Integrated Airline Scheduling: Considering Competition Effects and the Entry of the High Speed Rail

Luis Cadarso[1], Vikrant Vaze[2], Cynthia Barnhart[3], Ángel Marín[4]

[1] Rey Juan Carlos University, Departamental III Building, Camino del Molino s/n, Fuenlabrada 28943, Spain

[2] Thayer School of Engineering, Dartmouth College, 14 Engineering Drive, Hanover, NH 03755

[3] Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139

[4] Universidad Politécnica de Madrid, Plaza Cardenal Cisneros 3, Madrid 28040, Spain

Abstract

Airlines and high speed rail are increasingly competing for passengers, especially in Europe and Asia. Competition between them affects the number of captured passengers and, therefore, revenues. We consider multi-modal competition between airlines (legacy and low-cost) and high speed rail. We develop a new approach that generates airline schedules using an integrated mixed integer, non-linear optimization model that captures the impacts of airlines' decisions on passenger demand. We estimate the demand associated with a given schedule using a nested logit model. We report our computational results on realistic problem instances of the Spanish airline IBERIA, and show that the actual airline schedules are found to be reasonably close to the schedules generated by our approach.

Next, we use this optimization modeling approach under multi-modal competition to evaluate multiple scenarios involving entry of high speed rail into new markets. We account for the possibility of demand stimulation as a result of the new services. We validate our approach using data from markets that had an entry by high speed rail in the past. The out-of-sample validation results show a close match between the predicted and observed solutions. Finally, we use our validated model to predict the impacts of future entry by high speed rail in new markets. Our results provide several interesting and useful insights into the schedule changes, fleet composition changes, and fare changes that will help the airline cope effectively with the entry of high speed rail.

Keywords: airline scheduling, high speed rail, competition, integration, nested logit, demand stimulation.

1. Introduction

During the last few decades in Europe, High Speed Rail (HSR) has become an important competitor to airlines, presenting railway transport in a new form and notably improving the quality of service offered (Behrens and Pels, 2012; and Roman et al., 2007). An airline's ability to cope with HSR entry depends on a number of factors. Airline profitability is critically influenced by the airline's ability to: 1) estimate passenger demands; 2) construct profitable flight schedules (referred to as the airline schedule planning process); and 3) determine the fare levels for a set of products in an origin-destination market (referred to as the pricing process) and determine how many seats to make available at each fare level (referred to as revenue management). Because an investment in HSR often generates a redistribution of passengers between air and rail alternatives, varying the existing modal distribution, the impact of this investment on scheduled air transportation is quite uncertain and of great interest. This paper focuses on an integrated approach for airline passenger demand estimation and flight schedules construction when air and HSR modes compete.

1.1 Demand estimation

Transportation demand models are used to develop forecasts of passenger demand for each origin-destination (OD) pair (or market) as a function of attributes such as average fares, service frequencies, market demographics and economic conditions (Garrow, 2012). Given these total demand estimates, passenger choice models are used to estimate for each airline competitor and each market, the proportion of demand (or the share of market) it captures in that market, considering market-specific characteristics, including passengers' mode preferences and airline preferences, fares, flight frequencies and other market attributes.

There is widespread acceptance in the airline industry of an S-curve relationship between airline market share and frequency share. The S-curve describes how an airline's market share grows non-linearly with its frequency share in that market. Early theoretical development and empirical evidence that higher-frequency shares are associated with disproportionately higher market shares was provided in the 1970s before deregulation (Simpson, 1970). After deregulation, there exist a number of references to the S-curve in the aviation literature (Wei and Hansen, 2005; Belobaba, 2009; Vaze, 2011; and Vaze and Barnhart, 2012a). The most commonly used mathematical expression for the S-curve relationship is given by:

$$P_a^w = \frac{\left[f_a^w \right]^\alpha}{\sum_{a' \in A} \left[f_{a'}^w \right]^\alpha}, \tag{1}$$

where A is the set of airlines, P_a^w is the probability that a passenger in market w selects airline $a \in A$ among all the airlines, f_a^w is the frequency value of airline $a \in A$ in market w and $\alpha \geq 1$ is the frequency parameter.

1.2 Schedule planning

The schedule planning process typically starts from an existing schedule with a well-developed route structure and fleet composition. In constructing each new schedule, changes are introduced to the existing schedule to reflect changes in demands and the market environment. Due to the enormous size and complexity of the problem, schedule planning is a multi-step process, usually separated into four, sequentially solved subproblems: schedule design, fleet assignment, maintenance routing, and crew scheduling (Barnhart and Cohn, 2004; and Barnhart, 2009).

Schedule Design. Given available resources, the objective of schedule design is to develop a profit-maximizing schedule, defining an origin, a destination, a departure time, and an arrival time for each flight leg. It is a critical stage of an airline's planning process, as a major proportion of costs and revenues are fixed once a flight schedule is determined. Schedule design is typically composed of two sequential steps: frequency planning and timetable development (Lohatepanont and Barnhart, 2004).

Frequency Planning is the problem of determining the number of flight departures over a specified time period (e.g., over a week) for each route. This is a critical factor in an airlines' ability to compete for passengers, whose flight selection is primarily influenced by the frequency of flights, timetable, fares and quality of service. Increasing flight frequencies reduces wait time between flights, improves the convenience of air travel for passengers and increases the airline's market share relative to that of its competitors (Belobaba, 2009). This last effect results in an airline's market share being dependant not only on its own service but also on the services provided by other airlines in the market.

Timetable Development involves constructing a flight schedule that matches the frequencies determined in solving the frequency planning problem. Numerous factors must be considered in generating a timetable, including the trade-off between maximization of aircraft utilization (block hours per day) and schedule convenience for the passengers; minimum 'turnaround' times at each airport to deplane and enplane passengers, refuel and clean aircraft; and convenient passenger connections at hub airports (Belobaba, 2009).

Fleet Assignment. The fleet assignment problem is to determine the type of aircraft to be flown on each flight, given a planned flight network and specified timetable. The solution has a tremendous impact on an airline's profits, as it directly affects flight operating costs and passenger revenues. Important factors in assigning fleet types to flights include passenger demand, aircraft seating capacities, aircraft operating costs, and fleet size and composition (Hane et al., 1995).

Aircraft Routing. Given the assignment of fleet types to flights, the airline next determines the sequence of flights to be flown by each aircraft. The solution must ensure that each flight is flown exactly once, each aircraft visits maintenance stations at regular intervals, and only available aircraft of each type are utilized in the solution (Desaulniers et al., 1997).

Crew Scheduling. The goal of crew scheduling is to identify cost-minimizing crew schedules that provide the necessary crews for each flight, while satisfying the myriad of constraints imposed by government and labor work rules. A survey of optimization approaches for crew scheduling is provided in Barnhart et al. (2003).

For tractability purposes, schedule planning models typically consider demand for an airline's flights to be deterministic and invariant to schedule changes and competition. These assumptions, however, have been shown to lead to overestimates of the number of passengers served, the revenue captured, and schedule profitability (Yan et al., 2007; and Belobaba, 2009). An effective schedule planning process for an airline, then, depends critically on both the accurate estimation of the overall demand for travel in each market; and the accurate understanding of how passengers will choose between the airline's and its competitors' travel options in that market.

1.3 Schedule planning and demand estimation

Several researchers have developed enhanced schedule planning models to more accurately capture passenger demand for an airline's schedule and the resulting revenues (for example, Barnhart et al., 2002; Jacobs et al., 2008; Dumas et al., 2009; Barnhart et al., 2009; and Cadarso and Marín, 2013). These enhanced models, however, do not account for changes in passenger demand resulting from competition. Studies that have considered the impact of competition on travel demand include, for example, Hansen (1990), Hong and Harker (1992), Yan et al. (2007), Wei and Hansen (2007), Pita et al. (2012), and Vaze and Barnhart (2012a). These studies, however, do not consider multi-modal competition and do not differentiate between different types of airlines, i.e., Legacy Airlines (LAs) and Low-Cost Airlines (LCAs). LAs and LCAs provide different service levels, and hence, the resulting demand patterns are different. Previous econometric studies have investigated multi-modal competition using logit models to estimate the demand associated with a schedule (for example, Behrens and Pels, 2012; and Roman et al., 2007). However, there is limited research involving the integration of these types of multi-modal logit models with schedule design models. Zito el al. (2011) model how airlines make decisions on fares and frequencies of service in a multi-modal competitive environment, but competition between different types of airlines (such as that between legacy airlines and low-cost airlines) is not considered, schedule development is not modeled at the level of detail of schedule design models, and they do not apply their approach to real-world instances. Wang et al. (2012) present a new framework that incorporates the spill and recapture effects. The concept is derived from the classical multinomial logit model. Their preliminary computational study shows an increase in profitability and a better utilization of the network capacity, on a medium-size North American airline.

Our work differs from others in that it considers multi-modal competition including airline (legacy and low-cost) and high speed rail, and develops a new approach that: 1) estimates the demand associated with a given schedule using a nested logit model; and 2) generates airline schedules and fleet assignments using an integrated schedule design and fleet assignment optimization model that accounts for airline competition for passenger demand, and captures the impacts of schedule decisions on passenger demand, as suggested by the S-curve relationship. Our model, reflecting that passenger demand for an airline schedule depends not only on the airline's schedule but also on the schedule of its competitors, captures linkages between schedule competition and passenger demand. It is therefore able to drive profit maximization with improved estimates of revenues. Pricing and revenue management decisions, however, are out of the scope of our model, that is, the model uses average ticket fares as inputs.

The motivation for considering multi-modal competition stems from the fact that High Speed Rail (HSR) and airlines are increasingly competing for passengers in many parts of Europe and Asia, especially in short- to medium-haul markets. HSR often competes by providing similar or even greater service frequency and better connectivity to the city centers. Moreover, HSR is often perceived as the safer and more comfortable mode (Jehanno et al., 2011). In addition to modeling competition between air and rail modes, we model competition between legacy and low-cost airlines. These are perceived as different choices because level of service is different. Usually, legacy airlines include the following services: first class and/or business class, a frequent-flyer program, airport lounges, alliance partners that agree to provide these services to the passengers as well, etc. These services are often not associated with low-cost airlines.

1.4 Contributions

In this paper, we present a mixed integer, non-linear programming model for the schedule design and fleet assignment problem that includes a passenger choice model to capture multi-modal competition between high speed rail, low-cost airlines, and legacy airlines.

Our major contributions include:

1. Development of a tactical competition model for an airline — considering multi-modal competition between air and high speed rail, and airline competition

between legacy and low-cost carriers — using a nested logit model of demand behavior. We calibrate our model using real data.

2. Development of an integrated mixed integer, non-linear programming model for schedule optimization that includes frequency planning, approximate timetable development, fleet assignment and passenger demand choice. We solve this model using realistic problem instances obtained from the network of the Spanish airline IBERIA. Our instances also include other air and rail transportation options in Spain. We also perform sensitivity analysis on model parameters. Our experimental results show that the schedules generated by our approach are found to closely resemble the current decisions made by IBERIA with a reasonable level of accuracy. This indicates that the current decision-making by IBERIA does take into account the multi-modal competition aspects.
3. We evaluate multiple scenarios involving entry of HSR into new markets, and we account for the possibility of demand stimulation as a result of the new services. We validate our approach using data from markets that had an entry by high speed rail in the past. The out-of-sample validation results show a close match between the predicted and observed solutions. Finally, we use our validated model to predict the impacts of future entry by high speed rail in new markets. Our results provide several interesting and useful insights into the schedule changes, fleet composition changes, and fare changes that will help the airline cope effectively with the entry of high speed rail.

1.5 Key definitions

The remainder of the paper uses some important terms which are defined as follows.

- Tactical planning: resource assignment decision-making over a time horizon of several months prior to day of operations.
- Origin-Destination (OD) pair: combination of origin airport and destination airport.
- Route: a physical path through the nodes of the air network (defined by fixed geographic coordinates) linking an origin and a destination. We assume that an airline offers a unique route, which may be one-stop or non-stop, for every OD pair. This assumption holds true for our case study network.
- Market: combination of origin airport, destination airport and a desired departure time period.
- Flight leg: combination of a departure airport, a departure time period, an arrival airport and an arrival time period (i.e., a non-stop flight).
- Itinerary: sequence of one or more flight legs traversed by a passenger from the passenger's origin to the passenger's destination. Note that multiple itineraries may serve the same OD pair (within the same route). We assume that itineraries

are composed of at most two flight legs. This holds for all the itineraries in our case study. In contrast to a route which is a path in the spatial network of an airline, an itinerary is a path in the time-space network of an airline.

1.6 Outline of the paper

The remainder of this paper is organized as follows. In Sections 2 and 3, respectively, we present our demand modeling approach and our formulation for schedule optimization under multi-modal competition. In Section 4, computational experiments for a real-world problem using data provided by a legacy airline in Spain (IBERIA) are described, and the results are presented. In Section 5, we evaluate multiple scenarios involving entry of HSR into new markets. We conclude in Section 6 with a discussion of major findings.

2. Demand modeling

Service frequency is one of the most important attributes on which the airlines compete. An airline can attract more passengers in a market by increasing the service frequency. For a given unconstrained total demand, the market share of each airline depends, among other factors, on its own frequency and on the frequency of its competitors.

However, modeling the market share as simply a function of the frequency share is not enough to model passenger demand behavior in many markets. This is especially true in markets where the competitor fares are different from each other and the competing airlines are different from the perspectives of the passengers in other ways (Vaze and Barnhart, 2012a). There are other attributes, such as fares and travel times that can significantly affect passengers' airline choice (Behrens and Pels, 2012; and Roman et al. 2007). For instance, consider an Origin-Destination (OD) pair served by two different airlines: the first one operates a non-stop flight and the second one a one-stop flight. Passengers will likely prefer the non-stop flight over the one-stop flight, all else being equal, because non-stop flights typically have less travel time and are more convenient compared with one-stop flights.

2.1 Nested logit model development

Consequently, we extend the model in (1) in order to include fare and travel time as attributes. Many past studies have modeled market share as a function of the logarithms of the attributes such as frequency, price and travel time (Wei and Hansen, 2005; and Vaze and Barnhart, 2012a). The relationship in (1) may be extended and rewritten as a discrete choice multinomial logit model (see equation 2), where the choice probability is proportional to the exponential of the systematic utility ($v(a \mid w)$) of each airline a in market w.

$$v(a \mid w) = \log asc_a + \alpha \log f_a^w + \beta \log p_a^w + \gamma \log tt_a^w$$

$$P_a^w = \frac{e^{v(a|w)}}{\sum_{a' \in A} e^{v(a'|w)}}, \quad (2)$$

where p_a^w is the fare of airline $a \in A$ in market w, tt_a^w is the planned travel time of airline $a \in A$ in market w (for non-stop itineraries it is the flight time and for one-stop itineraries it is the sum of the two flight times and the average value of passenger connecting time), β is the fare parameter and γ is the flight time parameter. In addition, there could be other airline-specific factors that impact the passenger share. For example, some passengers might have a preference for legacy carriers over low-cost carriers, or some passengers might prefer one airline over the other due to frequent flyer program memberships, etc. In order to capture these factors, we include alternative specific constants asc_a for each airline $a \in A$.

