

# An integrated airline scheduling, fleet and pricing model for a monopolized market

Bilge Atasoy \*

Matteo Salani †

Michel Bierlaire \*

May 1, 2012

Report TRANSP-OR 120501  
Transport and Mobility Laboratory  
Ecole Polytechnique Fédérale de Lausanne  
`transp-or.epfl.ch`

## Abstract

In airline schedule planning models, the demand and price information are usually taken as inputs to the model. Therefore schedule and capacity decisions are taken separately from pricing decisions. In this paper we present an integrated scheduling, fleet and pricing model for a single airline where these decisions are taken simultaneously. This integration enables to explicitly model supply and demand interactions and take superior decisions. The model refers to a monopolized market. However, competing airlines are included in the model as a reference for the pricing decisions. The pricing decision is formulated through an itinerary choice model which determines the demand of the alternative itineraries in the same market according to their price, travel time, number of stops, and the departure time of the day. The demand model is estimated based on real data and is developed separately for economy and business classes. The seat allocation for these classes are optimized according to the behavior of the demand. The choice model is also used to appropriately model the spill and recapture effects. The resulting model is evaluated with different illustrations and the added value of the integrated approach is analyzed compared to a sequential approach. Results over a set of representative instances show that the integrated model is able to take superior decisions by jointly adjusting capacity and pricing.

Keywords: Fleet assignment; integrated schedule planning; pricing; supply-demand interactions; spill and recapture effects; itinerary choice

---

\*Transport and Mobility Laboratory (TRANSP-OR), School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. Email: {bilge.kucuk,michel.bierlaire}@epfl.ch

†Dalle Molle Institute for Artificial Intelligence (IDSIA), Switzerland. Email: matteo.salani@idsia.ch

# 1 Introduction

The increase in the mobility needs of individuals is an indispensable fact for the last decades. According to the statistics provided by the Association of European Airlines (AEA), air travel traffic has grown at an average rate of 5% per year over the last three decades. Similarly, Bureau of Transportation Statistics reports that the number of departures performed increased by 30% in the last decade. This increase in air travel demand justifies the need for improving the demand responsiveness of air transportation capacity. The underlying demand process should be understood and included in airline scheduling models for more profitable scheduling decisions. The air transportation capacity is determined by the fleet assignment process and is a good candidate to analyze the impacts of the integration of demand models. In this paper, we study the integrated fleet assignment and schedule design models where we further integrate pricing decisions to better represent the supply-demand interactions compared to state of the art studies.

In airline fleet assignment literature, there are various extensions of the basic fleet assignment model (Sherali et al., 2006). One of the extensions is the integration of schedule design decision in fleet assignment models. These integrated models are studied with the purpose of increasing the revenue by making simultaneous decisions on the schedule and the fleet assignment. Schedule design is handled in different ways according to the flexibility allowed for the changes in the schedule. Desaulniers et al. (1997) and Rexing et al. (2000) study the option of shifting departure and arrival times within a given time-windows. Lohatepanont and Barnhart (2004) work with sets of mandatory and optional flights where optional flights can be canceled when not sufficiently profitable.

In airline scheduling decisions, demand and price values are usually taken as inputs to the models. However, supply and demand depend on each other, that is decisions taken for supply influence the demand figures and vice versa. Therefore it is critical to analyze the underlying travel behavior to have more successful airline scheduling decisions. In the literature, random utility models have been used to model the choice of itinerary depending on various attributes. Coldren et al. (2003) propose logit models and Coldren and Koppelman (2005) extend the previous work with the introduction of Generalized Extreme Value (GEV) and nested logit models. Koppelman et al. (2008) model the time of day preferences under a logit setting in order to analyze the effect of schedule delay. Carrier (2008) and Wen and Lai (2010) propose some advanced demand modeling in which customer segmentation is modeled as a latent class. We refer to the work of Garrow (2010) for a comprehensive review of different specifications of choice behavior models for air travel demand.

Supply-demand interactions are considered in fleeting models from different perspectives. Yan and Tseng (2002) study an integrated schedule design and fleet assignment model in which the set of flight legs is built considering the itineraries under a given expected demand for every origin-destination pair. In the context of itinerary-based fleet assignment spill and recapture effects can be integrated in the model. These effects represent the potential number of passengers that could be redirected to alternative itineraries in the market when there is a capacity restriction on their desired itinerary. This information can be considered by the airlines in the planning phase in order to more effectively decide on the capacity. Barnhart et al. (2002) consider the spill and recapture effects separately for each fare class resulting from insufficient capacity. Similarly, Lohatepanont and Barnhart (2004) study the network effects including the demand adjustment in case of flight cancellations and spill effects. More recently, Dumas et al. (2009) model the passenger flow which gives the distribution of demand for each itinerary. This passenger flow

model is also used as an estimation for the recapture ratios between itineraries. Cadarsoa and Marín (2011) include passenger considerations through a schedule development based on passenger satisfaction. Their integrated schedule design and fleet assignment model takes into account the disrupted and misconnected passengers.

Advanced supply and demand interactions can be modeled by letting the model to optimize itinerary's attributes (e.g., the price, departure time). There are studies in the context of schedule planning of airlines where utility of passengers are considered when deciding on the frequency (Brueckner and Zhang, 2001; Brueckner and Flores-Fillol, 2006). Similarly, Vaze and Barnhart (2010) work on a game theoretical framework where they include an S-curve demand model to represent the impact of frequency on the demand. When we move the focus back to fleet assignment literature, Talluri and van Ryzin (2004a) integrate discrete choice modeling into the single-leg, multiple-fare-class revenue management model that determines the subset of fare products to offer at each point in time. They provide the characterization of optimal policies under a general choice model of demand. To overcome the missing no-purchase information in airline booking data, they use expectation-maximization (EM) method. Schön (2006) develops a market-oriented integrated schedule design and fleet assignment model with integrated pricing decisions. It is assumed that customers can be segmented according to some characteristics and different fares can be charged for these segments. Schön (2008) gives several specifications for the inverse price-demand function described in Schön (2006) including logit and nested logit models where the explanatory variable is the price of the itinerary. Budhiraja et al. (2006) also work on a similar topic where the change in unconstrained itinerary demand is incorporated into the model as a function of supply.

In this paper, we introduce an integrated scheduling, fleet and pricing model in a monopolized market. Integration of pricing decisions in schedule planning enables to capture supply-demand interactions and improve the profitability of the schedule plan. The pricing decisions are captured by a demand model which provides profitable average prices for each cabin class. This demand model is an itinerary choice model where the utilities of the itineraries are defined by the price, the travel time, the number of stops and the departure time of the day. The choice model is estimated with real data composed of two data sources: a revealed preferences (RP) data and a stated preferences (SP) survey. Different demand models are estimated for economy and business classes. The capacity allocated to each class is optimized according to their corresponding demand model. The developed itinerary choice model is adapted to model the spill and recapture effects. Since the demand model is endogenously included in the model, these effects are also elastic to the changes in the attributes of the itineraries. With all the listed considerations, the resulting model optimizes the schedule design, fleet assignment, average price, and seat allocation for each cabin class. The added value of the integrated model is analyzed through various illustrations and experiments. To the best of our knowledge the integrated model is not studied in literature. The schedule planning model is close to the work of Lohatepanont and Barnhart (2004). However they include the given demand as an input to the model so that the demand is inelastic to the attributes of the itineraries. Similarly, they use preprocessed recapture ratios to represent supply-demand interactions. A variant of the integration of pricing decision in schedule planning is presented by Schön (2008). However it is carried out with a demand model where the utility is defined by only the price of the itinerary. Spill and recapture effects are ignored. Moreover the demand model and the solution of the integrated model is based on synthetic data. In order to have a convex formulation Schön (2008) utilizes the

inverse demand function rather than the logit formula itself. However this restricts the model for the inclusion of more policy variables and socio-economic characteristics. The model presented in this article integrates the logit formula explicitly which brings flexibility for such extensions and allows for disaggregate models accounting for heterogeneity of behavior in the market.

