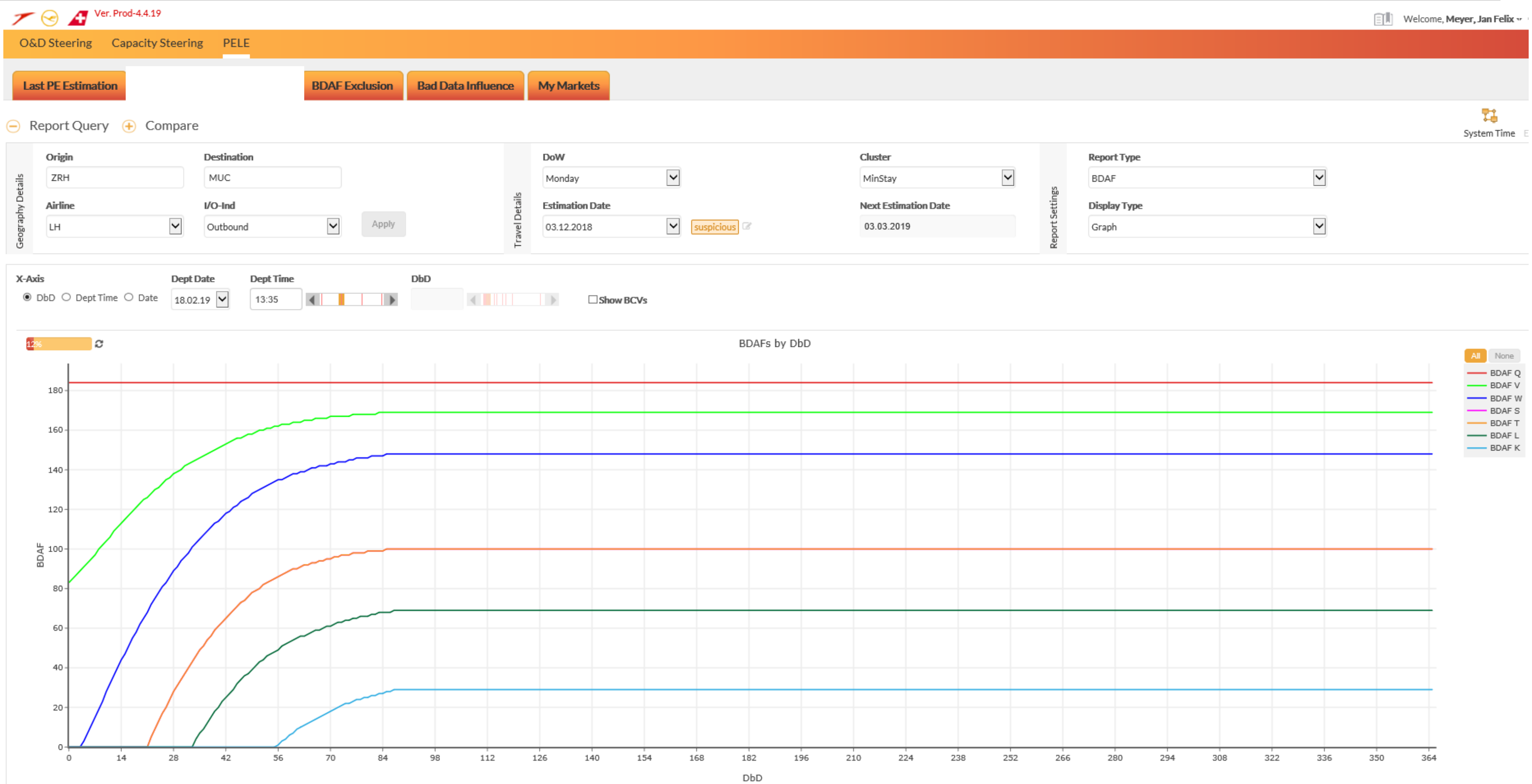
Elasticity model
controls prices (estimation
of willingness to pay) for all of
LHGs continental point-to-point routes.

The price elasticity model in action (user screen)



Demand intensity increases going towards the day of departure (DBD = 0). Change in slope is less obvious looking at 3D graphics.

The price elasticity model in action (user screen)



Much of LHGs distribution is still done via channels depending on booking-classes. Dynamic pricing is achieved by the adjustments of net-values to reduce its value below zero to make the class “unavailable for booking”.

THANK YOU

References

P-spline anova-type interaction models for spatio-temporal smoothing.
by Lee, Dae-Jin, Mara Durbán (2011)
Statistical Modelling 11(1) 49–69
doi:10.1177/1471082X1001100104.
http://smj.sagepub.com/content/11/1/49.abstract

A flexible instrumental variable approach
by Marra, Giampiero, Rosalba Radice (2011)
11(6) 581–603
doi:10.1177/1471082X1001100607

Estimating primary demand for substitutable products from sales transaction data.
by Vulcano, Gustavo, Garrett van Ryzin, Richard Ratliff (2012)
Oper. Res. 60(2) 313–334
doi:10.1287/opre.1110.1012
http://dx.doi.org/10.1287/opre.1110.1012

Demand forecasting and measuring forecast accuracy.
by Fiig, T., R. Hardling, S. Pölt, C. Hopperstad (2014)
Journal of Revenue and Pricing Management 13(6) 413–439
http://dx.doi.org/10.1057/rpm.2014.29.