In general, all passengers prefer lower fare and higher frequency. However, some passengers might value lower fare more than other passengers do, while others might give more importance to higher frequency. To incorporate these effects, we propose an extension of (2). Let Z be the set of passenger types and $\zeta \in Z$ represent a particular passenger type, e.g., business or leisure. Let H_w^ζ be the probability for an individual in market w of belonging to passenger type ζ, which is usually modeled with a logit expression as well, where the exponent is denoted by $\eta(\zeta \mid w)$ (Greene and Hensher, 2003; and Wen and Lai, 2010). Let $P_a^{w|\zeta}$ be the probability that a passenger of type ζ in market w selects airline $a \in A$ among all the airlines. Now the systematic utility of each airline a in market w for passenger type ζ is also conditioned on ζ and is denoted by $v(a \mid w, \zeta)$. Consequently, the joint probability for a passenger in market w of belonging to type ζ and selecting airline $a \in A$ ($P_a^{w,\zeta}$) among all the airlines will be as follows:

$$P_a^{w,\zeta} = H_w^\zeta P_a^{w|\zeta} = \frac{e^{\eta(\zeta|w)}}{\sum_{\zeta' \in Z} e^{\eta(\zeta'|w)}} \frac{e^{v(a|w,\zeta)}}{\sum_{a' \in A} e^{v(a'|w,\zeta)}}. \quad (3)$$

Due to lack of disaggregate data on passenger attributes (such as individual socioeconomic and trip characteristics) in our case studies, we assume that the exponent in the logit expression for the probability for a passenger of belonging to type ζ in market w is $\eta(\zeta \mid w) = \upsilon_\zeta + \mu_\zeta^b \delta_w^b + \mu_\zeta^l \delta_w^l + \mu_\zeta^d \bar{d}_w + \mu_\zeta^s \delta_w^s$, where υ_ζ is the alternative specific constant for each passenger type ζ, δ_w^b is a 0/1 dummy variable indicating whether market w is business-dominated (μ_ζ^b is the corresponding coefficient), δ_w^l is a 0/1 dummy variable indicating whether market w is tourism-dominated (μ_ζ^l is the corresponding coefficient), $\bar{d}_w$ is the average distance travelled (which depends on the route) in market w (μ_ζ^d is the distance coefficient) and δ_w^s is a 0/1 dummy variable indicating whether market w is served by a one-stop route (as against a non-stop route) (μ_ζ^s is the corresponding coefficient).

$v(a\,|\,w,\zeta) = \log asc_a^\zeta + \alpha^\zeta \log f_a^w + \beta^\zeta \log p_a^{w,\zeta} + \gamma^\zeta \log tt_a^w$ is the systematic component of the utility of an itinerary on airline a in market w for passenger type ζ, where α^ζ is the frequency parameter, $p_a^{w,\zeta}$ is the fare of airline $a \in A$ in market w for passenger type ζ, β^ζ is the price parameter, γ^ζ is the flight time parameter and asc_a^ζ is the alternative specific constant for each airline for each passenger type ζ. Therefore, the choice probability for airline a consists of summation (across different passenger types) of the products of two terms: the choice probability for the passenger type ζ ($P_a^{w|\zeta}$) and the probability for a passenger in market w of belonging to passenger type ζ (H_w^ζ).

Sometimes, HSR is considered the best transport mode for short distance trips, providing shorter travel times between cities, higher quality of service and reduced access times to city centers. We, therefore, introduce a modal choice model. Passengers decide between High Speed Rail (HSR) and air modes for the origin-destination pairs (OD pairs) where both are competing. We assume that each operator offers a unique route for each OD pair (this holds for 94.23% of the ODs in our case study). We also assume that there are no connecting flights in the OD pairs where the HSR operates (this holds for all the ODs where the HSR operates in our case study). Consequently, all airlines that serve the OD pairs where the HSR operates have the same flight time and the attribute tt_a^w is redundant and hence, it is not included in those specific cases within the systematic utility of the airline. However, the HSR and air modes are differentiated by modal travel times, tt_{rail}^w and tt_{air}^w (these times also include access and egress times), respectively. The passenger choice of a specific alternative in each OD pair can be modeled using the nested logit model. We have two levels in the nested logit model: mode (air vs. rail) level at top and operator level at bottom. At the top level, the passenger chooses from different alternative modes, including air and rail. Note that we have a unique rail operator. Therefore, if the passenger chooses rail mode then there is no more choice necessary. On the other hand, if the passenger chooses air mode, then the decision in the lower level of the nested logit model involves choice between the various competing airlines (equation 3). We model passengers' modal choice for each market and passenger type in equation 4:

$$P_{air}^{w|\zeta} = \frac{e^{\theta_{air} \log \sum_{a \in A} e^{v(a|w,\zeta)} + \gamma_{air} \log tt_{air}^{w}}}{e^{\theta_{air} \log \sum_{a \in A} e^{v(a|w,\zeta)} + \gamma_{air} \log tt_{air}^{w}} + e^{v(rail|w)} + e^{v(null|w)}}, \tag{4}$$

where $\theta_{air} \in [0,1]$ is a measure of correlation and substitution among alternatives in the nest. The lower the correlation between unobserved effects in the utilities of different air operators, the higher is the θ_{air} value. γ_{air} is the air travel time parameter. $v(rail\,|\,w) = \log asc_{rail} + \alpha_{rail} \log f_{rail}^{w} + \beta_{rail} \log p_{rail}^{w} + \gamma_{rail} \log tt_{rail}^{w}$ is the systematic utility of the alternative $rail$ for each market w. asc_{rail} is the rail alternative specific constant. f_{rail}^{w} is the rail service frequency and α_{rail} is the frequency parameter for the rail mode. p_{rail}^{w} is the rail mode's fare and β_{rail} is the fare parameter for the rail mode. Finally, γ_{rail} is the rail travel time parameter. $v(null\,|\,w) = \log asc_{null}$ is the systematic utility of the *null* alternative, which we define as the decision to travel by neither air nor rail in a market w. Such decision can include traveling by any other mode of transportation or not traveling at all. We treat the null alternative as a single alternative with no observable attributes except for an alternative specific constant due to lack of data. Note that, in general, the distance in each market could be considered to be an attribute. However, since the data for the case study considered in this paper does not exhibit significant variance in distance, we decided to omit this variable. The probability for a passenger in market w of belonging to type ζ and selecting airline a is: $P_{a}^{w,\zeta} = H_{w}^{\zeta} P_{a|air}^{w|\zeta} P_{air}^{w|\zeta}$, where $P_{a|air}^{w|\zeta}$ is the same as the expression for $P_{a}^{w|\zeta}$ in (3).

Note that at the mode choice level we are not differentiating between different passenger types. So the passenger type ζ appears only in the *log-sum* term in (4). This is consistent with our previously stated aim of investigating the differences in the relative preferences of the different passenger types for airfare and airline frequency. This modeling choice is made in order to strike the right balance between additional modeling complexity and the ability to get interesting insights using the available data. Note that one can extend the presented methodology in a number of ways, namely, inclusion of passenger types at the mode choice level, an increase in the number of passenger types at the airline choice level, etc.

2.2 Parameter estimation

We have estimated the nested logit model parameters using real data from year 2010 provided by IBERIA. This data was obtained by IBERIA from Amadeus Market Information Data Tapes (MIDT). This MIDT data includes data collected from a multitude of travel providers such as airlines, car rental companies, railway companies, ferry lines, cruise lines, and travel agencies. The data is composed of 104 different origin-destination pairs and 7 different periods of time where each period is a week. Consequently, there are 728 different sets of data, one corresponding to each week. Each set is composed of the data of the airline under study (IBERIA) and its competitors. The total number of rows in our dataset equals 30547, one corresponding to each market. Data includes market type (namely, business- or leisure-dominated), average distance travelled in the market, market service type (namely, one-stop vs. non-stop), overall demand, market share, frequency, travel time, and price (the last four by operator). The demand data from MIDT is constrained demand data in the sense that it includes sales that actually happened, rather than reflecting the total demand. So the sales that didn't happen don't get included. As a result, the total unconstrained demand is likely to exceed the total sales. We use a pragmatic approach for unconstraining demands. Our approach is backed by prior literature and is designed to make the best use of the limited availability of unconstrained demand data. We have available data on bookings and load factors for IBERIA flights. Therefore, we can use the approach in Wickham (1995), which is still widely used in practice (Guo et al., 2012), to find unconstrained demand data.

This approach consists of identifying departures in each market that are not constrained (i.e., those flights with load factor lower than 0.7). For these departures, the percentage of the bookings held at a given day prior to departure relative to the bookings held at departure day is computed. Then, for the constrained departures, the number of bookings held prior to departure is divided by this percentage to get an estimate of the total unconstrained demand for those departures. Due to data limitations, we have the information on the number of bookings held at a given day prior to departure only for a subset of all the markets within IBERIA network. Therefore, we use the following pragmatic approach that extends the approach proposed by Wickham (1995):

1. We first compute the percentage of the bookings held one week prior to departure relative to the bookings held at departure day for the departures that are not constrained and are in the subset for which we have the data on the prior number of bookings (these constitute 8.31% of the departures in the dataset).
2. Then, we compute the average of the percentages computed in step 1 and apply it to the constrained departures for which we have the data on the number of bookings one week prior to departure (these constitute 15.03% of the departures in the dataset), in order to obtain the unconstrained demand values for those departures. This is consistent with the approach proposed by Wickham (1995).

3. Next, we compute the ratio between the unconstrained and constrained demand values for the available constrained departures in step 2 and obtain its average value.
4. This average ratio is then applied to the rest of the constrained departures within IBERIA network to get the respective unconstrained demands (these constitute 45.81% of the departures in the dataset).
5. Because we do not have data on bookings held one week prior to departure for the rest of the airlines and the HSR, the average of the ratios computed from IBERIA network is used for this purpose. The HSR also uses revenue management techniques and the average load factor usually lies between 0.7 and 0.85 (Ferropedia, 2014). Therefore, we assume the demand unconstraining techniques in the airline industry are also valid for the HSR.

Note that this extended approach was developed with the aim of reusing a proven demand unconstraining method from past literature while making the best use of the limited data available at our disposal.

In this case study, there are two main types of competing airlines: low-cost and legacy airlines. There are three low-cost airlines (Vueling, Ryanair and easyJet) and three legacy airlines (IBERIA, Air Europa and Spanair). The rail competition is provided by a unique rail operator (RENFE) and it is not present in all the OD pairs. The HSR is present in the following OD pairs: Madrid-Barcelona, Barcelona-Madrid, Madrid-Seville, Seville-Madrid, Madrid-Malaga and Malaga-Madrid. Table 1 presents a data summary (the average and range) of overall unconstrained demand (total number of passengers interested in making a trip in each market); constrained market share (percentage of reservations made for each mode and/or airline); frequency (number of operated services by each operator per time period); travel time (service time in hours between each OD pair); price (average fare paid by each passenger in each OD pair). Note that each row in Table 1 presents the summary of one column in our dataset.

Table 1: Nested logit parameter estimation data summary

Type of Data	Average	Data range
Overall unconstrained demand	132.7	[1.0, 1086.1]
Market share	32.5%	[0.01%, 100%]
Frequency (per week)	27.6	[6, 144]
Travel time (hours)	2.42	[1.25, 5.76]

Price (Euro)	95.4	[32.6, 315.7]

The parameters to be estimated are: α^{ζ}, β^{ζ}, γ^{ζ}, asc_a^{ζ}, θ_{air}, asc_{rail}, γ_{air}, γ_{rail}, α_{rail}, β_{rail}, asc_{null}, $\upsilon_{\zeta}, \mu_{\zeta}^{b}$, μ_{ζ}^{l}, μ_{ζ}^{d} and μ_{ζ}^{s}. We assume two passenger types, indexed by $\zeta \in Z$ with ζ equal to 1 representing business passengers and ζ equal to 2 representing leisure passengers. As mentioned earlier, one can easily extend the presented methodology to include more passenger types. The alternative specific constant, asc_a^{ζ}, models the passenger perception of legacy and low-cost airlines. Thus, it takes value 1 if the airline is a legacy airline (without loss of generality) and if the airline is a low-cost airline then the alternative specific constant is a parameter to be estimated (see estimates in Table 2). Also, for identification, the parameters ($\left\{\upsilon_{\zeta}, \mu_{\zeta}^{b}, \mu_{\zeta}^{l}, \mu_{\zeta}^{d}, \mu_{\zeta}^{s}\right\}^{T}$) in the passenger type model's exponent are normalized to zero for one passenger type ($\zeta = 1$).

We have used the maximum likelihood estimation method in order to estimate the nested logit model parameters. The log-likelihood for all the observations is: $\log L(\sigma \mid x_1,...,x_n) = \sum_{i=1}^{n} \iota_{i,a}^{w} \log\left(\sum_{\zeta \in Z}\left(H_w^{\zeta}(x_i \mid \sigma) P_{a|air}^{w|\zeta}(x_i \mid \sigma) P_{air}^{w|\zeta}(x_i \mid \sigma)\right)\right) + \iota_{i,rail}^{w} \log\left(P_{rail}^{w}(x_i \mid \sigma)\right) + \iota_{i,null}^{w} \log\left(P_{null}^{w}(x_i \mid \sigma)\right)$, where $\iota_{i,a}^{w}, \iota_{i,rail}^{w}, \iota_{i,null}^{w}$ are indicator parameters (i.e., it is 1 if observation i is related to airline a, rail mode or the null alternative, respectively, for market w), $\sigma = \left\{\alpha^{\zeta}, \beta^{\zeta}, \gamma^{\zeta}, asc_a^{\zeta}, \theta_{air}, asc_{rail}, \gamma_{air}, \gamma_{rail}, \alpha_{rail}, \beta_{rail}, asc_{null}, \upsilon_{\zeta}, \mu_{\zeta}^{b}, \mu_{\zeta}^{l}, \mu_{\zeta}^{d}, \mu_{\zeta}^{s}\right\}^{T}$ are the parameters to be estimated, and $x_i = \left\{f_a^{w}, p_a^{w,\zeta}, tt_a^{w}, tt_{air}^{w}, f_{rail}^{w}, p_{rail}^{w}, tt_{rail}^{w}, \delta_w^{b}, \delta_w^{l}, \bar{d}_w, \delta_w^{s}\right\}_i^{T}$ is the data vector for the i th observation. The method of maximum likelihood estimation provides estimates of σ by finding a value of σ that maximizes $\log L(\sigma \mid x_1,...,x_n)$: $\hat{\sigma} \in \left\{\underset{\sigma \in \mathrm{P}}{\arg\max} \log L(\sigma \mid x_1,...,x_n)\right\}$, where P is the set of values of σ over which the likelihood maximization is conducted.