The rest of the paper is organized as follows. In section 2 we describe the demand model and explain how it is integrated with schedule planning decisions. In section 3 we present our integrated model. Section 4 provides reference models based on the state-of-the-art models in order to be compared with the integrated model. In section 5 we illustrate the added value of the integrated model in comparison to the reference models and provide computational experiments. Finally we conclude the paper in section 6.

## 2 Demand model

Demand forecasting models of airlines are critical in a profitable planning of the network and schedule. In the last decade discrete choice methodology has been introduced in the context of demand analysis of airlines (Garrow, 2010). It has been shown by Coldren et al. (2003) that discrete choice modeling leads to superior forecasts compared to a widely used Quality Service Index (QSI). Therefore in order to better represent the air travel demand we apply discrete choice methodology.

We develop an itinerary choice model which maintains the endogenous pricing decision in the integrated model. *Itinerary* is referred as each available product, which may include more than one flight leg, for a market segment. The market segments,  $s \in S^h$ , are defined by the origin and destination (OD) pairs and they are differentiated for each cabin class  $h$ . Considered classes are economy and business classes and therefore we have two segments for each OD pair. The choice situation is defined for each segment  $s$  with a choice set of all the alternative itineraries in the segment represented by  $I_s$ . The index  $i$  for each alternative itinerary in segment  $I_s$  carries the information of the cabin class of the itinerary due to the definition of the segments. In order to represent the market conditions, we include *no-revenue options* ( $I'_s \subset I_s$ ), which can be considered as the itineraries offered by competitive airlines. These options are dummy itineraries that are included as a reference market price for the demand model. They do not respond to the changes made by the airline which is an assumption we make and can be applicable depending on the context.

The utility of each alternative itinerary  $i$ , including the no-revenue options, is represented by  $V_i$  and the specification is provided in Table 1. The alternative specific constants,  $ASC_i$ , are included for each itinerary in each segment except one of them which is normalized to 0 for identification purposes. Other parameters are represented by  $\beta$  for each of the explanatory variables. Since we have different models for economy and business classes all the parameters and variables are specified accordingly. The parameters with a superscript  $E$  constitute the model for economy itineraries and the parameters with  $B$  represent the model for business itineraries. The superscripts  $NS$  and  $S$  are used to indicate whether the itinerary is a non-stop or a one-stop itinerary. The explanatory variables are described as follows:

- $p_i$  is the price of itinerary  $i$  in €, which is normalized by 100 for scaling purposes,
- $time_i$  is the elapsed time for itinerary  $i$  in hours,

- $non-stop_i$  is a dummy variable which is 1 if itinerary  $i$  is a non-stop itinerary, 0 otherwise,
- $stop_i$  is a dummy variable which is 1 if itinerary  $i$  is a one-stop itinerary, 0 otherwise,
- $economy_i$  is a dummy variable which is 1 if itinerary  $i$  is an economy itinerary, 0 otherwise,
- $business_i$  is a dummy variable which is 1 if itinerary  $i$  is a business itinerary, 0 otherwise,
- $morning_i$  is a dummy variable which is 1 if itinerary  $i$  is a morning itinerary departing between 07:00-11:00, 0 otherwise. The time slot is inspired by the studies in literature that show that the individuals have higher utility for the departures in this slot(Garrow, 2010).

Table 1: Specification table of the utilities

	Parameters	Explanatory variables
constants	$ASC_i^E$ $ASC_i^B$	$1 \times economy_i$ $1 \times business_i$
price	$\beta_p^{E,NS}$ $\beta_p^{B,NS}$ $\beta_p^{E,S}$ $\beta_p^{B,S}$	$\ln(p_i/100) \times non-stop_i \times economy_i$ $\ln(p_i/100) \times non-stop_i \times business_i$ $\ln(p_i/100) \times stop_i \times economy_i$ $\ln(p_i/100) \times stop_i \times business_i$
time	$\beta_{time}^{E,NS}$ $\beta_{time}^{B,NS}$ $\beta_{time}^{E,S}$ $\beta_{time}^{B,S}$	$time_i \times non-stop_i \times economy_i$ $time_i \times non-stop_i \times business_i$ $time_i \times stop_i \times economy_i$ $time_i \times stop_i \times business_i$
time-of-day	$\beta_{morning}^E$ $\beta_{morning}^B$	$morning_i \times economy_i$ $morning_i \times business_i$

As seen in Table 1 all the parameters are interacted with the *economy* and *business* dummies in order to be able to have two different models for the two classes. In addition to the interaction with the cabin class, the time and price variables are interacted with the number of stops, i.e. the dummies of *non-stop* and *stop* since there are strong correlations between the number of stops and both the time and price of the itinerary. Furthermore, the price variable is included as a log formulation since it improved the model significantly. The idea behind is that, the effect of the increase in price is not linear for different levels of the price.

The explanatory variables include the price,  $p_i$ , as a policy variable which can be controlled by the integrated model. The other explanatory variables are context variables which we denote by the vector  $z_i$ . These context variables provide information for the demand and improves the estimation of the market shares but can not be modified by the integrated model. In order to explicitly represent these variables we refer to the utilities  $V_i$  as  $V_i(p_i, z_i; \beta)$ , where  $\beta$  are parameters estimated from real data.

The choice model is formulated as a logit model. It gives the choice probability for each itinerary  $i$  in segment  $s$  and when multiplied with the total forecasted demand of

the segment,  $D_s$ , it provides the estimated demand of each itinerary as represented by equation 1.

$$\tilde{d}_i = D_s \frac{\exp(V_i(p_i, z_i; \beta))}{\sum_{j \in I_s} \exp(V_j(p_j, z_j; \beta))} \quad \forall h \in H, s \in S^h, i \in I_s \quad (1)$$

The itinerary choice model is also used to model the interactions between the itineraries in case of capacity shortage. Passengers, who can not be accommodated on their desired itineraries, may be redirected to other available itineraries in the same market segment in case of such shortages. This effect is referred as spill and recapture effect. It is important to note that, the spill effects are not considered in the day of operations but rather in the planning phase. Airlines can take advantage of this knowledge when planning for the schedule and the design of fleet capacity. They can keep their capacity at profitable levels by taking into account the possibility of redirecting passengers to the alternative itineraries. For example, for a flight with a forecasted demand of 100 passengers, the airline may investigate the option of assigning an aircraft with 70 seats. If there are similar alternatives in the same market by the same airline, the airline may assume that a portion of 30 spilled passengers will still fly on those itineraries. Therefore, the spill and recapture information is not communicated to the passengers but only investigated at the planning phase. The passengers will be aware of the capacity limits at the booking phase as usual.