References

Additive Models with Shape Constraint.
by Pya N. (2010)
Phd. Thesis, University of Bath

Flexible smoothing with B-splines and penalties.
by Eilers, Paul H. C., Brian D. Marx. (1996)
Stat. Science, 11 89–121

Better unconstraining of airline demand data in revenue management systems for improved forecast accuracy and greater revenues.
by Weatherford, Larry., Stefan Pölt. (2002)
Journal of Revenue and Pricing Management 1 234–254

Generalized Additive Models an introduction with R.
by Simon Wood (2006)
CRC Press

.

References

Modelling Retail Demand an Price Elastictiy for Passenger Flights.
by Felix Meyer, Mike Smith, Göran Kauermann
working paper LMU-Munich (2019)

Modeling price-sensitive demand: An application to continuous pricing.
by Felix Meyer, Göran Kauermann, Christopher Alder, Catherine Cleophas
working paper LMU-Munich (2019)

The unknown functions are approximated by the weighted sum of local P-spline basis functions (P-splines)

Univariate functions (example):

$f_k(.)$ is approximated by $\mathbf{x}_k(.)\boldsymbol{\gamma}_k$ where the $n \times m$ matrix $\mathbf{x}_k(.) = \left(\mathbf{x}_{k,1}(.), \mathbf{x}_{k,2}(.), \ldots, \mathbf{x}_{k,m}(.)\right)$ is represented by B-spline basis functions $\mathbf{x}_{k,j}(.)$.

Bivariate functions (example):

$f_{p,t}(.,.)$ is replaced by $\mathbf{x}_{p,t}(.,.)\boldsymbol{\gamma}_{p,t}$ where $\mathbf{x}_{p,t}(.,.) = \mathbf{x}_p(.)\Box\mathbf{x}_t(.)$ is the box-product (row-wise kronecker-product) of its marginals.

Model parameters are therefore $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma})^T$ where

- $\boldsymbol{\beta} = (\beta_0, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{p,I_1})^T$ concerns the parametric covariates and
- $\boldsymbol{\gamma} = \left(\gamma_p \gamma_t, \gamma_1, \ldots, \gamma_{I_2}, \gamma_{p,t}, \gamma_{p,1}, \ldots, \gamma_{p,I_2}, \gamma_{t,1}, \ldots, \gamma_{t,I_2}, \gamma_{1,2}, \ldots, \gamma_{I_2-1,I_2}\right)^T$ is the coefficient vector for the unknown functions which weight the corresponding B-spline basis functions.

The model parameters are derived by penalized maximum likelihood estimation.

We maximize the penalized log-likelihood:

$$\ell_p(\boldsymbol{\theta},\rho) = \ell(\boldsymbol{\theta}) + \rho_p \gamma_p^T D_p \gamma_p + \rho_{p,t} \gamma_{t,p}^T D_{t,p} \gamma_{t,p}$$
$$+ \sum_{k\in I_2} \rho_{p,k} \gamma_{p,k}^T D_{p,k} \gamma_{p,k} + \rho_t \gamma_t^T D_t \gamma_t + \sum_{k\in I_2} \rho_k \gamma_k^T D_k \gamma_k$$
$$+ \sum_{k\in I_2} \rho_{t,k} \gamma_{t,k}^T D_{t,k} \gamma_{t,k} + \sum_{\substack{k_1<k_2\\ k_1,k_2\in I_2}} \rho_{k_1,k_2} \gamma_{k_1,k_2}^T D_{k_1,k_2} \gamma_{k_1,k_2}$$

where the model log-likelihood equals to: $\ell(\boldsymbol{\theta}) = \sum_{i=1}^{M} \sum_{t=t_i^{\text{close}}}^{t_i^{\text{open}}} y_{i,t} \log\left(\lambda(\mathbf{x}_{i,t},t;\boldsymbol{\theta})\right) - \lambda(\mathbf{x}_{i,t},t;\boldsymbol{\theta})$

Smoothing matrices D result from taking differences of neighboring weighting coefficients γ (Eilers and Marx,1996).

The optimal penalty parameters

$$\boldsymbol{\rho} = (\rho_p, \rho_{p,t}, \rho_{p,1}, \ldots, \rho_{p,I_2}, \rho_t, \rho_1, \ldots, \rho_{I_2}, \rho_{t,1}, \ldots, \rho_{t,I_2}, \rho_{1,2}, \ldots, \rho_{I_2-1,I_2})^T$$

are selected by minimizing: $BIC(\boldsymbol{\rho}) = -2\ell(\hat{\boldsymbol{\theta}}) + \log(n)\mathrm{df}(\boldsymbol{\rho})$