The Newton-Raphson method is used to maximize the log-likelihood function with respect to the parameter vector σ. Almost all of the parameter estimates were found to be significant at the 95% confidence level using the classic Student's t-test. Table 2 shows the test results: the parameters are listed in the first column, the estimates in the second column, standard errors in the third column and the p-values in the last column. Because the problem is likely a non-convex optimization problem, there is a danger of the algorithm getting stuck in local maxima. Therefore, initial values could affect the results of likelihood maximization. In order to overcome this potential issue, we tried multiple initial values for the Newton-Raphson method, and picked the overall best solution obtained from these multiple tries. We found that the overall optimal solution results in intuitively reasonable relative values ($\alpha^1 > \alpha^2$, $\beta^1 > \beta^2$, and $\gamma^1 < \gamma^2$) of parameter estimates without enforcing such constraints explicitly. This was not found to be the case for some of the local maxima obtained in the process.

We use the likelihood ratio test statistic, following a Chi-squared distribution, with degrees of freedom equal to the number of extra estimated parameters, to test overall model significance. This test statistic equals 1831.21. Using the critical value of 25.00 at 5% significance level (for 15 extra parameter), we reject the null hypothesis that the naïve model (i.e., a model that contains no explanatory variables; it only contains the five alternative specific constants: asc_a^1, asc_a^2, υ_2, asc_{rail} and asc_{null}) is as good as the nested logit model. We also checked that the frequency, fare and travel time parameters are significantly different across passenger types. We performed the classical t-test of the difference between α^1 and α^2, β^1 and β^2, and γ^1 and γ^2. The p-values are 0.029, 0.047 and 0.053, respectively indicating that, for each of the three parameters, with approximately 95% or higher confidence we can reject the null hypothesis that the parameter value is equal across the two passenger types.

Table 2: Nested logit parameter estimation results

Parameter	Estimate	Standard Error	p-Value
α^1	1.5897	0.071	0.032
β^1	-0.5639	0.003	0.026
γ^1	-1.5801	0.204	0.019
asc_a^1	0.8239	0.281	0.027
α^2	1.1601	0.107	0.031
β^2	-0.8672	0.044	0.023
γ^2	-1.3195	0.156	0.014
asc_a^2	1.1385	0.358	0.044
υ_2	0.6106	0.109	0.052
μ_2^b	-0.2587	0.081	0.043
μ_2^l	0.8512	0.066	0.048
μ_2^d	0.0108	0.175	0.046

μ^{s}_{2}	0.0783	0.108	0.033
θ_{air}	0.7230	0.092	0.012
γ_{air}	-0.8028	0.131	0.030
α_{rail}	0.9541	0.122	0.016
β_{rail}	-0.4125	0.286	0.022
γ_{rail}	-0.8963	0.301	0.037
asc_{rail}	1.3574	0.127	0.021
asc_{null}	0.3953	0.027	0.053

The estimated inclusive value (i.e., $\log \sum_{a \in A} e^{v(a|w,\zeta)}$) parameter of the air nest (θ_{air}) is statistically significantly less than 1 (at 5% significance level). The corresponding likelihood ratio-test statistic equals 36.27, much greater than the critical value of 3.84 at 5% significance level (for one extra parameter). We thus reject the null hypothesis that the model with $\theta_{air} = 1$ is as good as our latent class nested logit model. $\theta_{air} = 1$ would mean that there is no correlation within the air nest and that the latent class nested logit model is equivalent to a latent class multinomial logit model. The interpretation of this result is that, for the presented case study, air alternatives have some correlation in the unobserved effects, but they are competing to some degree.

From the alternative specific constant for the rail mode, we can conclude that, all else being equal, passengers prefer the rail mode rather than the air mode when competing in the same origin-destination pair ($asc_{rail} > 1$). This is mainly due to unobserved effects such as the fact that railway security checks are less onerous for the passengers, and rail passenger travel is more comfortable and convenient (allowing the use of electronic devices throughout the journey, etc.).

From the estimation results we conclude that passengers' value of time is similar for different transport modes ($\gamma_{rail}, \gamma_{air}$). At first look, the parameter estimates might seem to suggest that passengers value railway frequency less than airline frequency ($\alpha_{rail} < 1$ and α^1, α^2 > 1). However, if we correct the scale of the parameters inside the nest to match the scale outside the nest, then the α_{rail} (0.9541) value is between $\alpha^1\theta_{air}$ (1.1493) and $\alpha^2\theta_{air}$ (0.8387), suggesting that passengers value railway frequency and airline frequency in roughly the same way. Note that α_{rail} parameter is used for a mixture of business and leisure travelers. The same holds for HSR passengers' sensitivity to prices, i.e., the β_{rail} value (-0.4125) is between the corrected $\beta^1\theta_{air}$ (-0.4076) and $\beta^2\theta_{air}$ (-0.6269) values.

The results show that, all else being equal, business passengers prefer legacy airlines over low-cost airlines ($asc_a^1 < 1$), while leisure passengers prefer low-cost airlines ($asc_a^2 > 1$). This suggests that in addition to the attributes that we explicitly captured in our model (that is, fare, frequency and travel time) there are other attributes, such as differences in airline images, loyalty programs etc., that make low-cost airlines relatively more attractive to leisure passengers and legacy airlines relatively more attractive to business passengers.

Passengers belonging to the business category are more sensitive to travel time than passengers belonging to the leisure category ($\gamma^1 < \gamma^2$), while business passengers are less sensitive to higher fares than leisure passengers ($\beta^1 > \beta^2$). Both these observations make intuitive sense. In addition, business passengers place more value than leisure passengers on higher frequencies ($\alpha^1 > \alpha^2$), presumably because higher frequency reduces effective travel time, which includes schedule displacement (Belobaba, 2009). Schedule displacement corresponds to time elapsed between when a passenger wants to travel and when service is offered. Thus, effective travel time and schedule displacement both change inversely with frequency.

Parameters in the passenger type model's exponent, i.e. $\left\{\upsilon_\zeta, \mu_\zeta^b, \mu_\zeta^l, \mu_\zeta^d, \mu_\zeta^s\right\}^T$, are normalized to zero for business passengers ($\zeta = 1$). For passengers belonging to the leisure category, the estimated values make intuitive sense. The business-dominated market coefficient for leisure passengers is lower than zero ($\mu_2^b < 0$), which means that, all else being equal, business-dominated markets have more business passengers than leisure passengers. The contrary holds for the tourism-dominated markets coefficient ($\mu_2^l > 0$). Finally the distance coefficient and the one-stop route coefficient are both slightly, but statistically significantly, greater than zero ($\mu_2^d > 0$, $\mu_2^s > 0$) implying that, all else being equal, there are slightly more leisure passengers than business passengers in markets with greater flight distances and in markets served by one-stop routes.

The captured demand Δ_a^w for airline a in market w is given by (5), where d_w is the unconstrained total demand for market w. Since we are studying a tactical problem for which the timetable is approximate, the OD attributes are substituted in the nested logit model for the frequency parameter, i.e., f_a^w is substituted by f_a^{od}, od being the OD pair of market w ($f_a^{od} = \sum_{w \in W_{od}} f_a^w$, where W_{od} is the set of markets in OD pair od).

$$\Delta_a^w = \sum_{\zeta \in Z} \left[\frac{e^{\eta(\zeta|w)}}{\sum_{\zeta' \in Z} e^{\eta(\zeta'|w)}} \frac{e^{v(a|w,\zeta)}}{\sum_{a' \in A} e^{v(a'|w,\zeta)}} \frac{e^{\theta_{air} \log \sum_{a' \in A} e^{v(a'|w,\zeta)} + \gamma_{air} \log\left(tt_{air}^w\right)}}{e^{\theta_{air} \log \sum_{a' \in A} e^{v(a'|w,\zeta)} + \gamma_{air} \log\left(tt_{air}^w\right)} + e^{v(rail|w)} + e^{v(null|w)}} \right] d_w$$

(5)

3. Optimization model formulation

Vaze and Barnhart (2012b) present a three-stage modeling framework for the airline planning process. The first stage deals with network design where decisions about the number and location of hubs, candidates for non-stop routes and allowable airports for passengers' connections are made. The second stage involves the frequency planning and fleet assignment problem, and the third stage addresses the timetable development and fleet balancing problem. The model presented in this paper combines the decisions in the second and third stages of this framework since the network design decisions in the first stage are beyond the scope of this paper. The integrated model addresses frequency planning, fleet assignment, and development of an approximate timetable (i.e., number of departures for each OD pair and for each time period). Since we are proposing a tactical competition model, we assume that fares are fixed at an average value and will not be part of the decision variables. We formulate the model as an optimization problem from an airline's point of view.

Air Network. The air network is formed by the airports and all the feasible flight legs linking them. The airports are characterized by the number of operations that can be performed at those airports. We assume that arrival and departure slot availability is known for each airport for each airline, and airlines offer the links (flight opportunities) between the airports.

In order to ensure tractability of the problem, we propose an aggregated network, similar to that presented in Harsha (2008). It allows different levels of time-discretization at each airport, depending on the level of airport congestion. An airport-time period is defined as a combination of an airport and a specific time period at that airport. More congested airports are modeled with a finer-level of time discretization (e.g., six time periods per day) to ensure that the number of flight operations does not exceed capacity during any (short) period of time. Uncongested airports are modeled with a lower level of fidelity (e.g., two time periods per day) because capacity constraints are typically not binding. Such an aggregation scheme reduces problem size compared to that using a single discretization level, without significantly compromising the modeling accuracy (Harsha, 2008).

A flight leg is a combination of departure airport-time period and arrival airport-time period. There are different fleet types for IBERIA, each characterized by its seating capacity. However, because of demand uncertainty and because of the effects of revenue management, the airlines are rarely able to sell all the seats on an aircraft. Therefore, we will assume a maximum allowable load factor for every flight leg (Vaze and Barnhart, 2012a). Our optimization formulation allows for practical constraints that ensure that certain aspects of an existing flight schedule are included in the new schedule. For example, an airline sometimes receives government grants to maintain a minimum level of service in some markets; an airline sometimes needs to schedule a pre-determined minimum number of flights in some OD pairs in order to maintain its competitive position in those markets, consistent with its corporate strategy; and an airline can lose some of its take-off and landing slots at an airport if some OD pairs are not operated. We assume that there are no flights being operated when the planning period (i.e., the time period for which operations are to be scheduled) starts because we study a domestic network where the planning period starts at night when none of the flights is flying.

Finally, we assume that the flight schedule will be periodic, that is, the schedule will repeat after the planning period ends. To satisfy this, the number of aircraft of each fleet type at each airport at the beginning and the end of the planning period must be the same.

Passenger Demand. As explained in Subsection 2.2, unconstrained demand is estimated for the combination of origin airport, destination airport and desired departure time period. This combination is defined as a market. Consequently, for each market the origin airport-time period is known. This combination of OD pair and time period is referred to as market-time period (note that airport-time period and market-time period are not the same because airport-time periods depend on the level of congestion and market-time periods don't). The terms market and market-time period will be used interchangeably in the rest of this paper. In each market, passengers can choose any of the corresponding non-stop or one-stop itineraries (or pick a competitor or the *null* alternative). Thus the proposed model is itinerary-based. The demand captured by a certain airline will depend on competition effects which are incorporated in our model through the nested logit model introduced in Section 2.

Not all the passengers are able to travel on their desired itinerary. Some passengers are spilled due to lack of available seating capacity and a fraction of those can be recaptured on either other itineraries by the same airline or on other alternatives including itineraries by other airlines, rail alternative or *null* alternative. We model recapture for a given market w and passenger type ζ using two different parameters: the split ratio ($\varphi_w^{w^+,\zeta}$) and the recapture rate (χ_w^ζ). We assume that once passengers are spilled, they are spilled to other available market-time periods (similar to Pita et al., 2012). The split ratio determines the split of passengers from the spilling time period to the time periods before and after it. It depends on the unconstrained demand of each passenger type (i.e., $H_w^\zeta d_w$) in the market-time periods immediately before (w^-) and immediately after (w^+) the spilling time period (w): $\varphi_w^{w^+,\zeta} = \dfrac{H_{w^+}^\zeta d_{w^+}}{H_{w^+}^\zeta d_{w^+} + H_{w^-}^\zeta d_{w^-}}$. In other words, passengers spilled from market-time period w may be recaptured in market-time periods w^- and w^+. The split ratios satisfy $\varphi_w^{w^-,\zeta} + \varphi_w^{w^+,\zeta} = 1$. Note that since these market-time periods immediately before and after the spilling market-time period, in turn, can also spill passengers, this modeling mechanism allows passengers from any market-time period to be recaptured in any other market-time period to a certain extent. Then, we model the recapture rate (χ_w^ζ) of an airline for a given market as a function of travel time and current market share of the spilling airline a_0 ($P_{a_0}^{w,\zeta}$) (Pita et al., 2012): $\chi_w^\zeta = \max\left\{0; P_{a_0}^{w,\zeta}\left(1 - \dfrac{1}{tt_a^w}\right)\right\}$. As shown in this expression, long-haul markets are less sensitive to departure/arrival times than short-haul markets, leading to higher recapture by the other itineraries on the spilling airline. Also, if an airline has a larger market share, it should be able to recapture a larger fraction of spilled demand. Therefore, the fraction of passengers of type ζ spilled from market-time period w whom the airline succeeds in redirecting to its own itineraries in market-time period w^+ is given by $\chi_{w^+}^\zeta \varphi_w^{w^+,\zeta}$. Thus, the airline schedule optimization problem depends on recapture rates, which in turn depend on airline market shares. Finally, the market shares themselves are dependent on the airline's optimized schedules. We solve this problem iteratively, as described in Section 4.

As mentioned above, we are integrating several sub-problems namely, frequency planning, approximate timetable development and fleet assignment, while explicitly accounting for passenger demand variation with schedule. Then, the aim of our Integrated Airline Scheduling under COmpetition Model (IASCOM) is to obtain, for all flight legs for one single airline, the frequency by fleet type, given airport slots availabilities, fleet sizes, average fares, unconstrained demands and competitors' schedules.