We assume that the spilled passengers are recaptured by the other itineraries with a recapture ratio based on the logit formulation. Therefore the recapture ratio is represented by equation (2).

$$b_{i,j} = \frac{\exp(V_j(p_j, z_j; \beta))}{\sum_{k \in I_s \setminus \{i\}} \exp(V_k(p_k, z_k; \beta))} \quad \forall h \in H, s \in S^h, i \in (I_s \setminus I'_s), j \in I_s. \quad (2)$$

The recapture ratios  $b_{i,j}$  represent the proportion of recaptured passengers by itinerary  $j$  among  $t_{i,j}$  number of spilled passengers from itinerary  $i$ . The recapture ratio is calculated for the itineraries that are in the same market segment where the desired itinerary  $i$  is excluded from the choice set. Therefore lost passengers may be recaptured by the remaining alternatives of the company or by the no-revenue options. Since no-revenue itineraries are out of the network we assume that no spill exist from them.

For the estimation of the demand model we use an RP data provided in the context of ROADEF Challenge 2009<sup>1</sup>. This is a booking data from a major European airline which provides the set of airports, flights, aircraft and itineraries. The information provided for the itineraries includes the corresponding flight legs. Therefore, we can deduce the information on the departure and arrival time of itinerary, the trip length and the number of stops. Additionally, we have information on the demand and average price (€) for each cabin class. We used a subset of this data which correspond to 3 market segments resulting with 30 available itineraries that have a total estimated demand of 904 passengers. The average price for all the 30 alternatives is 338 €. The RP data does not include any information concerning the competitive airlines. Therefore the no-revenue options are not considered in the estimation process. However for applying the model we assume that these itineraries have the same type of utility functions as presented in Table 1 and their attributes are assigned according to the other available itineraries in the market offered by competitive airlines.

---

<sup>1</sup><http://challenge.roadef.org/2009/en>

As it is common with RP data, the lack of variability in some attributes precludes a statistically significant estimation of key parameters of the choice models. Therefore, the RP data is combined with SP data, where the variability is obtained by design. This SP data is based on an Internet choice survey collected in 2004 in the US. The data includes information for 3609 passengers. Let us note that, the combined dataset therefore contains both European and US data. The Internet survey was organized to understand the sensitivity of air passengers to the attributes of an airline itinerary such as fare, travel time, number of stops, legroom, and aircraft. The respondents were presented hypothetical choice situations and offered three alternatives. The first is a non-stop itinerary, the second one is a one-stop itinerary with the same airline and the third is connecting with a different airline. By design, the data has enough variability in terms of price and other variables. The average price for all the proposed alternatives in the survey is 405 \$. For the estimation, the parameters of the logit model corresponding to the RP data are constrained to be the same as those of the SP data. The estimation of the two logit models for the two data sets is carried out simultaneously. For the simultaneous estimation we introduce a scale parameter for the utilities of the SP model to capture the vector of variances for the error terms of the two models. The SP model is only used to take the advantage of the elastic behavior of the SP data. It is not used for the integrated model. The details on the SP model are provided in A.

Table 2: Estimated parameters for the model with joint RP and SP data

	$\beta_p$		$\beta_{time}$		$\beta_{morning}$
	non-stop	stop	non-stop	stop	
economy	-2.23	-2.17	-0.102	-0.0762	0.0283
business	-1.97	-1.96	-0.104	-0.0821	0.0790

The estimation of the parameters is done with a maximum likelihood estimation using the software BIOGEME (Bierlaire and Feterison, 2009). The resulting parameters can be seen in Table 2. The cost and time parameters have negative signs as expected since the increase in the price or the time of an itinerary decreases its utility. They also indicate that, economy demand is more sensitive to price and less sensitive to time compared to business demand as expected (Belobaba et al., 2009). Departure time of the day parameter,  $\beta_{morning}$ , is higher for business demand compared to the economy demand, which means that business passengers have a higher tendency to chose morning itineraries.

In addition to the estimated parameters, the price elasticities are computed to have more insight on the underlying demand behavior. For an arbitrary OD pair we present the price elasticities for 4 alternatives with the given attributes in Table 3. It is observed that business demand is less elastic to price compared to the economy demand as expected (Belobaba et al., 2009). Moreover the demand for non-stop itineraries is less elastic to the increase in price. In other words, passengers are ready to pay more for the non-stop alternatives compared to the connecting ones (Garrow, 2010).

Table 3: Price elasticities of different type of itineraries

Alternatives	stops	class	price (€)	price elasticity
1	one-stop	business	656	-1.95
2	one-stop	economy	564	-2.15
3	non-stop	business	410	-1.90
4	non-stop	economy	175	-2.01

The details on the demand model and results on the demand indicators such as the price and time elasticities as well as the willingness to pay are provided in Atasoy and Bierlaire (2012).

In order to illustrate the behavior of the demand model together with the spill and recapture effects we choose an arbitrary OD pair A-B. There are two alternatives of economy itineraries which are both nonstop itineraries with the same travel time. We include the no-revenue itinerary A-B'. The values of attributes can be seen in Table 4. According to the attributes, the resulting choice probability, which is referred as the *market share*, is presented in the last column. The itinerary 2 has the lowest price and is a morning itinerary. Therefore it attracts the biggest number of passengers.

With the same example we illustrate the spill and recapture effects. The resulting ratios according to the given attributes are presented in Table 5. For example, in case of capacity shortage for itinerary 1, at most 55% of spilled passengers will be recaptured by itinerary 2 and 45% will be lost to the itineraries offered by competitive airlines. Since the price of itinerary 2 is lower than the price of competitors, the probability to be recaptured by itinerary 2 is higher.

Table 4: Attributes of the itineraries and the resulting market shares

OD	price	morning	market share
A-B <sub>1</sub>	225	0	0.26
A-B <sub>2</sub>	203	1	0.44
A-B'	220	0	0.30

Table 5: Resulting recapture ratios

	A-B <sub>1</sub>	A-B <sub>2</sub>	A-B'
A-B <sub>1</sub>	0	0.552	0.448
A-B <sub>2</sub>	0.487	0	0.513

### 3 Integrated scheduling, fleet and pricing model

In this section, we introduce an integrated scheduling, fleet and pricing model for a single airline. We explicitly model the demand and integrate it in the schedule planning which enables to make use of the interaction between supply and demand.

