ISSN 2281-4299

Dipartimento di Ingegneria informatica automatica e gestionale Antonio Ruberti

Sapienza Università di Roma

The Sales Based Integer Program for Post-Departure Analysis in Airline Revenue Management: model and solution

Giorgio Grani
Gianmaria Leo
Laura Palagi
Mauro Piacentini
Hunkar Toyoglu

Technical Report n. 5, 2019

The Sales Based Integer Program for Post-Departure Analysis in Airline Revenue Management: model and solution

Giorgio Grani[a], Gianmaria Leo[b], Laura Palagi[a], Mauro Piacentini[c], Hunkar Toyoglu[d]

[a]*Sapienza University of Rome, Dep. of Computer Science, Control and Management Engineering, Rome, Italy*
[b]*IBM Analytics Data Science Elite, Munich, Germany*
[c]*Amazon EU SARL, Luxembourg, Luxembourg*
[d] *Sabre Airline Solutions, Operations Research Consulting, Southlake, TX, United States*

Abstract

Airline revenue management (RM) departments pay remarkable attention to many different applications based on sales-based linear program (SBLP). SBLP is mainly used as the optimization core to solve network revenue management problems in RM decision support systems. In this study we consider a post-departure analysis, when there is no more stochasticity in the problem and we can tackle SBLP with integrality constraints on the variables (SBIP) in order to understand which should be the best possible solution.
We propose a new formulation based on a market-service decomposition that allows to solve large instances of SBIP using LP-based branch-and-bound paradigm. We strengthen the bound obtained with the linear relaxations by introducing effective Chvátal-Gomory cuts. Main idea is to optimally allocate the capacity to the markets by transforming the market subproblems into a piecewise linear objective function. Major advantages are significant reduction of the problem size and the possibility of deriving a concave objective function which is strengthened dynamically. Numerical results are reported. Providing realistic integral solutions move forward the network revenue management state of the art.

Keywords: revenue management, mixed-integer programming, decomposition, airline, piecewise linear

---

Email addresses: `g.grani@uniroma1.it` (Giorgio Grani), `gianmaria.leo@ibm.com` (Gianmaria Leo), `laura.palagi@uniroma1.it` (Laura Palagi), `piacm@amazon.lu` (Mauro Piacentini), `hunkar.toyoglu@sabre.com` (Hunkar Toyoglu)

1. Introduction

Network revenue management (RM) has been playing an increasingly crucial role in both strategic and tactical decisions of airlines over the recent years. Successful RM processes aim to achieve the revenue maximization by leveraging huge amount of data, upcoming technologies and more sophisticated approaches to measure the RM performance. This leads top carriers to invest millions of dollars every year to face the challenge of catching new revenue opportunities. An overview of RM can be found at Talluri and Van Ryzin (2006), McGill and Van Ryzin (1999) and Walczak et al. (2012). RM models present some drawbacks, for example the large dimension of the network, stochasticity of the environment and difficulty of fitting the real behavior of customers. Furthermore, product demand is uncertain, so as the demand dependencies and competition among carriers.

One of the latest challenges in RM problems is to incorporate a costumer choice model (CCM) into the RM problem statement, to capture customer behavior with respect to airline's booking strategy. Indeed CCMs are fundamental tools to include in the RM analysis because the assumption that the demand for any product is independent of the availability of other products is no more viable because flight options become interchangeable due to possible *recapture* and/or *spill* effects. (see e.g. Strauss et al. (2018), Vulcano et al. (2010)). Recapture refers to redirected demand to an available product different from the unavailable first choice of the customer rather than to buy up to a higher fare on the same flight. Spill refers to lost demand due to unavailability. Spilled demand can either (i) be recaptured on other available alternatives according to their attractiveness or (ii) are lost to competition or no-purchase option. The two principal CCMs that are used in RM are the basic attraction model (BAM, see Luce (1959)) and the general attraction model (GAM, see Gallego et al. (2015)) that differ in considering product specific shadow attractions and hence in evaluating the attraction value of the no-purchase alternative.

Two main RM models including CCMs in their formulation have been proposed in the literature: choice-based and sales-based Programs. Choice-based linear program (CBLP) was originally proposed by Gallego and Phillips (2004) and Liu and Van Ryzin (2008) for the network RM problem. In the (CBLP), the decision variables represent the portion of time that a certain offer set of possible flight options is open for booking. If n is the number of possible flight options, in the worst case there are 2^n offer sets (corresponding to variables), including the empty one, thus leading to hard LP problems due to the exponential number of variables. Methods trying to circumvent the size of the problem include e.g. column generation (see Bront et al. (2009)), heuristic methods (see Kunnumkal and Topaloglu (2008), Kunnumkal and Topaloglu (2010), Liu and Van Ryzin (2008)), simulation as proposed in Van Ryzin and Vulcano (2008), linear program reductions in Vossen and Zhang (2015) or relaxation techniques as in Kunnumkal and Talluri (2015) for the dynamic programming case and Meissner et al. (2013) for CBLP.