The notation in the IASCOM is defined as follows:

Sets

- A : set of airline operators indexed by a .
- G : set of airports indexed by g .
- OD : set of origin-destination pairs indexed by od .
- COD : subset of origin-destination pairs served by one-stop routes.
- W : set of markets indexed by w .
- F : set of flight legs indexed by f .
- I : set of itineraries indexed by i .
- I_w : subset of itineraries serving market w .
- I^f : subset of itineraries using flight leg f .
- Π : set of fleet types indexed by π .
- Z : set of passenger types indexed by ζ .
- K : set of nodes in our space-time network indexed by k ; each node is an airport-time period.
- K_g : subset of nodes belonging to airport g .
- PK_k : subset of nodes including node k and all those that precede it in time at the same airport.
- AF_k : subset of flight legs arriving in node k .
- DF_k : subset of flight legs departing from node k .
- $F1_{od}$: subset of flight legs which serve OD pair od as the first flight leg in the itinerary. Note that each OD pair is only served by a unique route in the airline network.
- $F2_{od}$: subset of flight legs which serve OD pair od as the second flight leg in the itinerary. Note that each OD pair is only served by a unique route in the airline network.
- RF_f : subset of flight legs that are subject to some regulation/other considerations such as government grants, competitive position, and slot rights.

Parameters

- p_w^{ζ} : average ticket price for passenger type ζ in market w .

- c_f^π: operating cost for flight leg f with fleet type π.
- d_w: total unconstrained demand in market w.
- alf_f: maximum allowable load factor for flight leg f.
- q_π: seating capacity of fleet type π.
- qa_k^a: the number of available arrival slots in node k for airline a.
- qd_k^a: the number of available departure slots in node k for airline a.
- τ_f: block time for flight leg f.
- b_π: average block time across the planning period per aircraft of fleet type π (i.e., actual average number of block hours for which each aircraft was operated in reality).
- n_π: fleet size for fleet type π.
- m_f: minimum frequency to be operated for flight leg f because of some regulation/other consideration.
- u_g^π: maximum number of planes of fleet type π on the ground at airport g at the beginning of the planning period (there are some airports where the airline does not overnight planes; u_g^π is 0 for those airports).
- H_w^ζ: probability for an individual in market w of belonging to passenger type ζ.
- $P_{air}^{w,\zeta}$: probability for a passenger of type ζ in market w of selecting nest air at the top level of the nested logit model. Note that $P_{air}^{w,\zeta}$ depends on f_a^{od} which is a decision variable of the mathematical model.
- $P_{a|air}^{w,\zeta}$: probability for a passenger of type ζ from market w of selecting alternative a among all the air alternatives. Note that $P_{a|air}^{w,\zeta}$ depends on f_a^{od} which is a decision variable of the model.
- χ_w^ζ: the recapture rate in market-time period w for passenger type ζ.
- $\varphi_w^{w^+,\zeta}$ ($\varphi_w^{w^-,\zeta}$): the split ratio from market-time period w to w^+ (w to w^-) for spilled passengers of type ζ.

Variables

- z_f^π: frequency of flight leg f with fleet type π.
- y_g^π: number of planes of fleet type π on the ground at the beginning of the planning period in airport g.
- h_i^ζ: number of passengers of type ζ flown on itinerary i.

- g_w^{ζ} : number of passengers of type ζ spilled from market w.
- f_a^{od} : frequency operated by airline a in OD pair od.

The IASCOM for an operator (an airline) a is as follows:

$$\max \Theta_a = \sum_{w \in W} \sum_{i \in I_w} \sum_{\zeta \in Z} p_w^{\zeta} h_i^{\zeta} - \sum_{f \in F} \sum_{\pi \in \Pi} c_f^{\pi} z_f^{\pi} \tag{6}$$

subject to:

$$\sum_{i \in I_w} h_i^{\zeta} \le P_{a|air}^{w,\zeta} P_{air}^{w,\zeta} H_w^{\zeta} d_w - g_w^{\zeta} + \chi_w^{\zeta} \left(\varphi_{w^-}^{w,\zeta} g_{w^-}^{\zeta} + \varphi_{w^+}^{w,\zeta} g_{w^+}^{\zeta} \right) \quad \forall w \in W, \zeta \in Z \tag{7}$$

$$\sum_{i \in I^f} \sum_{\zeta \in Z} h_i^{\zeta} \le \sum_{\pi \in \Pi} alf_f q_{\pi} z_f^{\pi} \quad \forall f \in F \tag{8}$$

$$\sum_{f \in AF_k} \sum_{\pi \in \Pi} z_f^{\pi} \le qa_k^a \quad \forall k \in K \tag{9}$$

$$\sum_{f \in DF_k} \sum_{\pi \in \Pi} z_f^{\pi} \le qd_k^a \quad \forall k \in K \tag{10}$$

$$f_a^{od} \le \sum_{f \in F1_{od}} \sum_{\pi \in \Pi} z_f^{\pi} \quad \forall od \in OD \tag{11}$$

$$f_a^{od} \le \sum_{f \in F2_{od}} \sum_{\pi \in \Pi} z_f^{\pi} \quad \forall od \in COD \tag{12}$$

$$\sum_{k \in K_g} \sum_{f \in AF_k} z_f^{\pi} = \sum_{k \in K_g} \sum_{f \in DF_k} z_f^{\pi} \quad \forall g \in G, \pi \in \Pi \tag{13}$$

$$\sum_{f \in F} \tau_f z_f^{\pi} \le b_{\pi} n_{\pi} \quad \forall \pi \in \Pi \tag{14}$$

$$y_g^{\pi} \ge \sum_{k' \in PK_k} \left(\sum_{f \in DF_{k'}} z_f^{\pi} - \sum_{f \in AF_{k'}} z_f^{\pi} \right) \quad \forall g \in G, k \in K_g, \pi \in \Pi \tag{15}$$

$$\sum_{g \in G} y_g^{\pi} \le n_{\pi} \quad \forall \pi \in \Pi \tag{16}$$

$$y_g^{\pi} \le u_g^{\pi} \quad \forall g \in G, \pi \in \Pi \tag{17}$$

$$\sum_{\pi \in \Pi} z_f^{\pi} \ge m_f \quad \forall f \in RF_f \tag{18}$$

$$z_f^{\pi} \in \mathrm{Z}^+ \quad \forall f \in F, \pi \in \Pi \tag{19}$$

$$y_g^{\pi} \in \mathrm{R}^+ \quad \forall g \in G, \pi \in \Pi \tag{20}$$

$$h_i^{\zeta} \in \mathrm{R}^+ \quad \forall i \in I, \zeta \in Z \tag{21}$$

$$g_w^{\zeta} \in \mathrm{R}^+ \quad \forall w \in W, \zeta \in Z \tag{22}$$

$$f_a^{od} \in \mathrm{R}^+ \quad \forall od \in OD \tag{23}$$

The objective function (6) maximizes the airline's operating profit. Operating profit is the difference between the fare revenue, given by the first term in the objective function, and the operating costs, given by the second term in the objective function. Constraints (7) ensure that the number of passengers of each type flown in each market does not exceed those allowed by the demand model minus those spilled in that market plus those recaptured from other market-time periods. The served demand is the number of passengers willing to travel in that market multiplied by the airline's market share in that market, minus the spilled passengers, plus the passengers that are recaptured from the other market-time periods. Note that $P_{a|air}^{w,\zeta} P_{air}^{w,\zeta} H_w^{\zeta} d_w$ depends on the decision variables f_a^{od}. Constraints (8) ensure that the number of passengers on a flight leg must be at most equal to the aircraft's total capacity multiplied by the maximum allowable load factor. Constraints (9) and (10) are slot constraints for arrivals and departures, respectively. Constraints (11) and (12) ensure that the frequency in each OD pair cannot be greater than the total number of flight legs operated in that OD pair by all fleet types. Note that, as mentioned earlier, the airline under study has only one route serving each OD pair. So the constraints (11) ensure that the frequency in each OD pair does not exceed the total frequency of the first flight leg and constraints (12) ensure that it also does not exceed the total frequency of the second flight leg (wherever applicable). Constraints (12) are redundant for OD pairs served by non-stop itineraries. Constraints (13) state that the schedule must be symmetric, that is, the number of departures and arrivals in every airport must be the same within the planning period ensuring the number of planes in each location at the beginning and end of a cycle is the same. Similar to Harsha (2008) we ensure the fleet capacity constraint in an indirect way: constraints (14) are fleet utilization constraints which ensure that the utilization of each fleet type must not be greater than the available total block hours during the planning period. Constraints (15) count the number of planes of fleet type π on the ground in airport g at every node $k \in K_g$. Constraints (16) limit the number of planes to the fleet size. Constraints (17) limit the number of planes on the ground at each airport to the maximum as specified by airline's requirements (i.e., there are airports where planes do not overnight, or only a pre-determined maximum number of them overnight, because there are neither crew bases nor maintenance opportunities). Constraints (18) capture regulations or some other restrictions on minimum frequencies on certain flight legs. Constraints (19)-(23) are variable value constraints.

4. Case study and base case scenario

We evaluate our model's performance with case studies focusing on a single airline's perspective. We use data provided by IBERIA, representing its operations for the year 2010. The dataset consists of operating schedule information, operating expenses, unconstrained demand values (as estimated using the procedure explained in Subsection 2.2), frequencies from other operators, and the available fleet. Air and rail competition has been considered. There are two main types of competing airlines: low-cost and legacy airlines. There are three low-cost airlines and three legacy airlines in this case study. The air-rail competition is present in six origin-destination pairs: Madrid-Barcelona, Barcelona-Madrid, Madrid-Seville, Seville-Madrid, Madrid-Malaga and Malaga-Madrid. There is one unique rail operator. Due to lack of data, we assume in this case study that there is no response from the competitors to the airline's schedule changes suggested by our model. However, each operator will likely respond to each competitor's schedule change. Therefore, the airline under study will be a part of a multi-operator, non-cooperative game. Each player (operator) in this game will choose a strategy (i.e., a set of scheduling decisions) that maximizes its own pay-off function (profit). Such a pay-off maximizing strategy for each player will depend on other players' chosen strategies. The airline under study will participate in this game by developing a pay-off maximizing strategy every time a competitor changes its schedule. In order to do so it will solve a problem modeled using IASCOM to obtain a new strategy each time.

The air network is a pure hub-and-spoke network with 23 airports and 104 OD pairs. The only hub is located in Madrid. There is no flight leg bypassing the hub airport. We discretize time at different airports at different levels based on the levels of operations by the carrier and the congestion levels of the airports. There are some airports in the Spanish network which have very low utilization (i.e., two operations per day for IBERIA). For these airports a discretization of two time periods per day is employed (one time period for each half-day). However, for congested airports, such as airports in Madrid and Barcelona, a discretization of six time periods is implemented. There are 44 possible non-stop routes and 104 total routes within the network. There are three different fleet types available for IBERIA in this case study: an A-319 fleet with 141 seats per aircraft, an A-320 fleet with 171 seats per aircraft, and an A-321 fleet with 200 seats per aircraft. We have considered a planning period of seven days. We require the planning to be periodic, that is, the fleet distribution across the airports must be the same at the beginning and the end of the planning period.

The IASCOM is a non-linear mixed integer programming model. The non-linearity is in constraints (7). The captured demand is a non-linear function of the airline's frequency values (f_a^{od}). In order to solve the model, we linearize this expression using piecewise linear functions (note that for each market w and passenger type ζ the term $P_{a|air}^{w,\zeta} P_{air}^{w,\zeta} H_w^{\zeta} d_w$ depends on one frequency decision variable: f_a^{od}). Consequently, we approximate the relationship between the fraction of passengers selecting the airline and the frequency values by a piecewise linear function. The piece sizes are selected to be the same for every origin-destination pair, equal to one frequency value. In order to linearize the non-linear relationship we use *special ordered set variables*, which is an ordered set of non-negative variables, of which at most two can be non-zero, and if two are non-zero these must be consecutive in their ordering. Special ordered sets are typically used to approximately incorporate non-linear functions of a variable into a linear model. When embedded in a Branch-and-Bound code these variables enable truly global optima to be found, and not just local optima (Beale and Tomlin, 1970). We coded the IASCOM in GAMS, using CPLEX 12.1 as the optimization solver, on a computer with 8 GB RAM and solved all models to a maximum 1% optimality gap. The computational times across all the test cases never exceeded 1837 seconds.

Now, we present our base case scenario and measure how closely our model solutions match reality, as described by IBERIA's actual schedule.

Base case scenario

An airline's revenue management practices indirectly lead to restrictions on the number of passengers to some fraction of the seating capacity. This fraction is a function of the airline's revenue management system and the passenger demand patterns. If there were no revenue management system then we would have all flights filled up to the minimum of demand and available seats. Therefore, an important parameter in the IASCOM is the maximum allowable load factor on every flight leg, designated by alf_f. Although there are historical data available on load factors, it is not obvious how to make use of it to ascertain this maximum allowable load factor. Some recent studies have used arbitrary constant values, e.g. a value of 0.85 was used by Vaze and Barnhart (2012a). We acknowledge the fact that the demand patterns, and hence the response of the revenue management systems, is likely to be different for the different markets. Consequently, we employ two different values of the maximum allowable load factors. In order to set these two base values, we look for the ratio of average fare to operating cost per seat in each market so as to split them into two groups. Figure 1 shows this ratio for each market in the network (using total fare and total operating cost in case of one-stop markets). We calculate this average ratio for the flight legs (by taking ratio of weighted averages of all markets using that flight leg) and we classify them into the two groups: for flight legs with ratio less than 1.5, we set the average maximum allowable load factor (alf_f) equal to 0.85; otherwise it is set to 1. This assumption is consistent with our dataset, where among flight legs with the ratio (of average fare to operating cost per seat) less than 1.5, 85.1% had a load factor below 0.85, and 97.7% had a load factor below 0.9. Similarly, among flight legs with the ratio (of average fare to operating cost per seat) greater than 1.5, 94.3% had a load factor between 0.85 and 1. Furthermore, our value of 0.85 is also consistent with the values used in prior literature (e.g., Vaze and Barnhart, 2012a).