Let  $F$  be the set of flight legs, there are two subsets of flights: mandatory flights ( $F^M$ ), which should be flown, and optional flights ( $F^O$ ) which can be canceled. The included schedule design context is solely related to the optional flights, apart from that the schedule is known and assumed to be used without any change.  $A$  represents the set of airports and  $K$  is for the fleet where each type of aircraft in the fleet is indexed by

$k$ . The schedule is represented by time-space network such that  $N(k, a, t)$  is the set of nodes in the time-line network for plane type  $k$ , airport  $a$  and time  $t \in T$ .  $\text{In}(k, a, t)$  and  $\text{Out}(k, a, t)$  are the sets of inbound and outbound flight legs for node  $(k, a, t)$ .

$$\max \sum_{h \in H} \sum_{s \in S^h} \sum_{i \in (I_s \setminus I'_s)} (d_i - \sum_{j \in I_s} t_{i,j} + \sum_{j \in (I_s \setminus I'_s)} t_{j,i} b_{j,i}) p_i - \sum_{\substack{k \in K \\ f \in F}} C_{k,f} x_{k,f} \quad (3)$$

$$\text{s.t.} \sum_{k \in K} x_{k,f} = 1 \quad \forall f \in F^M \quad (4)$$

$$\sum_{k \in K} x_{k,f} \leq 1 \quad \forall f \in F^O \quad (5)$$

$$y_{k,a,t^-} + \sum_{f \in \text{In}(k,a,t)} x_{k,f} = y_{k,a,t^+} + \sum_{f \in \text{Out}(k,a,t)} x_{k,f} \quad \forall [k, a, t] \in N \quad (6)$$

$$\sum_{a \in A} y_{k,a, \min E_a^-} + \sum_{f \in CT} x_{k,f} \leq R_k \quad \forall k \in K \quad (7)$$

$$y_{k,a, \min E_a^-} = y_{k,a, \max E_a^+} \quad \forall k \in K, a \in A \quad (8)$$

$$\sum_{s \in S^h} \sum_{i \in (I_s \setminus I'_s)} \delta_{i,f} d_i - \sum_{j \in I_s} \delta_{i,f} t_{i,j} + \sum_{j \in (I_s \setminus I'_s)} \delta_{i,f} t_{j,i} b_{j,i} \leq \sum_{k \in K} \pi_{k,f}^h \quad \forall h \in H, f \in F \quad (9)$$

$$\sum_{h \in H} \pi_{k,f}^h \leq Q_k x_{k,f} \quad \forall f \in F, k \in K \quad (10)$$

$$\sum_{j \in I_s} t_{i,j} \leq d_i \quad \forall h \in H, s \in S^h, i \in I_s \quad (11)$$

$$\tilde{d}_i = D_s \frac{\exp(V_i(p_i, z_i; \beta))}{\sum_{j \in I_s} \exp(V_j(p_j, z_j; \beta))} \quad \forall h \in H, s \in S^h, i \in I_s \quad (12)$$

$$b_{i,j} = \frac{\exp(V_j(p_j, z_j; \beta))}{\sum_{k \in I_s \setminus \{i\}} \exp(V_k(p_k, z_k; \beta))} \quad \forall h \in H, s \in S^h, i \in (I_s \setminus I'_s), j \in I_s \quad (13)$$

$$x_{k,f} \in \{0, 1\} \quad \forall k \in K, f \in F \quad (14)$$

$$y_{k,a,t} \geq 0 \quad \forall [k, a, t] \in N \quad (15)$$

$$\pi_{k,f}^h \geq 0 \quad \forall h \in H, k \in K, f \in F \quad (16)$$

$$0 \leq d_i \leq \tilde{d}_i \quad \forall h \in H, s \in S^h, i \in I_s \quad (17)$$

$$0 \leq p_i \leq UB_i \quad \forall h \in H, s \in S^h, i \in I_s \quad (18)$$

$$t_{i,j} \geq 0 \quad \forall h \in H, s \in S^h, i \in (I_s \setminus I'_s), j \in I_s \quad (19)$$

$$b_{i,j} \geq 0 \quad \forall h \in H, s \in S^h, i \in (I_s \setminus I'_s), j \in I_s \quad (20)$$

Objective (3) is to maximize the profit calculated as revenue minus operating costs. The revenue is the sum of the revenues for business and economy passengers taking into account the lost revenue due to spill. The price of the itinerary  $i$  is represented by  $p_i$ . Operating cost for flight  $f$  when using fleet type  $k$  is represented by  $C_{k,f}$  which is associated with a binary variable of  $x_{k,f}$  that is one if a plane of type  $k$  is assigned to flight  $f$ .

Firstly, we have the fleet assignment constraints. Constraints (4) ensure the coverage of mandatory flights which must be served according to the schedule development.

Constraints (5) are for the optional flights that have the possibility to be canceled. Constraints (6) are for the flow conservation of fleet, where  $y_{k,a,t^-}$  and  $y_{k,a,t^+}$  are the variables representing the number of type  $k$  planes at airport  $a$  just before and just after time  $t$ . Constraints (7) ensure that for each fleet type  $k$ , the number of used planes does not exceed the number of available planes represented by  $R_k$ .  $\min E_a^-$  represents the time just before the first event at airport  $a$  and CT is the set of flights flying at count time. It is assumed that the network configuration at the beginning and at the end of the day is the same in terms of the number of planes at each airport. This is ensured by the constraints (8) where  $\max E_a^+$  represents the time just after the last event at airport  $a$ .

The relation between the supply capacity and the actual demand should be maintained. Therefore we have the constraints (9) which maintain that the assigned capacity for a flight should satisfy the demand for the corresponding itineraries. The assigned capacity for flight  $f$  by a plane type  $k$  for class  $h$  passengers is represented by  $\pi_{k,f}^h$ . The actual demand is composed of the original demand of the itinerary minus the spilled passengers plus the recaptured passengers from other itineraries. The same constraints ensure that the itineraries do not realize any demand if any of the corresponding flight leg is canceled.  $\delta_{i,f}$  is a binary parameter which is one if itinerary  $i$  uses flight  $f$  and enables us to have itinerary-based demand. We let the allocation of business and economy seats to be decided by the model as a revenue management decision. Therefore, we need to make sure that the total allocated seats does not exceed the capacity of the aircraft. This is ensured by the constraints (10) where  $Q_k$  is the capacity of plane type  $k$ .

Demand related constraints include constraints (11) which maintain that the total redirected passengers from itinerary  $i$  to all other itineraries including the no-revenue options do not exceed its realized demand. We have already explained the constraints (12) and (13) in section 2.

Finally, we have the non-negativity constraints and upper bounds (14)-(20) for the decision variables. Demand value provided by the logit model,  $\tilde{d}_i$ , serves as an upper bound for the realized demand,  $d_i$ . The price of each itinerary has an upper bound  $UB_i$ , which is assumed to be the average market price plus the standard deviation.

## 4 Reference models

In order to quantify the impact of the presented integrated model, we consider the state-of-the-art models as reference models which are already cited in section 1. Firstly, we consider the model of Lohatepanont and Barnhart (2004). This model considers the demand and price as inputs to the schedule planning model. We compare our integrated model to a similar model in order to show the added-value of integrating explicit supply-demand interactions. Lohatepanont and Barnhart (2004) use QSI index to model the recapture ratios. Since we do not have access to the parameters of these ratios, we formulate spill and recapture with our itinerary choice model. We refer to this model as *price-inelastic schedule planning model*.

Secondly, we consider the model of Schön (2008) which is an integrated schedule planning and pricing model. This model does not include spill and recapture effects. We compared our integrated model with a similar model, named *integrated model w/o spill*, in order to analyze the added value of the spill and recapture effects. Schön (2008) uses a synthetic data and we do not have access to this data. Therefore, for this reference model we use our estimated parameters. Since our demand model is specific to the cabin class we keep the revenue management decision on the allocation of seats to the classes.