Sales-based programs (SBPs) are more recent models that are first proposed in Guillermo et al. (2011) and Gallego et al. (2015). In SBPs the decision variables are the seats that can be allocated. The number of variables is polynomial in the network size. If n is the number of flight options the model will have $n + 1$ variables, including the possibility of rejecting all the options. Gallego et al. (2015) consider the linear formulation SBLP and proved that it is equivalent to CBLP by showing both formulations have same objective value at the optimum. Thanks to this equivalence, SBLP became the preferred optimization model in many airline network RM problems.

SBLP is mainly used to solve big airline network RM problems with possiblly different objectives. For example, one objective could be to find out what could have been done better to increse the revenue last week so that the RM strategy can be improved for future. Briefly the usage is a post-departure analysis to find out improvement opportunities. This usage is based on the solution of a linear model that is representing the integer seat sales and hence such linear solutions may deviate from reality. This in return may have an adverse impact on future decisions an airline may adopt to improve its revenue and RM performance. After several years of real life implementation, it has been observed that there is benefit in improving the network RM linear models to obtain integer solutions for better accuracy and real life representation and for eliminating impractical solutions and thus recommendations. These aspects are explained better with the help of a toy example in the Section 3.

Hence in this paper we address the sales based integer programming (SBIP) model instead of linear SBLP. We further note that in post-departure analysis scenarios we don't need to perform real time computations so that the computational time, although limited, does not represent a hard restriction. Many approaches have been proposed in the literature to solve large instances of RM problems. Among these approaches decomposition is not new in RM and often it is carried out at leg level, or origin and destination (O&D) network level. Leg based decomposition approaches have been proposed in Smith and Penn (1988) and Smith et al. (1992). They represent the network problem as multiple independent single flight problems. They focus on local optimization of single O&D flights and hence lose information about the network effects due to shared capacity on multiflight itineraries. A more recent approach of Birbil et al. (2013) consists of a two-stage O&D based decomposition method which first allocates network capacity to multiflight O&D pairs and then optimizes over the itinerary as it is a single flight problem.

We propose a decomposition based on market-service concept rather than leg or O&D based decomposition. Our main goal is to introduce effective heuristic methods that can solve large-scale network SBIP with a low optimality gap. Unfortunately the subproblems obtained by our decomposition are still hard because they are still nonlinear integer programs. However, it is possible to define a piecewise linear concave approximation to make the problems tractable. The first preliminary results obtained by substituting linear SBLP with integer SBIP, although solved approximatively, seems to be quite promising and gives better indication on advisable policy adjustments.

Before we dig deep into details it is necessary to define why such integer problem is of interest in RM. As already mentioned classical network RM models have been adopted in order to control booking limits in real time applications. In this framework

demands and willingness to pay of customers must be seen as random variables and so it would seem unreasonable to include integrality when talking of expected values. This is not the case, since our model is useful after-sales, when all the quantities have been observed. SBIP has to be seen as a model to evaluate a revenue performance. In other words we would like to understand how good our pricing strategy was. Commercial softwares use SBLP as an ideal approximation of the best revenue, but since there is no more stochasticity we need to consider the variables as actual seat, and not expected ones.
This paper is organized as follows. In section 2 we investigate the motivations that moved the idea of SBIP. Section 3 describes the standard model settings and terminology for the sales-based integer program (SBIP). In section 4 we propose the market-service decomposition approach and explain different approaches, properties and developments. Finally in section 5 we show some empirical results.

2. Motivations regarding the model

The Sales Based Linear Program (SBLP) has been introduced in Gallego et al. (2010) as a formulation able to overcome the typical intractability of Choice Based models. It is widely used in the industry since it can incorporate Customer Choice models (CCMs) in the formulation. The model finds out the best expected booking limits with respect to expected revenue and it is used directly in the booking control process.
In post-departures analysis however it could be important to evaluate the RM strategy adopted. By the structure of the industry it is not easy to estimate good benchmarks from other companies, especially when addressing daily revenues. In this situation it turns out to be fundamental the possibility to evaluate revenues based on the observed demands. In other words we could ask to an optimization model to calculate the best possible combination of sold tickets, according to the observed demands and the observed willingness. From this point of view the new model do not need to tackle stochasticity at all, since all the parameters are realized ones.
It is indeed unreasonable to think of a solution of such problem as a continuous one, since the optimal selling strategy can only distribute seats among fare classes in a discrete way. In section 5 some results are reported supporting the misleading optimistic high revenues obtained from using a linear model instead of an integer one.
It is not our intention to underestimate the importance of SBLP in RM. In fact SBIP cannot address real time booking problems, since it is not giving back expected revenues and it is not tackling stochastic problems. Furthermore SBIP cannot return bid prices, since it is an integer problem. SBIP usage must rely on after-sales or post-departure analysis only.