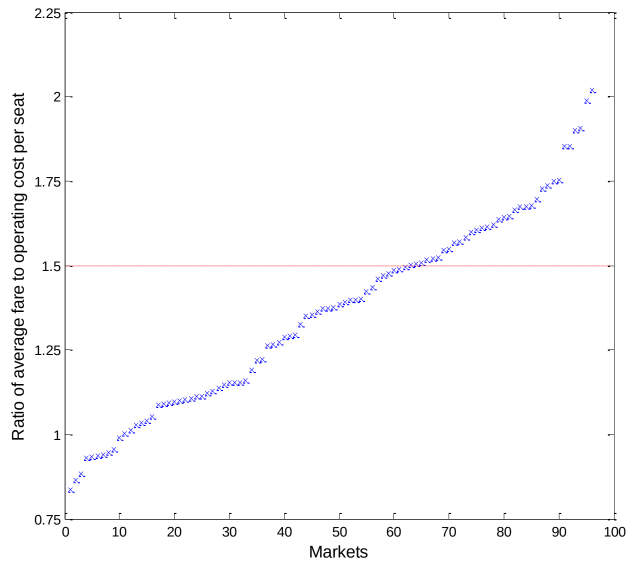


Figure 1: Ratio of average fare to operating cost per seat for each market

We further validate this assumption by performing a sensitivity analysis of the model results to variations in this assumption. The aim of this sensitivity analysis is twofold: first, to validate our choice of the maximum allowable load factor values and second, to verify the robustness of the model formulation to some perturbations in this underlying assumption about parameter values. Seven different experiment runs are conducted. The maximum allowable load factor for each flight leg is assigned a value equal to $min(\rho \cdot alf_f, 1)$, where ρ takes one of the following values for each of the seven experiment runs: $0.85, 0.90, 0.95, 1, 1.05, 1.10, 1.15$. Note that the base scenario corresponds with $\rho = 1$.

We compare the solutions provided by the IASCOM with the solution operated by IBERIA. Similar to Vaze and Barnhart (2012a), we use Mean Absolute Percentage Error (MAPE) to measure the discrepancies between model output and actual schedules, with $MAPE_1$ (equation 24) measuring the difference in the total frequency value in each origin-destination pair over the planning period; $MAPE_2$ (equation 25) measuring the difference in the frequency value per fleet type per flight leg; $MAPE_3$ (equation 26) measuring the difference in the total number of flown seats per flight leg; and $MAPE_4$ (equation 27) measuring the difference in frequency of flight legs per origin-destination pair and time period combination. Note that the subscript a has been dropped to simplify notation.

$$MAPE_1 = \frac{\sum_{od \in OD} \left| \hat{f}_{od} - f_{od} \right|}{\sum_{od \in OD} f_{od}} \tag{24}$$

$$MAPE_2 = \frac{\sum_{f \in F} \sum_{\pi \in \Pi} \left| \hat{z}_f^{\pi} - z_f^{\pi} \right|}{\sum_{f \in F} \sum_{\pi \in \Pi} z_f^{\pi}} \tag{25}$$

$$MAPE_3 = \frac{\sum_{f \in F} \left| \sum_{\pi \in \Pi} q_{\pi} \left(\hat{z}_f^{\pi} - z_f^{\pi} \right) \right|}{\sum_{f \in F} \sum_{\pi \in \Pi} q_{\pi} z_f^{\pi}} \tag{26}$$

$$MAPE_4 = \frac{\sum_{od \in OD} \sum_{k \in K_{od}} \left| \sum_{f \in DF_k} \sum_{\pi \in \Pi} \hat{z}_f^{\pi} - z_f^{\pi} \right|}{\sum_{od \in OD} \sum_{k \in K_{od}} \sum_{f \in DF_k} \sum_{\pi \in \Pi} z_f^{\pi}} \tag{27}$$

where $\hat{f}_{od}, \hat{z}_f^{\pi}$ are the solutions provided by the IASCOM optimization, f_{od}, z_f^{π} are based on the current schedules operated by the airline, and K_{od} is the subset of airport-time periods which belong to the origin airport of the OD pair od.

Table 3: Comparison of model-predicted schedules with the actual IBERIA schedule

ρ	$MAPE_1$ (%)	$MAPE_2$ (%)	$MAPE_3$ (%)	$MAPE_4$ (%)	AMS (%)			PAX (%)	$PROFIT$ (%)	$RASK$ (%)	RPK (%)	$FLEET$ (%)		
					BCN	SVQ	AGP					A-319	A-320	A-321
0.85	12.05	25.71	15.06	29.13	0.00	-6.45	-2.31	-9.95	-16.83	-10.81	5.45	-2.53	-1.41	-17.02
0.9	8.01	16.29	9.37	18.72	1.29	-4.51	-1.06	-5.03	-8.91	-6.19	4.52	-1.69	-0.63	-10.01
0.95	5.31	8.92	6.82	12.03	1.71	-2.42	-0.49	-1.06	-3.11	-2.15	1.97	0.67	0.33	-3.41
1	4.25	6.74	4.96	9.46	1.29	-1.26	-0.49	2.15	2.93	0.82	0.45	-0.09	-0.58	0.18
1.05	4.63	9.08	5.87	13.29	1.71	-1.56	0.68	6.91	8.35	3.65	-0.72	-2.57	0.82	0.65
1.1	7.41	12.91	8.18	18.99	1.29	-1.56	1.25	8.24	14.12	5.79	-2.03	-4.53	-0.71	1.21
1.15	11.07	24.36	13.71	26.16	0.97	-2.42	1.51	11.71	18.34	6.41	-2.91	-7.65	-1.53	0.71

Table 3 shows a summary of the results obtained for the base case scenario. The first column of the table lists ρ values. The values of the Mean Absolute Percentage Errors are in the second, third, fourth and fifth columns. Frequency values predicted by our model IASCOM (see $MAPE_1$) are found to be close to the current schedule operated by IBERIA. The difference in fleet assignment (see $MAPE_2$) is slightly higher than that in frequencies but is still small in an absolute sense, especially for the ρ=1 case. The variation in the number of seats flown (see $MAPE_3$) is closely related to the variation in frequencies. The variation in frequency of flight legs per origin-destination and time period (see $MAPE_4$) is the highest as compared with the rest of MAPEs, reflecting that the approximate optimal timetable slightly differs from the current one operated by the airline when compared at a very disaggregate level. Even then, the difference is less than 10% for the ρ=1 case. Similar to the mean absolute percentage error values, the individual differences between predicted and actual flight frequencies were also found to be small for the ρ=1 case.

The sixth column ($AMS = \frac{\sum_{w \in W} \sum_{\zeta \in Z} \left(\hat{P}_{air}^{w,\zeta} - P_{air}^{w,\zeta} \right)}{\sum_{w \in W} \sum_{\zeta \in Z} P_{air}^{w,\zeta}}$) in Table 3 shows the mean percentage difference in the air mode's market share as provided by the IASCOM optimization ($\hat{P}_{air}^{w,\zeta}$) compared to the actual values as provided in the data obtained from IBERIA ($P_{air}^{w,\zeta}$), for the markets in which HSR currently operates, namely, Madrid-Barcelona, Madrid-Seville and Madrid-Malaga. For the Madrid-Barcelona case (BCN), the predicted market share is never below the actual value. This is due to the fact that in this market there is a constraint (imposed by the airline to ensure their competitive position in that market) that requires a minimum level of service to be maintained and therefore, the frequency number is always above a minimum number (constraints imposing this minimum level are found to be binding only in the optimization model solution for $\rho = 0.85$). In the Madrid-Seville market (SVQ), the air market share is comparatively more sensitive to the schedule predicted by the model, which in turn is dependent on ρ . This is due to the fact that in this market the only operating airline is IBERIA. The model predicts that the airline should offer somewhat lower frequencies due to the competition of HSR compared to those actually offered by the airline currently. Finally, in the Madrid-Malaga case (AGP), the model predicts that the market share grows with the ρ value. Note that the errors in market shares in Barcelona-Madrid, Seville-Madrid and Malaga-Madrid ODs are not presented here because the results are very similar as those for the corresponding opposite directional ones presented here.

Other important metrics of interest include the mean percentage difference in the number of passengers transported ($PAX = \frac{\sum_{i \in I} \sum_{\zeta \in Z} \left(\hat{h}_i^{\zeta} - h_i^{\zeta} \right)}{\sum_{i \in I} \sum_{\zeta \in Z} h_i^{\zeta}}$) and in IBERIA's total profit ($PROFIT = \frac{\hat{\Theta}_a - \Theta_a}{\Theta_a}$) (Table 3: columns seven and eight, respectively). Both behave in a similar way except that the percentage difference in profits is more sensitive than that in passengers, which is expected. Note that when the difference in the number of passengers transported is close to zero, the difference in profit is also close to zero. It is so because the predicted frequency and the predicted number of passengers transported in every market is very close to those in the IBERIA schedule.

The quality of the schedule can also be analyzed by looking at the percentage difference (compared to the actual values obtained from IBERIA) in the average revenue per available seat kilometer ($RASK$), and the average revenue per passenger kilometer (RPK) (Table 3: columns nine and ten, respectively). $RASK$ and RPK are calculated as average values over the whole network and their mathematical expressions are the same as the one for $PROFIT$ (replacing Θ_a with the average revenue per available seat kilometer and the average revenue per passenger kilometer, respectively). The $RASK$ increases monotonically as ρ increases, that is, as we increase the maximum allowable load factor, the aircraft can accommodate more passengers and the revenue per seat kilometer increases. When the maximum allowable load factor is lower, the percentage difference with actual values is negative, which means that airplanes have more empty seats. The RPK has the opposite behavior. When the number of captured passengers is low, the model solution preferentially serves the higher fare passengers, and when the maximum allowable load factor goes up, it also serves an increasing number of lower fare passengers. Note that we are using average fares for each market for each passenger type: business and leisure. So the impacts of changes in ρ on $RASK$ and RPK are manifestations of two different effects: 1) a preference for accommodating more passengers on higher fare markets by providing additional flights or larger aircraft on such markets when not all the passenger demand on all markets can be accommodated due to limited fleet size (especially when ρ is low); and 2) a preference for accommodating more business passengers than leisure passengers when all the passengers demand for both types of passengers cannot be accommodated.

Finally, the last column ($FLEET$) in Table 3 shows the percentage difference in fleet utilization with respect to IBERIA's actual values (fleet utilization is measured as the average block hours flown per day). Its mathematical expression is the same as the one for $PROFIT$ (replacing Θ_a with the fleet utilization for that fleet type). Fleet utilization depends on the schedule operated. In general, fleet utilization is lower when the maximum allowable capacity in flight legs is low (due to ρ being low). This can be attributed to the fact that when the maximum allowable number of seats to be sold on each flight leg goes down, it becomes less profitable to operate certain flight legs. As a result, in some cases, it becomes more profitable to keep the aircraft on the ground rather than to operate flight legs that are marginally profitable at higher values of maximum allowable load factor. As a result, the number of scheduled flight legs is also lower at lower values of maximum allowable load factors. The scheduled frequencies increase as the value of ρ increases, thereby increasing fleet utilization. Nevertheless, when the value of ρ becomes too high, scheduled frequencies decrease. This is so because the offered capacity in each flight leg is now so large that passenger demand is satisfied with fewer flight legs, thus decreasing fleet utilization (except for the A-321 fleet, which has the highest capacity). At higher values of ρ, it becomes more economical to fly the largest planes. The cost per seat is lower at higher values of ρ because the planes are allowed to carry more passengers.

For almost all the metrics compared in Table 3, the difference between the model predictions and IBERIA's actual schedules is found to be the lowest (or close to the lowest) for a ρ value of 1. This provides credibility to our choice of the two base values (0.85 and 1) of maximum allowable load factors. Additionally, for most of the comparison metrics in Table 3, the differences between model predictions and IBERIA's actual schedules are found to be reasonably small for all but the extreme (i.e., 0.85 and 1.15) values of ρ. This shows that our model predictions are quite robust across multiple potential values of ρ. We will use ρ = 1 value, that is, we will use the two base values (0.85 and 1) of maximum allowable load factors, for all our subsequent computational experiments in this paper.

As stated in Section 3, the recapture rate for a given market and passenger type is dependent on the airline's market shares. However, if airline's flight frequencies change, market shares will also change. In order to overcome this drawback, we solve the IASCOM model iteratively and update the airline's market shares at each iteration (we take averages of the market share values corresponding to the previous two iterations in order to update them). Let $P_a^{w,\zeta}(\nu)$ be the airline's market share at iteration ν in market w and for passenger type ζ. Again we use Mean Absolute Percentage Error (MAPE) to measure the discrepancies of the optimization model output between iterations ν and $\nu - 1$:

$$MAPE_a^{\nu} = \frac{\sum_{w \in W} \sum_{\zeta \in Z} \left| P_a^{w,\zeta}(\nu) - P_a^{w,\zeta}(\nu - 1) \right|}{\sum_{w \in W} \sum_{\zeta \in Z} P_a^{w,\zeta}(\nu - 1)}. \tag{28}$$

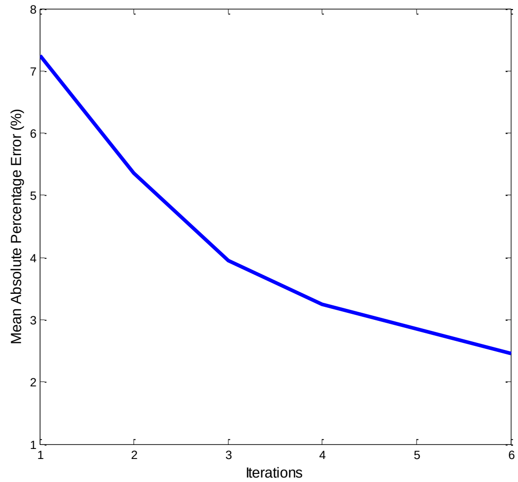


Figure 2: Trend in Mean Absolute Percentage Error in the airline's market shares across iterations.

Figure 2 shows the variation of $MAPE_a^v$ across the iterative procedure. Note that the $MAPE_a^v$ value drops rapidly to less than 3% within the first five iterations indicating that our model yields reasonably accurate results within a handful of iterations. After this threshold, $MAPE_a^v$ continues decreasing but at a slower rate. All results presented in the earlier parts of this section correspond with the last (we conducted six iterations) iteration in the iterative procedure.

5. Impacts of the entry of High Speed Rail

In Spain, HSRs, mainly developed by the government, are a strong source of competition to airlines. They first operated in the origin-destination pairs Madrid-Seville and Seville-Madrid. Then, the government expanded HSR operations to the other OD pairs including Madrid-Malaga, Malaga-Madrid, Madrid-Barcelona, Barcelona-Madrid, Madrid-Valencia and Valencia-Madrid. In all these cases, the introduction of HSRs translated into a loss in market share for the airlines. Consequently, airlines had to change their operations to match the lower levels of demand by offering lower frequencies and smaller fleet sizes. In this section, we apply our IASCOM model, developed in Section 3, to understand the changes in the competitive landscape of airline markets upon entry of HSR by investigating the airline response to HSR entry. We use data provided by IBERIA, representing its operations for the years 2008 (in order to study past HSR entries) and 2010 (in order to predict post-HSR entry scenario).