Schön (2008) does not include this decision since she does not consider different cabin classes.

The presented integrated model includes explicit interactions between supply capacity and the demand. On the other hand, many revenue management models assume that the capacity is fixed and provided by the schedule planning process (Talluri and van Ryzin, 2004b). Therefore we consider this current practice of airlines through a *sequential approach* and compare it with the integrated model. In the sequential approach, firstly, the schedule planning model is optimized to obtain an optimal fleet assignment. Then as a sequential step, the revenue maximization is performed with this optimal fleet assignment. In other words, the fleet assignment is decided with an assumption of inelastic demand and then we expose the model to elastic demand. The integrated model is compared with the sequential approach in order to evaluate the advantage of simultaneously optimizing the schedule planning and revenue related decisions. A similar analysis is performed by Lohatepanont (2002) for the comparison of leg-based fleet assignment and itinerary-based fleet assignment.

## 5 Results

In this section we provide illustrations and results to quantify the impact of the integrated model. The data instances are based on the RP data source explained in section 2. We focus on a daily circular schedule.

The presented integrated model is a mixed integer nonlinear problem. Nonlinearity is due to the explicit integration of the demand model. The model is implemented in AMPL and BONMIN solver (Bonami et al., 2008) is used for the solution of the problem. BONMIN solver applies several algorithms depending on the nature of the problem including branch-and-bound, branch-and-cut and outer-approximation. It serves as a heuristic approach since we cannot guarantee the convexity of the problem.

Table 6: The data instance used for the illustrations

Number of airports:	3
Number of flights:	26
Flight density:	4.33 flights per route
Average demand:	56.12 passengers per flight
Number of itineraries:	36
Cabin classes:	Economy and business
Level of service:	All itineraries are nonstop
Available fleet:	3 types of aircraft (100, 50 and 37 seats)

### 5.1 Illustrative example for the impacts of the integrated demand model

In order to analyze the added value of the integrated scheduling, fleet and pricing model, we compare our model with the price-inelastic schedule planning model defined in section 4. The two models are run with the data instance provided in Table 6. In Table 7, we provide the results for the price-inelastic schedule planning model and two sets of results for the integrated model. The integrated model has the flexibility to change the prices of the itineraries which might be higher than the average values used by the price-inelastic schedule planning model. Therefore we first present the results of the integrated

model where we constrain the itinerary prices not to be higher than the average prices used by the price-inelastic schedule planning model (*integrated model with limited prices*). The motivation for constraining the prices is to show that the strength of the integrated model is not only due to the ability to increase the prices, but also the simultaneous decisions which lead to superior schedule planning. Finally we present the results for the integrated model where the prices can be increased above the average prices. We arbitrarily select 2 of the 3 airports and present the realized price and demand values for the bi-directional flights (A-B and B-A) between these airports.

The results in Table 7 indicate that the integrated model has the flexibility to change the fleet assignment decisions simultaneously with the pricing decisions in order to have more profitable planning. For the case of limited prices, the integrated model decreases the prices of itineraries 10 and 11 and assigns larger capacity to them. It is observed that these itineraries are morning itineraries and therefore more attractive itineraries according to the logit model. The model decides to increase the capacity of these itineraries since it can be maintained without significant decrease in the price. It is also observed that the decrease in the prices of the itineraries 10 and 11 affects the demand of the itineraries 6 and 8 respectively and they are assigned smaller capacity. With similar decisions for the other OD pairs, the resulting profit and the number of transported passengers are higher compared to the price-inelastic schedule planning model. When the integrated model is allowed to increase the prices beyond the average prices, the resulting profit and the served demand increases more significantly. The decisions on the prices of the itineraries show that the integrated model increases the prices whenever it sees a potential and decreases the prices when assigning a larger capacity is more profitable.

## **5.2 Illustrative example for the reaction of the integrated model to the market conditions**

One of the most important factors for airlines in their revenue management is the alternative itineraries provided by competitive airlines. As explained in section 2, we introduce no-revenue options in our model to represent the attributes of the competitors' itineraries. Therefore the integrated model takes into account those competitive itineraries offered by other airlines, while optimizing the revenue decisions. In order to illustrate this phenomenon we compare the price-inelastic schedule planning model with the integrated demand model in three different market conditions based on the data instance provided in Table 6. Compared to the actual itineraries, the competitors have lower, similar and higher prices respectively in the presented scenarios.

Table 7: Results on the impact of the integrated demand model

					Price-inelastic schedule planning model			Integrated model - limited prices			Integrated model		
Revenue					204,553			214,380			244,924		
Operating costs					150,603			160,003			173,349		
Profit					<b>53,949</b>			<b>54,377 (+ 0.8%)</b>			<b>71,575 (+ 32.7%)</b>		
Number of flights					<b>22</b>			<b>22</b>			<b>24</b>		
Transported passengers					<b>943</b>			<b>1031 (+ 9.3%)</b>			<b>1064 (+ 12.7%)</b>		
Economy-Business passengers					882 E - 61 B			970 E - 61 B			997 E - 67 B		
Allocated seats					<b>274</b>			<b>324</b>			<b>324</b>		
	OD pair	Class	Morning	Forecasted demand	Average price	Realized demand	Assigned capacity	Realized price	Realized demand	Assigned capacity	Realized price	Realized demand	Assigned capacity
1	B-A	E	0	59	225	50	50	225	50	50	250	50	50
2	A-B	E	1	68	212.5	50	50	212.5	50	50	225	50	50
3	B-A	E	0	59	225	50	50	225	50	50	250	50	50
4	A-B	E	0	66	212.5	50	50	212.5	50	50	225	50	50
5	B-A	B	1	6	526.5	6	<b>50</b>	526.5	6	<b>37</b>	550	6	<b>37</b>
6	A-B	E	0	66	497.2	2	<b>50</b>	497.2	2	<b>37</b>	507.34	2	<b>37</b>
7	B-A	E	0	66	212.5	50	50	212.5	50	50	225	50	50
8	A-B	E	1	60	<b>225</b>	<b>50</b>	50	<b>182.4</b>	<b>100</b>	<b>100</b>	<b>195.15</b>	<b>100</b>	<b>100</b>
9	B-A	E	1	68	<b>212.5</b>	<b>50</b>	50	<b>194</b>	<b>100</b>	<b>100</b>	<b>201.56</b>	<b>100</b>	<b>100</b>
10	A-B	E	0	59	225	50	50	225	50	50	250	50	50
11	B-A	E	0	66	212.5	50	50	212.5	50	50	225	50	50
12	A-B	E	0	59	225	50	50	225	50	50	250	50	50

Table 8: The results with changing market conditions

<b>Competitors with higher prices</b>		
	Price-inelastic schedule planning model	Integrated model
Revenue	206,001	247,269
Operating costs	150,604	173,349
Profit	<b>55,397</b>	<b>73,920 (+ 33%)</b>
Number of flights	<b>22</b>	<b>24</b>
Transported passengers	<b>951</b>	<b>1,076 (+ 13%)</b>
Economy-Business passengers	888 E - 63 B	1007 E - 69 B
Allocated seats	<b>274</b>	<b>324</b>
<b>Competitors with similar prices</b>		
	Price-inelastic schedule planning model	Integrated model
Revenue	202,645	218,456
Operating costs	150,604	149,656
Profit	<b>52,401</b>	<b>68,800 (+ 31%)</b>
Number of flights	<b>22</b>	<b>22</b>
Transported passengers	<b>935</b>	<b>935</b>
Economy-Business passengers	876 E - 59 B	878 E - 57 B
Allocated seats	<b>274</b>	<b>274</b>
<b>Competitors with lower prices</b>		
	Price-inelastic schedule planning model	Integrated model
Revenue	190,590	215,429
Operating costs	140,822	149,656
Profit	<b>49,768</b>	<b>65,773 (+ 32%)</b>
Number of flights	<b>20</b>	<b>22</b>
Transported passengers	<b>871</b>	<b>926 (+ 6%)</b>
Economy-Business passengers	815 E - 56 B	871 E - 55 B
Allocated seats	<b>274</b>	<b>274</b>

The results are provided in Table 8. For the scenario with similar prices, the scheduling decisions of the price-inelastic schedule planning model and the integrated model are the same and therefore the realized demand is similar. In the scenario where the competitors are more expensive, price-inelastic schedule planning model keeps the same scheduling decisions which results with an improvement in the profit and realized demand since the competitors are less attractive. For the same scenario, the integrated model allocates higher capacity and operates one more flight which results with a significant increase in realized demand. The advantage of the integrated model emerges from the fact that either it finds room to attract more passengers or it has a potential to increase the prices. When we analyze the results in the case of cheaper competitors, it is observed that the inelastic demand operates less flights. Since it can not compete with the cheap prices it carries less passengers compared to the other scenarios. However the integrated model can still accommodate a similar level of passengers thanks to the flexibility of decreasing the prices in order to attract passengers.

### 5.3 Illustrative example for the spill and recapture effects

We compare our integrated approach with the integrated model w/o spill described in section 4 in order to analyze the design flexibility of the schedule gained by airlines with the spill and recapture effects. In order to be able to see the impact of the spill more clearly, in this analysis the upper bound on the prices is set to the average price.

The results are presented in Table 9. We select the OD pair C-D arbitrarily among 6 OD pairs to present the impact of spill from an itinerary level. Spill values with “+” sign means that the itinerary recaptures passengers that are spilled from other itineraries. On the other hand, spill values with “-” sign corresponds to the total redirected passengers from the itinerary to the remaining alternatives in the same market. It is observed that the integrated model with spill modifies the prices of the itineraries relatively to capture the passengers of the other itineraries for the same market. For example the itinerary 4 attracts both economy and business passengers from other itineraries and therefore assigned a larger aircraft compared to the integrated model w/o spill. The flexibility of redirecting passengers to other itineraries enables to keep the prices at higher prices. As an example, the integrated model assigns higher price to itinerary 1 compared to the integrated model w/o spill, but realizes the same demand due to recaptured passengers. Furthermore it is observed that itinerary 5 is not operated since some of its passengers were recaptured by the other itineraries and it was more profitable to cancel it. With similar decisions for other OD pairs, the integrated model has higher profit and carries more passengers in the presence of spill and recapture.

### 5.4 Experiments on the added value of the integrated model

In order to see the added value of the integration of the demand model we need to support our observations with a comprehensive set of experiments. For that purpose we identified 15 data instances with different characteristics that are listed in Table 10. For the experiments, we present the number of airports and the number of flights in the network. Moreover, the flight density stands for the average number of flights per route. The average demand gives the average number of passengers per flight according to demand forecast. The fleet composition provides information on the number of different plane types in the fleet together with the seat capacity for each type.

Table 9: Illustration for the spill and recapture effects

					Integrated model w/o spill			Integrated model			
<b>Revenue</b>								214,380			
<b>Operating costs</b>								160,003			
<b>Profit</b>					<b>50,514</b>			<b>54,377 (+ 7.7%)</b>			
<b>Number of flights</b>					24			22			
<b>Transported passengers</b>					<b>972</b>			<b>1031 (+ 6.1%)</b>			
<b>Economy-Business passengers</b>					902 E - 70 B			970 E - 61 B			
<b>Allocated seats</b>					<b>224</b>			<b>324</b>			
	<b>OD pair</b>	<b>Class</b>	<b>Morning</b>	<b>Forecasted demand</b>	<b>Realized price</b>	<b>Realized demand</b>	<b>Assigned capacity</b>	<b>Realized price</b>	<b>Spill</b>	<b>Realized demand</b>	<b>Assigned capacity</b>
1	C-D	E	0	34	<b>160.2</b>	<b>37</b>	37	<b>169</b>	<b>+3</b>	<b>37</b>	37
2	C-D	E	0	32	-	0	0	-	-32	0	0
3	C-D	E	1	34	<b>158.6</b>	<b>39</b>	<b>50</b>	<b>171.4</b>	<b>+3</b>	<b>37</b>	<b>37</b>
4	C-D	B	1	7	<b>409.5</b>	<b>7</b>	<b>37</b>	<b>404.3</b>	<b>+3</b>	<b>10</b>	<b>50</b>
		E	1	35	<b>175</b>	<b>30</b>		<b>168.7</b>	<b>+15</b>	<b>40</b>	
5	C-D	B	0	7	409.5	7	37	-	-7	0	0
		E	0	32	175	30		-	-32	0	
6	C-D	B	0	7	409.5	7	37	409.5		7	37
		E	0	32	175	30		175	-2	30	

Table 10: The experiments

No	Airports	Flights	Flight density	Average demand	Fleet composition
1	3	10	1.67	51.9	2 50-37 seats
2	3	11	2.75	83.1	2 117-50 seats
3	3	12	2	113.8	2 164-100 seats
4	3	26	4.33	56.1	3 100-50-37 seats
5	3	19	3.17	96.7	3 164-117-72 seats
6	3	12	3	193.4	3 293-195-164 seats
7	3	33	8.25	71.9	3 117-70-37 seats
8	3	32	5.33	100.5	3 164-117-85 seats
9	2	11	5.5	173.7	3 293-164-127 seats
10	4	39	4.88	64.5	4 117-85-50-37 seats
11	4	23	3.83	86.1	4 117-85-70-50 seats
12	4	19	3.17	101.4	4 134-117-100-85 seats
13	4	15	1.88	58.1	5 117-85-70-50-37 seats
14	4	14	2.33	87.6	5 134-117-85-70-50 seats
15	4	13	2.6	100.1	5 164-134-117-100-85 seats

For the considered data instances, we compared the sequential approach presented in section 4 and the integrated model. The comparative results are presented in Table 11. It is observed that for 6 of these 15 instances there is an improvement with the integrated model in terms of the profit and transported number of passengers. These are the cases where the simultaneous optimization of the schedule planning and pricing lead to different scheduling decisions such as the operated number of flights or the number of allocated capacity.