We divide this section into three subsections. In Subsection 5.1, we develop a demand stimulation model to model the changes in the unconstrained demand due to HSR entry. In Subsection 5.2, we validate our model's ability to predict the airline response to HSR entry by using past data from a time period before an actual HSR entry into a market to predict the post-HSR entry scenario and compare it with the actual post-HSR entry scenario. In Subsection 5.3, we consider a scenario with HSR entry in six new origin-destination pairs, and we solve the IASCOM model in order to study the predicted responses.

5.1 Demand stimulation modeling

In the overview of IASCOM model, presented in (6)-(23), we assume that the unconstrained demand is fixed for each market. This assumption is reasonable in many markets where the entry of new operators is unlikely. However, if new operator entry occurs, the unconstrained demand will likely change due to demand stimulation. Consequently, we modify our model to account for this demand stimulation effect.

The objective is to explain the variation in demand, by identifying the attributes that have the greatest and the most direct impact on total unconstrained demand. The major factors affecting the total travel demand in a market are the price of travel, total trip time and demographics related to the market itself (Belobaba, 2009). With the entry of HSR, the average price of the trip in the market and the total frequency value in the OD ($\sum_{a \in A} f_a^{od} + f_{rail}^{od}$) are likely to vary. No significant demographic changes are expected to occur because we are looking at a planning period of several months, rather than several years. Consequently, we model the variation in unconstrained demand with the total frequency (as a measure of schedule displacement, which is a part of the total trip time) and average fare in the market. In other words we model the elasticity of unconstrained demand to frequency and price.

The concepts of price and time elasticity of demand for air travel can be incorporated into an unconstrained demand function. Consider the following multiplicative model of demand for travel in a given market (Grosche et al., 2007; and Belobaba, 2009):

$$d_w = M_w \cdot \overline{p}_{od}^{\,\varepsilon} \cdot \left[\overline{tt}_{od} + \frac{sd}{f_{od}} \right]^{\phi} \quad , \tag{29}$$

where M_w is a market sizing parameter (constant), $\overline{p}_{od}$ is the average price of travel, $\overline{tt}_{od}$ is the average trip time, f_{od} (= $\sum_{a \in A} f_a^{od} + f_{rail}^{od}$) is the frequency value, and ε and ϕ are price elasticity and time elasticity of demand, respectively. sd is a constant expressed in hours. The schedule displacement component of total trip time may be expressed as sd / f_{od} . We assume that the desired departure times of passengers are distributed uniformly. Therefore, the mean schedule displacement time for a typical passenger can be easily calculated, assuming the departure times are chosen to minimize mean schedule displacement times (Belobaba, 2009; and Vaze and Barnhart, 2012). Using this approach, the value of constant sd equals 31.5 for our case studies. Values of ε and ϕ can be estimated from a historical data sample of d_w, $\overline{p}_{od}$, $\overline{tt}_{od}$ and f_{od} from similar markets over a period of time. Statistical estimation techniques like ordinary least squares regression applied to historical data (provided by the airline) provide us with the best fit curve (Belobaba, 2009).

We developed this model extension to be applied to the case when HSR enters as a new operator in a market where the airline was already operating. Consequently, we used a sample of similar markets over a period of time in order to estimate the values of M_w, ε and ϕ. These similar markets where IBERIA also faced the entry of the HSR are Madrid-Malaga/Malaga-Madrid, Madrid-Barcelona/Barcelona-Madrid, and Madrid-Seville/Seville-Madrid. The entry of the HSR in these markets occurred in years 2007, 2008 and 1992, respectively. The total number of rows in our dataset equaled 4539, one corresponding to each combination of OD pair, day of week and time period. Other elements in the dataset include overall demand, price, travel time and frequency. Table 4 presents a summary of the data used for parameter estimation (averages and ranges): overall unconstrained demand (number of requested trips in each market before and after the rail entry), price (average fare paid by each passenger in each OD pair before and after the rail entry), travel time (average service time between each OD pair before and after the rail entry), and frequency (total number of operated services in each OD before and after the rail entry). Taking logarithms of both sides of equation (29) (i.e., $\log d_w = \log M_w + \varepsilon \log \overline{p}_{od} + \phi \log\left[\overline{tt}_{od} + \frac{sd}{f_{od}}\right]$) and then regressing using the ordinary least squares estimation method, we obtained the numerical values for ε and ϕ of -0.8759 and -1.0961, respectively. Values of M_w for each OD pair-day of week-time period combination are also obtained from this estimation. For the markets where HSR is expected to enter in future, we substitute the estimated values ε, ϕ and market-specific values of $d_w, \overline{p}_{od}, \left[\overline{tt}_{od} + \frac{sd}{f_{od}}\right]$ in (29) to calculate the M_w values for those markets. A subset of those values (for the markets in Madrid-La Coruña and La Coruña-Madrid OD pairs) is presented in Table 7 in Appendix A. We use the classical F-test for linear models to test the statistical significance of this model. This test statistic equals 4.9817. Using the critical value of 3.4668 (with two degrees of freedom in the numerator and 21 degrees of freedom in the denominator; 0.95 confidence level), we reject the null hypothesis that the naïve model (with only the market sizing parameter) is as good as the full model. The r-squared value for the linear model fit described here was found to be 66.83%.

Table 4: Demand stimulation model estimation data summary

Type of Data	Before		After	
	Average	Data range	Average	Data range
Overall unconstrained demand	445.3	[318.7, 693.5]	578	[459.3, 1086.1]
Price (Euro)	182.2	[102.5, 392.1]	148.3	[62.1, 291.3]
Travel time (hours)	1.30	[1.25, 1.36]	1.98	[1.25, 2.98]
Frequency (per week)	75.4	[48, 144]	101.3	[77, 252]

The unconstrained passenger demand in each market w (d_w) is assumed to vary as described in (29). Therefore, the captured demand (Δ_a^w) by airline a in each market w is $\Delta_a^w = \sum_{\zeta \in Z} P_{a|air}^{od,\zeta} P_{air}^{od,\zeta} H_w^\zeta d_w$, where $d_w = M_w \cdot \overline{p}_{od}^{\,\varepsilon} \cdot \left[\overline{tt}_{od} + sd \big/ \left(\sum_{a \in A} f_a^{od} + f_{rail}^{od} \right) \right]^\phi$. Thus, in the extended IASCOM formulation, the first term on the right side of constraints (7) is replaced by the following (note that for each market w it still depends on one frequency decision variable: f_a^{od}):

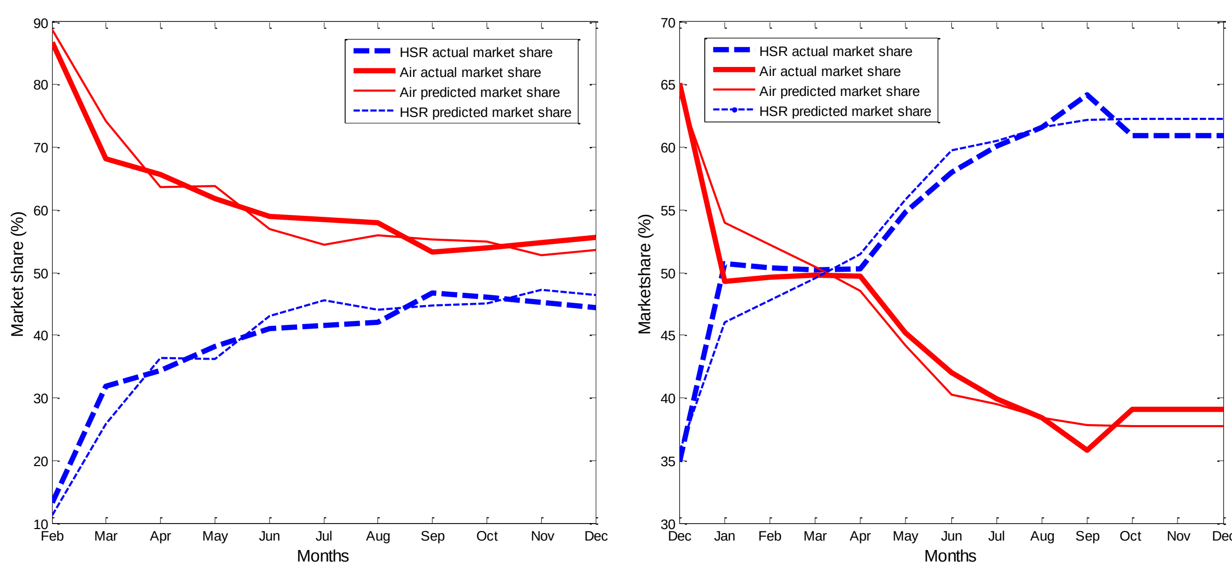
$$\frac{e^{\eta(\zeta|w)}}{\sum_{\zeta' \in Z} e^{\eta(\zeta'|w)}} \frac{e^{v(a|w,\zeta)}}{\sum_{a' \in A} e^{v(a'|w,\zeta)}} \frac{e^{\theta_{air} \log \sum_{a' \in A} e^{v(a'|w,\zeta)} + \gamma_{air} \log\left(tt_{air}^w\right)}}{e^{\theta_{air} \log \sum_{a' \in A} e^{v(a'|w,\zeta)} + \gamma_{air} \log\left(tt_{air}^w\right)} + e^{v(rail|w)} + e^{v(null|w)}} M_w \overline{p}_{od}^{\,\varepsilon} \left[\overline{tt}_{od} + \frac{sd}{\sum_{a' \in A} f_{a'}^{od} + f_{rail}^{od}} \right]^\phi \tag{30}$$

.

5.2 Model validation against post-HSR entry scenario

In order to validate the optimization model's prediction accuracy, we compare the model predictions about the airline response to HSR entry with the actual response by the airline after the HSR entry. We study the entry of the HSR in two different case studies: Madrid-Barcelona/Barcelona-Madrid and Madrid-Malaga/Malaga-Madrid. The results for Barcelona-Madrid and Malaga-Madrid OD pairs are found to be similar to Madrid-Barcelona and Madrid-Malaga respectively, and hence are omitted here. The OD pair Madrid-Seville is not included in this validation because HSR-entry in it took place several years ago (in 1992) when the airline's resource availability (e.g., aircraft availability and capacity) was significantly different than what it is now. In order to estimate the demand stimulation model parameters for each of the two case studies, we use the data before and after the HSR entry for other OD pairs. For example, in the Madrid-Barcelona/Barcelona-Madrid case study, we use the data from Madrid-Seville/Seville-Madrid, and Madrid-Malaga/Malaga-Madrid OD pairs.

The expansion of the HSR into the Madrid-Barcelona and Barcelona-Madrid markets took place during February 2008. Consequently, we study different scenarios drawn from year 2008. Each scenario corresponds to one week of each month of year 2008 (from February to December, that is, eleven scenarios). The expansion of the HSR into the Madrid-Malaga and Malaga-Madrid markets took place during December 2007. Consequently, each scenario corresponds to one week of December 2007 and one week of each month of year 2008 (that is, thirteen scenarios) (Ferropedia, 2014). For each of the scenarios we solve a restricted version of the extended IASCOM with the demand model in (30). This restricted version of the optimization model optimizes the schedule in the OD pair under study and assumes a fixed schedule for the rest of the network (provided by the airline). It is important to note that in predicting the post-HSR entry scenario for the market under study, we use information from other similar markets and the pre-HSR entry data from the market under study. Note that this is similar to the type of information that an airline will have at its disposal when making decisions about its response to an impending entry by an HSR operator. So we use these results to validate our approach.



(a) (b)

Figure 3: Variation in (a) air and rail market shares over an 11 month period after the entry of the HSR in the Madrid-Barcelona OD pair, and (b) air and rail market shares over a 13

month period after the entry of the HSR in the Madrid-Malaga OD pair

Figure 3 shows the variation in the air and rail market shares with the entry of the HSR in (a) the Madrid-Barcelona OD pair and (b) the Madrid-Malaga OD pair. Note that the HSR slowly increased its frequency over time in these markets and then stabilized. The solid red lines show the air market share variation (the thicker one the actual market share and the thinner one the predicted market share) and the dashed blue lines show the rail market share variation (the thicker one the actual market share and the thinner one the predicted market share). In the Madrid-Barcelona OD pair the HSR entry took place during February 2008. The air market share dropped and the rail market share increased from February through September almost uniformly and then somewhat stabilized from October through December. In the Madrid-Malaga OD pair the HSR entry took place during December 2007. The air market share dropped and the rail market share increased from this point until September and then stabilized from October to December. As shown in Figure 3, the model-predicted market shares for the air and rail modes are found to be reasonably close to their actual values.

Table 5 shows a summary of the results obtained in this validation process. For each OD pair under study, we show two different items: $MAPE_1$ (see equation 24) and $MAPE_{air}$ (see equation 31), both calculated for the OD showed in the first column of the table.

$$MAPE_{air} = \frac{\sum_{w \in W_h} \sum_{\zeta \in Z} \left| \hat{P}_{air}^{w,\zeta} - P_{air}^{w,\zeta} \right|}{\sum_{w \in W_h} \sum_{\zeta \in Z} P_{air}^{w,\zeta}}, \tag{31}$$

where W_h is the set of the markets where the HSR was introduced, $\hat{P}_{air}^{w,\zeta}$ is based on the solution provided by the IASCOM optimization, and $P_{air}^{w,\zeta}$ is based on the actual schedule operated by the airline. The table shows the values of $MAPE_1$ and $MAPE_{air}$ for every scenario. The last 13 columns of Table 5 represent the months of December 2007 through December 2008. As explained earlier, the values for Madrid-Barcelona and Barcelona-Madrid markets start from February 2008. Note that the values are reasonably small. The average MAPE values in terms of frequencies and air shares are 5.18% and 5.71% respectively and all the MAPE values are found to be in the range between 3.63% and 8.29%. Thus the optimization model predictions are found to be quite close to the actual post-HSR entry scenario. $MAPE_1$ measures the error in the frequency values. Therefore, the values for ODs Madrid-Barcelona and Barcelona-Madrid, and Madrid-Malaga and Malaga-Madrid are symmetric.

Table 5: Mean Absolute Percentage Errors in IASCOM predictions in terms of frequencies

and air market shares

OD	Item	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Madrid-Barcelona	$MAPE_1$ (%)	-	-	4.52	6.65	4.70	4.29	5.03	5.42	5.56	4.71	4.69	4.51	4.89
	$MAPE_{air}$ (%)	-	-	6.07	8.29	4.95	5.57	5.36	7.54	6.03	6.70	5.38	5.75	6.01
Barcelona-Madrid	$MAPE_1$ (%)	-	-	4.52	6.65	4.70	4.29	5.03	5.42	5.56	4.71	4.69	4.51	4.89
	$MAPE_{air}$ (%)	-	-	6.12	8.26	5.54	5.39	5.71	6.91	6.15	6.42	4.92	5.41	6.09
Madrid-Malaga	$MAPE_1$ (%)	4.03	7.51	6.07	5.80	5.09	5.61	5.13	5.26	4.19	5.45	5.10	5.10	5.10
	$MAPE_{air}$ (%)	4.15	7.85	5.89	5.67	4.83	5.48	5.60	4.90	3.63	5.39	5.31	5.31	5.31
Malaga-Madrid	$MAPE_1$ (%)	4.03	7.51	6.07	5.80	5.09	5.61	5.13	5.26	4.19	5.45	5.10	5.10	5.10
	$MAPE_{air}$ (%)	5.06	7.53	6.01	5.83	5.39	4.88	5.28	4.47	3.80	6.52	5.16	5.16	5.16

The impacts of the HSR entry in the Madrid-Barcelona/Barcelona-Madrid OD pairs, and in Madrid-Malaga/Malaga-Madrid OD pairs are significantly different. There are several fundamental differences between the two OD pairs in terms of the exogenous data. Firstly Madrid-Barcelona distance is 618 km compared with Madrid-Malaga distance which is 534 km. Secondly, passenger attributes are significantly different. In particular, the proportion of passengers belonging to business segment and the unconstrained demand values are both much greater for Madrid-Barcelona/Barcelona-Madrid OD pairs. Finally, products offered are also considerably different. The daily frequency in Madrid-Barcelona/Barcelona-Madrid OD pairs is usually significantly higher than the frequency value in Madrid-Malaga/Malaga-Madrid OD pairs. This is reflected in the qualitative differences between the Figures 3 (a) and 3 (b). Upon the entry of HSR in the Madrid-Barcelona market, the air market share continues to be above 50% while in Madrid-Malaga OD pair it drops considerably below 50%.

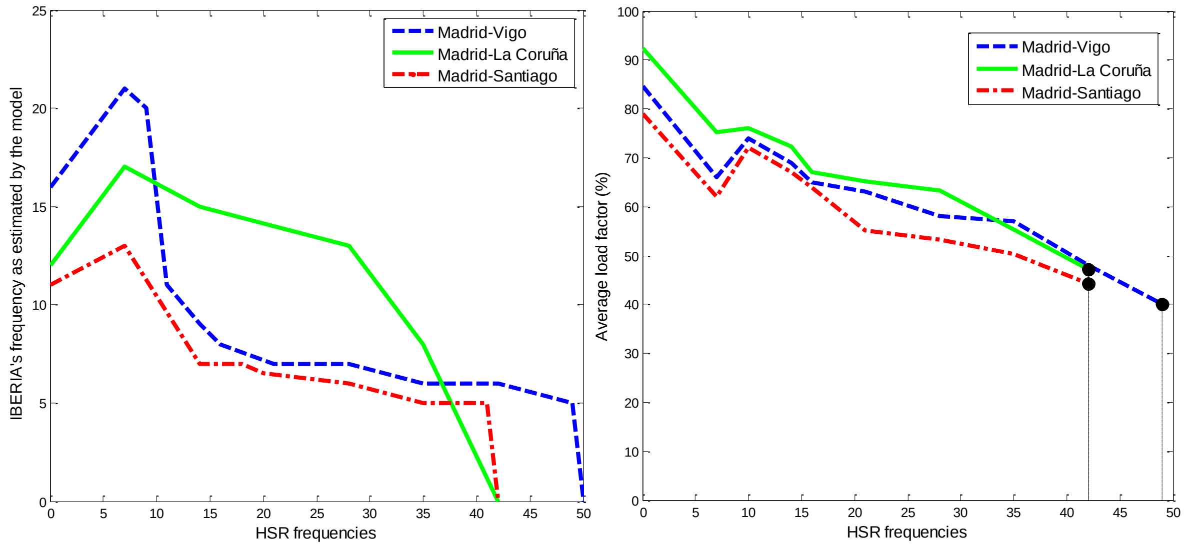
5.3 Predicting airline response to a future entry of high speed rail

In the near future, the Spanish government is planning to operate HSRs between Madrid and the region of Galicia, in the northwest of Spain (Ministerio de Fomento, 2012). Consequently, IBERIA will face the challenge again of modifying its schedule to compete not only against other airlines, but also against HSR, in these additional markets.

The origin-destination pairs affected by this new competition are those in the Galicia region: Madrid-Vigo, Vigo-Madrid, Madrid-La Coruña, La Coruña-Madrid, Madrid-Santiago and Santiago-Madrid. For these experiments, we assume that the airlines do not change their available fleet. However, if we know about any changes to an airline’s fleet, then the IASCOM model makes it straightforward to test the impact of such changes on the optimal schedules. Initially, we assume that the airline will maintain pricing and revenue management practices to be the same as in the base case scenario and thus the base fares will be similar. We relax this assumption later in this section and provide insights into the patterns in optimal fare changes.

We solve the extended IASCOM with the demand stimulation model introduced in Subsection 5.1. Figure 4(a) shows the model's solution in terms of IBERIA frequencies (on the y-axis) for several hypothetical scenarios characterized by different frequency values of operation of the HSR (on the x-axis). We assume a simultaneous entry of HSR into all six markets because they will be served along the same corridor. Therefore, we assume that the offered frequency in every market will be the same. Note that this is not due to any limitation of our models or approach, and can be easily modified to fit any other scenario. We make this assumption simply because this is the most likely scenario in reality. As before, the planning period is seven days. The average fare for HSR is assumed to be proportional to the number of kilometers in the trip and this proportionality constant is obtained as an average from the current HSR markets (Garcia and Luceño, 2010).

The response is found to be similar for every market directly affected by the new competitor. As HSR starts operating at a small frequency in a new market, the airline's frequency, as predicted by the optimization model, increases with increase in HSR frequency. By increasing the frequency of its services, the airline strives to make its schedule attractive to the passengers and attempts to maintain a large share of the market. As HSR frequency increases further, there is a threshold point beyond which the predicted frequency of the airline starts decreasing until a second threshold frequency value is reached by HSR. Beyond that first threshold point it becomes too expensive for the airline to maintain a very high market share and finds it unprofitable to keep competing at a high frequency level against the rail operator. Beyond the second threshold point, the airline holds its frequency relatively steady until another (third) frequency threshold is reached by HSR. Beyond this third threshold, the airline stops serving that route and its frequency falls to 0. As the connecting passengers are not affected by the entry of HSR, they always choose to fly. Between the second and the third threshold points the airline continues to carry a relatively steady share of the connecting passengers but a uniformly decreasing proportion of the direct passengers. As fewer direct passengers select the airline due to competition from HSR, the flight leg starts becoming unprofitable, and thus, beyond the third threshold point, the optimal schedule for the airline does not serve those origin-destination pairs. Of course, the airlines can decide to continue operations in certain origin-destination pairs for reasons other than short-term profitability maximization, such as, maintaining the airline's image or long-term growth strategy, etc. In such cases, the frequency will never be allowed to drop below a certain minimum value.



(a) (b)

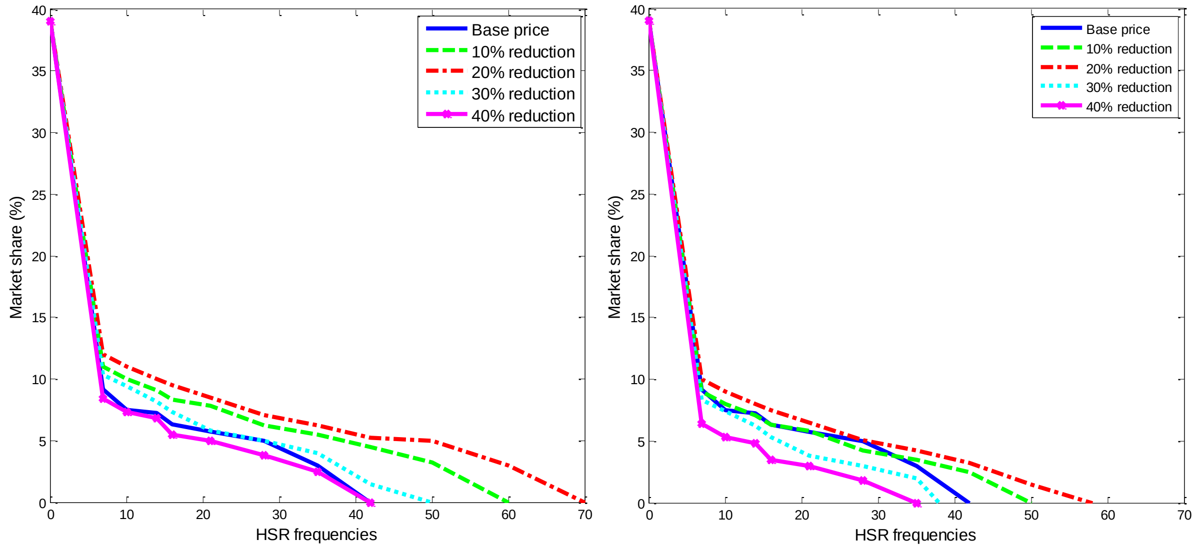
Figure 4: (a) Model-predicted IBERIA frequency in response to HSR frequency variation in origin-destination pairs of Madrid-Vigo, Madrid-La Coruña and Madrid-Santiago; (b) Model-predicted average IBERIA load factor in response to HSR frequency variation in the origin-destination pairs of Madrid-Vigo, Madrid-La Coruña and Madrid-Santiago

Figure 4(b) depicts the average load factors for IBERIA for the origin-destination pairs Madrid-Vigo, Madrid-La Coruña and Madrid-Santiago as a function of the HSR frequency, as per the IASCOM solution. The curves for the remaining (opposite directional) origin-destination pairs with HSR entry were also found to be similar to those presented in Figures 4(a) and 4(b). Up to the first threshold point, average load factor decreases with increasing schedule frequencies as the airline attempts to maintain passenger share by increasing its own frequency (see Figure 4(a)). When the airline's frequency is decreased beyond the first threshold point, the average load factor increases at first. However, it starts decreasing again until the point when flight legs are not profitable and service is discontinued. This decrease in the average load factors occurs while the frequency values are constant. This is due to the fact that connecting passengers always choose to fly and as the HSR frequency increases, increasingly fewer direct passengers choose to fly.

With the IASCOM, we study tactical competition in which average fare is assumed to have a constant value. However, with the entry of new competitors IBERIA is likely to use pricing (in addition to schedule changes) to better compete with HSR, and consequently changes in average fare values are possible through changes in the airline's pricing and revenue management practices for these markets (note that applying price reductions makes the alf_f values change). We analyze what might happen under four different price reduction scenarios, corresponding to a 10%, 20%, 30% and 40% reduction in IBERIA's average prices in markets where the new HSR competitor has entered. We compare these four scenarios with the base scenario (with original prices). The reduction in average prices is only applied to the tickets of the airline under study, that is, IBERIA. We assume that IBERIA will be the first airline responding to the entry of the HSR. Then, the rest of the airlines are likely to propose a new schedule. This would lead to an iterative process until an equilibrium point is reached.

In Figure 5(a), we show the IBERIA market share predicted by our model in response to HSR entry in the Madrid-La Coruña origin-destination pair for the base price and for 10%, 20%, 30% and 40% reductions in base price. The curves for the remaining origin-destination pairs with HSR entry were found to be similar. As the ticket price is lowered from 0% through 20% of the original price, IBERIA market share increases. By reducing average ticket price, our model predicts that IBERIA could compete for a larger set of scenarios against HSR. However, this is not true for further price reductions from 20% to 40%. By offering greater discounts in average prices, flight legs start becoming non-profitable. Our model results show that the number of frequencies is lowered and therefore, market share drops. In general, it is clear that for any scenario involving HSR entry, the loss of market share for IBERIA would be significant compared to the pre-HSR scenario.

In order to evaluate a scenario where the other airlines match the price reduction by IBERIA, we run another set of four scenarios with price reduction matching by other airlines. Figure 5(b) shows the market share predicted by our model in response to HSR entry in the origin-destination pair Madrid-La Coruña with the price reduction by IBERIA also matched by the rest of the airlines. The results are found to be similar to those in Figure 5(a), except that IBERIA captures a lower number of passengers, compared to that in Figure 5(a), due to price reductions by the rest of the airlines and is able to compete for a smaller set of scenarios. Also, the market share gains achieved through moderate ticket price reductions were found to be lower when they are matched by the other airlines.

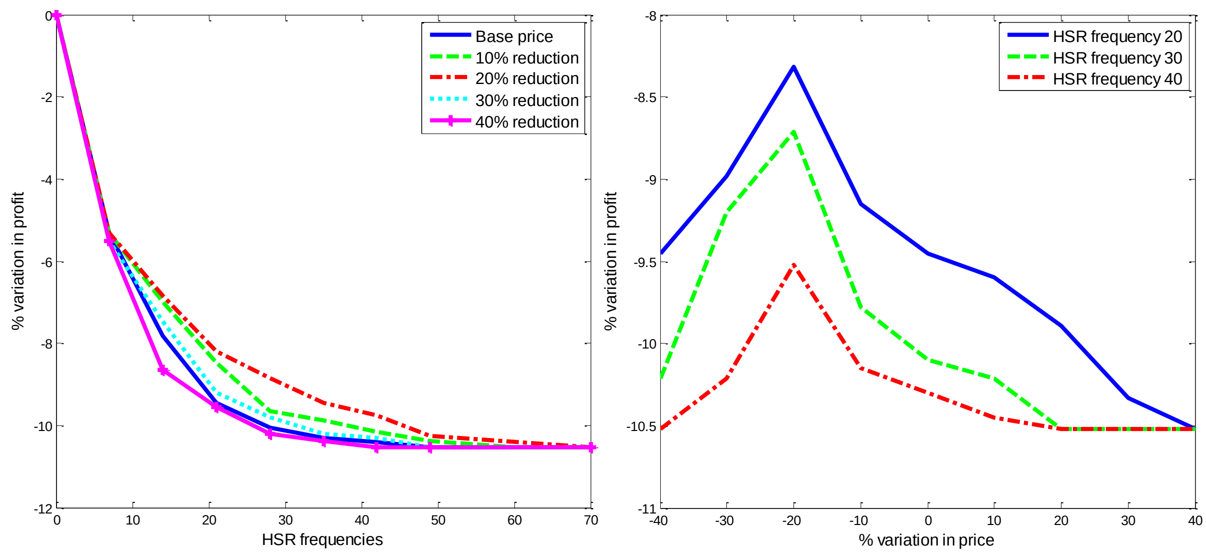


(a) (b)

Figure 5: Model-predicted IBERIA market share in response to HSR frequency variation in the origin-destination pair of Madrid-La Coruña and with different levels of fare reductions: (a) by IBERIA alone, and (b) by IBERIA and the rest of the airlines

The overall impact on total IBERIA profits of the scenario where the reduction in average fares is only applied to the IBERIA tickets is shown in Figure 6(a). With an increase in HSR frequency value, the predicted total profit decreases and eventually reaches a constant value beyond which the HSR frequency ceases to have any impact on our model's predicted profit. This is due to the fact that our model predicts that IBERIA should exit from those markets where HSR has entered. Thus, there comes a point beyond which the increase in HSR's frequency value has no effect on IBERIA's profit. For the scenarios with proposed moderate fare reductions (10% and 20%), the total profit is greater than the profit in the base scenario. In particular, 20% fare reduction results in the highest predicted profit at all levels of HSR frequencies. However, more aggressive discounts in the average prices do not result in further profit increases because the operation of flight legs on those discounted markets might not be profitable due to the decreasing revenue per passenger. Moreover, due to these discounts, other markets might become relatively more profitable and the optimized schedule might increase the number of flight legs in those other markets at the expense of these markets with HSR entry.