Table 11: The results of the experiments

No	Sequential approach				Integrated model				Improvement	
	Profit	Pax.	Flights	Seats	Profit	Pax.	Flights	Seats	Profit	Pax.
1	15,091	284	8	124	15,091	284	8	124	-	-
2	35,372	400	8	<b>150</b>	37,335	534	8	<b>217</b>	<b>5.55%</b>	<b>33.50%</b>
3	50,149	859	10	300	50,149	859	10	300	-	-
4	69,901	931	<b>22</b>	<b>274</b>	70,904	1063	<b>24</b>	<b>324</b>	<b>1.43%</b>	<b>14.18%</b>
5	82,311	1145	16	333	82,311	1145	16	333	-	-
6	779,819	1448	10	1148	779,819	1448	10	1148	-	-
7	135,656	1814	32	498	135,656	1814	32	498	-	-
8	107,927	2236	26	691	107,927	2236	26	691	-	-
9	854,902	1270	10	<b>1016</b>	858,544	1344	10	<b>1090</b>	<b>0.43%</b>	<b>5.83%</b>
10	109,906	1448	<b>32</b>	391	112,881	1541	<b>34</b>	391	<b>2.71%</b>	<b>6.42%</b>
11	82,440	1135	20	387	85,808	1164	20	387	<b>4.09%</b>	<b>2.56%</b>
12	49,448	1050	12	370	49,448	1050	12	370	-	-
13	27,076	448	10	207	27,076	448	10	207	-	-
14	44,339	599	<b>10</b>	267	45,070	699	<b>12</b>	267	<b>1.65%</b>	<b>16.69%</b>
15	26,486	504	6	185	26,486	504	6	185	-	-

We observe that the improvement is higher for the experiments where the demand levels for the flights has high variation but there is a few number of plane types. In those cases, the integrated model is able to adjust the capacity according to the demand

and has significant improvement over the sequential approach. Experiment 2 is a good example for this phenomenon. There are 2 different fleet types with 50 and 117 seats. The sequential approach does not use the larger aircraft which is costlier to fly. On the other hand the integrated model uses this large aircraft thanks to its flexibility in controlling the demand by pricing decisions. As a result, there is a 5.55% increase in profit and 33.5% more passengers are transported. Similarly, for the experiments 4 and 9, the integrated model decides to use more capacity with the knowledge on the demand behavior. In addition to the decision on the allocated capacity, the integrated model may decide to operate less/more flights by changing the attractiveness of the corresponding itineraries as seen in experiments 4, 10, and 14. For example, for experiment 14, the integrated model operates 2 more flights with the same overall capacity compared to the sequential approach.

## 6 Conclusions

In this paper an integrated scheduling, fleetling and pricing model is presented for a single airline, which enables to take the advantage of explicit supply-demand interactions in decision making. The novelty of the model is due to the modeling of the demand through an itinerary choice model based on a real data and the integration of this demand model in a scheduling and fleetling framework for airlines. The demand model is utilized for pricing as well as the spill and recapture effects which gives flexibility to airlines in determining their transportation capacity.

The impacts of the integrated model is evaluated on a European air transportation network with several illustrations. It is observed that the integrated model has more flexibility on the decisions thanks to the simultaneous optimization. The pricing is determined according to the market conditions and whenever there is a potential in increasing the profit by altering the price the integrated model benefits from it. Therefore the integrated model is elastic to the market conditions.

The added value of the integrated model is analyzed in comparison to a sequential approach which mimics the current practice of airlines. It is shown that the integrated model may decide on different scheduling and/or fleetling compared to the sequential approach by making use of the supply-demand interactions. These differences may be in terms of the number of flights operated or the assigned capacity which result with increase in profit and increase in served demand.

The presented analysis shows that the airlines should consider the demand related information earlier in their planning phase when deciding on the schedule and capacity. Our model is a proof of concept for the integration of scheduling, fleetling and pricing decisions which is expected to improve the efficiency of decision support tools of airlines. This effort can motivate the integration of more detailed demand information through disaggregate demand models as a future extension.

The presented integrated model is a mixed integer nonconvex problem which is highly complex. When we go beyond the instances provided in Table 10 in terms of size, the solver can not provide good quality feasible solutions for the integrated model. The non-convexity is due to the explicit logit formula which allows for the integration of more advanced demand models with more policy variables and/or socio-economic characteristics. As a future work we are working on appropriate transformations of the logit formula allowing to convexify the problem while keeping the disaggregate nature of the demand model.

The considered demand model has only one policy variable which is the price of the itinerary. In other words the integrated model can only control the price of the itineraries in order to maximize the profit. However there is a variable for the departure time of the day which indicates whether the itinerary is a morning itinerary or not. As a future work the departure time can be introduced as a policy variable in addition to the price. This will enable the integrated model to take the advantage of the flexibility in changing the departure time of the flights. As another promising research direction, the presented model can be embedded in a competitive framework with a game theoretical approach in order to represent the response of each airline in the market segment.

## Acknowledgements

We are thankful to the reviewers for their valuable feedbacks which helped us to improve the paper considerably.

## References

- Atasoy, B. and Bierlaire, M. (2012). An air itinerary choice model based on a mixed RP/SP dataset. Technical Report TRANSP-OR 120426, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.
- Barnhart, C., Kniker, T. S., and Lohatepanont, M. (2002). Itinerary-based airline fleet assignment. *Transportation Science*, 36:199–217.
- Belobaba, P., Odoni, A., and Barnhart, C. (2009). *The Global Airline Industry*. John Wiley & Sons, Ltd, Chichester, UK, first edition.
- Bierlaire, M. and Fetiariison, M. (2009). Estimation of discrete choice models: extending BIOGEME. In *Swiss Transport Research Conference (STRC)*.
- Bonami, P., Biegler, L. T., Conn, A. R., Cornuéjols, G., Grossmann, I. E., Laird, C. D., Lee, J., Lodi, A., Margot, F., Sawaya, N., and Wächter, A. (2008). An algorithmic framework for convex mixed integer nonlinear programs. *Discrete Optimization*, 5(2):186 – 204.
- Brueckner, J. K. and Flores-Fillol, R. (2006). Airline schedule competition: Product-quality choice in a duopoly model. Working Papers 050629, University of California-Irvine, Department of Economics.
- Brueckner, J. K. and Zhang, Y. (2001). A model of scheduling in airline networks: How a hub-and-spoke system affects flight frequency, fares and welfare. *Journal of Transport Economics and Policy*, 35(2):195–222.
- Budhiraja, N., Pal, M. N., and Pal, A. K. (2006). Airline network design and fleet assignment: using logit-based dynamic demand-supply interaction. Working Paper, Indian Institute of Management Calcutta.
- Cadarsoa, L. and Marín, A. (2011). Integrated robust airline schedule development. *Procedia - Social and Behavioral Sciences*, 20:1041 – 1050.