Figure 6(b) shows the predicted profit variation in response to variations in IBERIA's average fares in markets affected by the HSR entry. Three different curves are depicted: each one represents a different HSR frequency value, namely, 20, 30 and 40. Note that the most likely frequency value for a real future scenario lies between these values (Ministerio de Fomento, 2012). The results show that the profit reduction due to HSR entry is the least for a price reduction of the 20%, thus these results indicate that IBERIA will find it the most profitable to decrease the average fares by approximately 20% in these six markets with future HSR entry. Note that this study applies the same fare reductions for all six markets.



(a) (b)

Figure 6: (a) Model-predicted total IBERIA profit variation as a function of HSR frequency for different reductions in average fares by IBERIA alone; (b) Model-predicted total IBERIA profit variation as a function of fare changes for IBERIA alone for different HSR frequencies

Table 6 shows the predicted IBERIA frequency values for different HSR frequency values and different IBERIA fare reductions for La Coruña-Madrid and Madrid-La Coruña OD pairs (results are identical for both these OD pairs due to symmetric schedules). The results for the rest of the OD pairs showed similar patterns. Each column (among columns 2, 4, 5, 6, 7, 8, and 9) in Table 6 corresponds to a different HSR frequency (f_{rail}^{od}). Each row shows the model-predicted IBERIA frequency value, for the OD shown in the first column and for the IBERIA fare reduction shown in the third column in that row. The frequency value per fleet type is displayed in each cell of second row onward. For example, $0/1/10$ implies that there are 0 flight legs per week with A-319, 1 with A-320 and 10 with A-321 fleet type. The second column in the table considers the scenario before HSR entry.

These experiments have been performed assuming a fleet of size and composition matching that of IBERIA and assuming that the other airlines do not reduce their average fares. The results suggest that it might be advisable for IBERIA to make changes in its fleet to compete in markets with HSR entry. In the absence of HSR competition, the optimized schedule utilizes a heterogeneous fleet in the markets in the Galicia region (see Table 6). However, post-HSR entry, only the aircraft with the smallest number of seats are utilized in these markets especially in the scenarios without price reduction. This effect is also observed in the scenarios with price reductions. However, the price reductions slightly change the utilized fleet. For the scenarios with low HSR frequency and low price reductions, the optimized schedule utilizes a somewhat heterogeneous fleet. Nevertheless, as HSR competes with higher frequencies, the optimized fleet again almost exclusively uses the smallest-sized aircraft type. This suggests that the availability of smaller aircraft would make it possible for the airline carrier to compete more effectively against HSR.

Table 6: Predicted IBERIA frequency values for different HSR frequency values and different IBERIA fare reductions for La Coruña-Madrid and Madrid-La Coruña OD pairs assuming that the other airlines do not reduce their average fares

O.D\ f_{rail}^{od}	0	Price reduction (%)	7	14	21	28	35	42
La Coruña-Madrid	7/7/0	0	20/0/0	13/3/1	14/2/0	15/1/0	7/0/0	0/0/0
		10	7/9/5	7/9/3	7/7/0	7/6/0	7/2/0	12/0/0
		20	7/7/7	7/7/4	6/7/2	7/5/0	10/3/0	12/0/0
		30	10/6/1	7/4/2	7/4/0	7/0/0	7/0/0	7/0/0
		40	14/4/0	8/4/0	7/0/0	7/0/0	7/0/0	0/0/0
Madrid-La Coruña	7/7/0	0	20/0/0	13/3/1	14/2/0	15/1/0	7/0/0	0/0/0
		10	7/9/5	7/9/3	7/7/0	7/6/0	7/2/0	12/0/0
		20	7/7/7	7/7/4	6/7/2	7/5/0	10/3/0	12/0/0
		30	10/6/1	7/4/2	7/4/0	7/0/0	7/0/0	7/0/0
		40	14/4/0	8/4/0	7/0/0	7/0/0	7/0/0	0/0/0

6. Conclusions

The airline planning process involves solving problems such as frequency planning, timetable development and fleet assignment. All of them are directly related to passenger demand in that passenger demand influences the schedule decisions and the schedule influences passenger demand. Consequently, integration of appropriate customer behavior modeling into the schedule optimization models is needed in order to be able to solve these problems effectively. Moreover, competition from different modes and operators makes it difficult to estimate the number of passengers that an airline will capture.

We developed a tactical competition model for an airline considering multi-modal competition between air and high speed rail, and airline competition between legacy and low-cost carriers. At the core of our modeling approach was an integrated schedule optimization model that includes frequency planning, approximate timetable development, fleet assignment and passenger demand choice. The model accounts for passenger demand share competition, and captures the impacts of schedule decisions on passenger demand. We developed and estimated a nested logit model of multi-modal demand behavior and integrated it into the schedule optimization model.

We calibrated the nested logit model using real data provided by IBERIA (the major Spanish airline). Then, we tested the integrated optimization model in a real network from IBERIA including other air and rail transportation options in Spain. We conducted sensitivity analysis to demonstrate the choice of model parameters and the robustness of our modeling approach to small changes in the parameter values. We found that the actual decisions taken by the airline are reasonably close to those predicted by our model. We evaluated multiple scenarios involving the entry of high speed rail in some markets, and we accounted for the possibilities of demand stimulation and airfare reductions as a result of the new services. We also validated our results using out-of-sample validation data from markets that had an entry of high speed rail in the past. Contingent on the competing airlines’ and the high speed rail's offered attributes, the model predicts the optimal decisions for IBERIA in order to retain passengers and maximize profits. The model provides interesting insights into the schedule changes, fleet composition changes, and fare changes that will help the airline cope effectively with the entry of high speed rail.

In summary, we showed that the solution predicted by our integrated optimization model is close to the actual decisions being taken by the airline, implying that the current decision-making at IBERIA does take into account the multi-modal competition aspects. More importantly, our strong out-of-sample validation results indicate that the model is able to predict the airline response to HSR entry, and the resulting network performance metrics with a good level of accuracy. Consequently, the proposed modeling framework is attractive from the perspective of the airline operators. It allows them to plan better for the impending HSR entry by fine-tuning schedules, fleets and fares. The framework can facilitate careful evaluation of various scenarios (such as competitor actions, fleet changes, fare changes, etc.) before the actual HSR entry, allowing the airline to be better prepared to adapt to the changing competitive environment.

7. Acknowledgments

This research was supported by project grant TRA2011-27791-C03-01 by the "Ministerio de Economía y Competitividad, Spain". We want to thank Virot Ta Chiraphadhanakul (MIT) for his helpful comments when the research was conducted. We also want to thank the associate editor and the anonymous referees for their helpful comments on the paper.

References

Barnhart, C., Kniker, T., Lohatepanont, M. (2002). Itinerary-Based Airline Fleet Assignment. Transportation Science, 36(2), 199–217.

Barnhart, C., Cohn, A.M, Johnson, E.L., Klabjan, D., Nemhauser, G.L., Vance, P.H. (2003). Airline Crew Scheduling. In: Hall, R.W., Editor(s), Handbook of Transportation Science, International Series in Operations Research & Management Science 56, 517-560.

Barnhart, C and Cohn, A. (2004). Airline Schedule Planning: Accomplishments and Opportunities. Manufacturing & Service Operations Management, 6(1), 3-22.

Barnhart, C. (2009). Airline Schedule Optimization. In: Peter Belobaba, Amedeo Odoni and Cynthia Barnhart, Editor(s), The Global Airline Industry, Wiley. ISBN: 978-0-470-74077-4.

Barnhart, C., Farahat, A., Lohatepanont, M. (2009). Airline Fleet Assignment with Enhanced Revenue Modeling. Operations Research, 57(1), 231–244.

Beale, E.M.L, Tomlin, J.A. (1970). Special facilities in a general mathematical programming system for non-convex problems using ordered sets of variables. Mathematical Methods of Optimization. In Proceedings of the fifth international conference on operational research; Lawrence, J., Ed.; Tavistock Publications: London, England; pp 447-454.

Behrens, C., Pels, E. (2012). Intermodal competition in the London–Paris passenger market: High-Speed Rail and air transport. Journal of Urban Economics, 71, 278–288.

Belobaba, P. (2009). Overview of Airline Economics, Markets and Demand. In: Peter Belobaba, Amedeo Odoni and Cynthia Barnhart, Editor(s), The Global Airline Industry, Wiley. ISBN: 978-0-470-74077-4.

Cadarso, L., Marín, Á. (2013). Robust passenger oriented timetable and fleet assignment integration in airline planning. Journal of Air Transport Management, 26, 44–49.

Desaulniers, G., Desrosiers, J., Dumas, Y., Solomon, M. M., Soumis, F. (1997). Daily aircraft routing and scheduling. Management Science, 43(6), 841–855.

Dumas, J., Aithnard, F., Soumis, F. (2009). Improving the objective function of the fleet assignment problem. Transportation Research Part B: Methodological, 43(4), 466-475.

Ferropedia. (2014). Accessed October 2014. http://www.ferropedia.es/wiki/Tr%C3%A1ficos_AVE_y_LD_corredor_Barcelona-Madrid

Garcia, A., Luceño, B. (2010). Relación entre el precio básico medio del billete de tren con la velocidad media y la distancia recorrida por el viajero. Proceedings of Congress of Transport Engineering, Madrid 2010. ISBN: 978-84-96398-41-2.

Garrow, L. (2012). Customer Modeling. In: Barnhart, C. and Smith, B.C., Editor(s), Quantitative Problem Solving Methods in the Airline Industry, International Series in Operations Research & Management Science, Springer, 169, 1-34. ISBN 978-1-4614-1607-4.

Greene, W. H., Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological, 37(8), 681-698.

Grosche, T., Rothlauf, F., Heinzl, A. (2007). Gravity models for airline passenger volume estimation. Journal of Air Transport Management, 13, 175–183.

Guo, P., Xiao, B., & Li, J. (2012). Unconstraining methods in revenue management systems: Research overview and prospects. Advances in Operations Research, doi:10.1155/2012/270910.

Hane, C.A., Barnhart, C., Johnson, L.E., Marsten, R.E., Nemhauser, G.L., Sigismondi, G. (1995). The fleet assignment problem: solving a large scale integer program. Mathematical Programming, 70, 211-232.

Hansen, M. (1990). Airline competition in a hub-dominated environment: an application of noncooperative game theory. Transportation Research Part B, 248(I), 27-43.

Harsha, P. (2008). Mitigating airport congestion: Market mechanisms and airline response models. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Hong, S., Harker, P.T. (1992). Air traffic network equilibrium: toward frequency, price and slot priority analysis. Transportation Research Part B, 26B(4), 307-323.

Jacobs, T.L., Smith, B.C., Johnson, E.L. (2008). Incorporating Network Flow Effects into the Airline Fleet Assignment Process. Transportation Science, 42(4), 514–529.

Jehanno, A., Palmer, D., and James, C. (2011). High Speed Rail and Sustainability. International Union of Railways. http://www.uic.org/etf/publication/publication-detail.php?code_pub=531E

Lohatepanont, M., Barnhart, C. (2004). Airline Schedule Planning: Integrated Models and

Algorithms for Schedule Design and Fleet Assignment. Transportation Science, 38(1), 19-32.

Ministerio de Fomento, Spain. (2012). Plan de Infraestructuras, Transporte y Vivienda (PITVI 2012-2024). http://www.adif.es/es_ES/infraestructuras/planes_infraestructura/planes_infraestructura.shtml

Pita, J., Barnhart, C., Antunes, A.P. (2012). Integrated Flight Scheduling and Fleet Assignment Under Airport Congestion. Transportation Science, Articles in Advance, 1-16 http://dx.doi.org/10.1287/trsc.1120.0442.

Román, C., Espino, R., Martín, J.C. (2007). Competition of high-speed train with air transport: The case of Madrid-Barcelona. Journal of Air Transport Management, 13, 277-284.

Simpson, RW. (1970). A market share model for US domestic airline competitive markets. Working paper, MIT Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Vaze, V. (2011). Competition and Congestion in the National Aviation System: Multi-agent, Multi-stakeholder Approaches for Evaluation and Mitigation. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Vaze, V., Barnhart, C. (2012a). Modeling Airline Frequency Competition for Airport Congestion Mitigation. Transportation Science, 46(4), 512-535.

Vaze, V., Barnhart, C. (2012b). An Assessment of the Impact of Demand Management Strategies for Efficient Allocation of Airport Capacity. International Journal of Revenue Management, 6(1/2), 5-27.

Wei, W., M. Hansen. (2005). Impact of aircraft size and seat availability on airlines' demand and market share in duopoly markets. Transportation Research Part E: Logistics and Transportation Review, 41(3), 315-327.

Wei, W., Hansen, M. (2007). Airlines' competition in aircraft size and service frequency in duopoly markets. Transportation Research Part E: Logistics and Transportation Review, 43, 409-424.

Wen, C. H., Lai, S. C. (2010). Latent class models of international air carrier choice. Transportation Research Part E: Logistics and Transportation Review, 46(2), 211-221.

Wickham, R. R. (1995). Evaluation of forecasting techniques for short-term demand of air transportation. Report R95-7, Flight Transportation Laboratory, Massachusetts Institute of Technology.

Yan, S., Tang, C-H., Lee, M-C. (2007). A flight scheduling model for Taiwan airlines under market competitions. Omega, 35, 61-74.

Wang, D., Shebalov, S., Klabjan, D. (2012). Attractiveness-Based Airline Network Models with Embedded Spill and Recapture. Working paper (submitted for publication).

Zito, P., Salvo, G., La Franca, L. (2011). Modelling Airlines Competition on Fares and Frequencies of Service by Bi-level Optimization. Procedia – Social and Behavioral Sciences, 20, 1080-1089.