- Carrier, E. (2008). *Modeling the choice of an airline itinerary and fare product using booking and seat availability data*. PhD thesis, Massachusetts Institute of Technology.
- Coldren, G. M. and Koppelman, F. S. (2005). Modeling the competition among air-travel itinerary shares: GEV model development. *Transportation Research Part A: Policy and Practice*, 39(4):345–365.
- Coldren, G. M., Koppelman, F. S., Kasturirangan, K., and Mukherjee, A. (2003). Modeling aggregate air-travel itinerary shares: logit model development at a major US airline. *Journal of Air Transport Management*, 9:361–369.
- Desaulniers, G., Desrosiers, J., Dumas, Y., Solomon, M. M., and Soumis, F. (1997). Daily aircraft routing and scheduling. *Management Science*, 43:841–855.
- Dumas, J., Aithnard, F., and Soumis, F. (2009). Improving the objective function of the fleet assignment problem. *Transportation Research Part B: Methodological*, 43(4):466–475.
- Garrow, L. A. (2010). *Discrete Choice Modelling and Air Travel Demand: Theory and Applications*. Ashgate Publishing: Aldershot, United Kingdom.
- Koppelman, F. S., Coldren, G. M., Kasturirangan, K., and Parker, R. A. (2008). Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B: Methodological*, 42:263–273.
- Lohatepanont, M. (2002). *Airline Fleet Assignment and Schedule Design: Integrated Models and Algorithms*. PhD thesis, Massachusetts Institute of Technology.
- Lohatepanont, M. and Barnhart, C. (2004). Airline schedule planning: Integrated models and algorithms for the schedule design and fleet assignment. *Transportation Science*, 38:19–32.
- Rexing, B., Barnhart, C., Kniker, T. S., Jarrah, T., and Krishnamurthy, N. (2000). Airline fleet assignment with time windows. *Transportation Science*, 34:1–20.
- Schön, C. (2006). Market-oriented airline service design. *Operations Research Proceedings*, pages 361–366.
- Schön, C. (2008). Integrated airline schedule design, fleet assignment and strategic pricing. In *Multikonferenz Wirtschaftsinformatik (MKWI)*.
- Sherali, H. D., Bish, E. K., and Zhu, X. (2006). Airline fleet assignment concepts, models, and algorithms. *European Journal of Operational Research*, 172(1):1–30.
- Talluri, K. T. and van Ryzin, G. J. (2004a). Revenue management under a general discrete choice model of customer behavior. *Management Science*, 50(1):15–33.
- Talluri, K. T. and van Ryzin, G. J. (2004b). *The Theory and Practice of Revenue Management*. Kluwer Academic Publishers, Boston, first edition.
- Vaze, V. and Barnhart, C. (2010). Competitive airline scheduling under airport demand management strategies. Working paper.

- Wen, C.-H. and Lai, S.-C. (2010). Latent class models of international air carrier choice. *Transportation Research Part E: Logistics and Transportation Review*, 46:211–221.
- Yan, S. and Tseng, C.-H. (2002). A passenger demand model for airline flight scheduling and fleet routing. *Computers and Operations Research*, 29:1559–1581.

## A Appendix: The logit model based on the SP data

As mentioned in section 2 the SP data is combined with the RP data in order to take the advantage of its elasticity. In this section we provide the logit model which is estimated on the SP data. The logit model has three alternatives. The first one is a nonstop itinerary. The second alternative is a one-stop itinerary both flights being operated by the same airline. The third alternative is also a one-stop itinerary where the connection is provided by another airline. The utilities for these alternatives are provided by the equations 21, 22, and 23 respectively.

Similar to the RP model, the parameters are specified as economy and business. The parameters of the price variables for each of the alternatives ( $\beta_p^{E,NS}$ ,  $\beta_p^{B,NS}$ ,  $\beta_p^{E,S}$ ,  $\beta_p^{B,S}$ ) are constrained to be the same as the price parameters of the RP model presented in section 2. Similarly the parameters of the time variables ( $\beta_{\text{time}}^{E,NS}$ ,  $\beta_{\text{time}}^{B,NS}$ ,  $\beta_{\text{time}}^{E,S}$ ,  $\beta_{\text{time}}^{B,S}$ ) and the parameters of the morning variables ( $\beta_{\text{morning}}^E$ ,  $\beta_{\text{morning}}^B$ ) are also designed to be the same as the parameters of the RP model. Since the models for RP and SP datasets are estimated simultaneously, we need to estimate a scale parameter,  $scale_{SP}$ , to ensure that their variances are the same.

In the SP model, there are additional explanatory variables since it is based on a rich data set. For business passengers we have the information whether they pay their ticket or their company pay for that. Therefore there is an additional dummy variable, *business/others-pay*, which is 1 if the business passenger’s ticket is not paid by himself. There are other explanatory variables which are represented by  $v$ . These variables include the legroom provided in the airplane, the delay of the flight in case of late or early arrival and the variable representing whether the passenger is a frequent flyer or not.

$$\begin{aligned}
V_1 = & \text{scale}_{SP} \times (\beta_p^{E,NS} \times \ln(p_1/100) \times \text{economy} + \beta_p^{B,NS} \times \ln(p_1/100) \times \text{business}) \\
& + \beta_p^{B-OP} \times \ln(p_1/100) \times \text{business/others-pay} \\
& + \beta_{\text{time}}^{E,NS} \times \text{time}_1 \times \text{economy} + \beta_{\text{time}}^{B,NS} \times \text{time}_1 \times \text{business} \\
& + \beta_{\text{morning}}^E \times \text{morning}_1 \times \text{economy} \\
& + \beta_{\text{morning}}^B \times \text{morning}_1 \times \text{business} \\
& + \sum_i \beta_i^E \times v_1^i \times \text{economy} + \beta_i^B \times v_1^i \times \text{business}
\end{aligned} \tag{21}$$

$$\begin{aligned}
V_2 = & \text{scale}_{SP} \times (\text{ASC}_2^E \times \text{economy} + \text{ASC}_2^B \times \text{business}) \\
& + \beta_p^{E,S} \times \ln(p_2/100) \times \text{economy} + \beta_p^{B,S} \times \ln(p_2/100) \times \text{business} \\
& + \beta_p^{B-OP} \times \ln(p_2/100) \times \text{business/others-pay} \\
& + \beta_{\text{time}}^{E,S} \times \text{time}_2 \times \text{economy} + \beta_{\text{time}}^{B,S} \times \text{time}_2 \times \text{business} \\
& + \beta_{\text{morning}}^E \times \text{morning}_2 \times \text{economy} \\
& + \beta_{\text{morning}}^B \times \text{morning}_2 \times \text{business} \\
& + \sum_i \beta_i^E \times v_2^i \times \text{economy} + \beta_i^B \times v_2^i \times \text{business}
\end{aligned} \tag{22}$$

$$\begin{aligned}
V_3 = & \text{scale}_{SP} \times (\text{ASC}_3^E \times \text{economy} + \text{ASC}_3^B \times \text{business}) \\
& + \beta_p^{E,S} \times \ln(p_3/100) \times \text{economy} + \beta_p^{B,S} \times \ln(p_3/100) \times \text{business} \\
& + \beta_p^{B-OP} \times \ln(p_3/100) \times \text{business/others-pay} \\
& + \beta_{\text{time}}^{E,S} \times \text{time}_3 \times \text{economy} + \beta_{\text{time}}^{B,S} \times \text{time}_3 \times \text{business} \\
& + \beta_{\text{morning}}^E \times \text{morning}_3 \times \text{economy} \\
& + \beta_{\text{morning}}^B \times \text{morning}_3 \times \text{business} \\
& + \sum_i \beta_i^E \times v_3^i \times \text{economy} + \beta_i^B \times v_3^i \times \text{business}
\end{aligned} \tag{23}$$