3. Sales-Based Integer Program

The sales-based linear program (SBLP) was introduced by Gallego et al. (2015) and further studied in Liang et al. (2017) as an effective reformulation of deterministic choice-based revenue management problem.

As we mentioned in the introduction we are interested in the sales-based integer program (SBIP) which is the integer version of the SBLP. We also focus on the post-departure analysis scenario that aims to find the optimal (integer) number of seats that should have been allocated to get the best revenue with full awareness of the actual demand and hence we use BAM as CCM. In this section we report the SBIP formulation as presented in Gallego et al. (2015) to introduce the main definitions and notations. At the end of the section we report a simple example to explain the fundamental role of integrality within network RM solutions.

Flights and seats are the main resources of the host airline to be sold or allocated to passengers. A *Flight* f is defined as the direct connection between a certain origin and destination. It is fully characterized by flight number, origin and destination and departure date. Let $\mathcal{F}$ be the set of all flights operated or at least marketed by the airline.

In order to be more accurate, we introduce also the differentiation of seats depending on the cabins. Indeed flight seats are differentiated respect to different cabins b in a set $\mathcal{B}_f$ (e.g. business and economy cabins). In particular, each cabin $b \in \mathcal{B}_f$ has its own capacity c_b, meaning that only c_b seats are available for sale in cabin b.

Flights may be combined into itineraries to allow more complex travels for the passengers going from one origin to one destination, with possible intermediate stops and served by a sequence of flights. A *Service* identifies an itinerary O&D and it is defined by the sequence of flights. The service set $\mathcal{S}$ refers to all relevant airline itineraries within a year.

As passengers are different not only in terms of itinerary, but also in travel purposes and willingness to pay, it is important to differentiate the seat product in terms of itinerary, travel restrictions and price. To this aim we assume that there exist customer market segments and that customers in a particular market segment are interested in a specific set of options. In particular a *Market* m is defined as the specific group of travelers characterized by a specific trip, itinerary O&D, point-of-sale, cabin, product group (e.g. normal, frequent fliers, staff or groups), departure date and time period before departure for completing the reservation (*booking period*). Set $\mathcal{M}$ denotes all possible markets.

Markets play a fundamental role in the formulation addressed in this paper. In particular, $\mathcal{M}_b \subseteq \mathcal{M}$ is the subset of markets referring to cabin $b \in \mathcal{B}_f$ for a fixed $f \in \mathcal{F}$, and $\mathcal{M}_{fb} \subseteq \mathcal{M}$ is the subset of markets sharing a flight f and cabin b.

To discriminate the price within the same market m, airline seats are used to define inventory products called booking classes or simply *classes*, differentiated by fare and buying restrictions (e.g. advanced purchase and length of stay)). An *Alternative* a within a *market* refers to a specific service and corresponding booking class for a specific market m. Each alternative is characterized by the selling price p_a, first choice demand d_a (in the sense that we are not taking into account recapturing effects) which is assumed to be deterministic, and an inherent attractiveness $\zeta_a \geq 0$ depending on itinerary quality and price.

Customers in market segment m select from $\mathcal{A}_m$ and the nopurchase (or null) alternative

$\bar{a}_m$ which aggregates all the no-host and no-fly alternatives in the market. Note that null alternative is assumed to be always available for sale. The total available demand d_m within market m results from the sum of all alternative demands, namely

$$d_m = \sum_{a \in \mathcal{A}_m} d_a + \bar{d}_m \tag{1}$$

where we denote as $\bar{d}_m$ the lost demand associated with the null alternative $\bar{a}_m$ with attractiveness $\bar{\zeta}_m$. Set $\mathcal{A}$ defines all possible market alternatives, namely $\mathcal{A} = \bigcup_{m \in \mathcal{M}} \mathcal{A}_m$.

Given the set of available alternatives in a market $\hat{\mathcal{A}}_m \subseteq \mathcal{A}_m$, BAM defines the probability of selection of the alternative a as

$$\rho_a(\hat{\mathcal{A}}_m) = \begin{cases} \dfrac{\zeta_a}{\sum_{a' \in \hat{\mathcal{A}}_m} \zeta_{a'} + \bar{\zeta}_m} & \text{if } a \in \hat{\mathcal{A}}_m \\ 0 & \text{otherwise} \end{cases}$$

Attraction and demand are assumed to be input parameters of the optimization model. Actually, they can be estimated by applying expectation-maximization algorithm to the incomplete data log-likelihood function, as it is explained in Vulcano et al. (2012). Note that estimation algorithms require only historical data for the host airline, such as bookings, inventory controls and market share.

Now we introduce SBIP formulation. Main decision variables are the sales allowed for different market alternatives, namely $x_a \in \mathbb{N}, \quad \forall\, a \in \mathcal{A}$. Indirectly, it is important to keep track of the total market demand recaptured by the null market alternatives. For this purpose simple continuous variables can be added, namely $z_m, \quad m \in \mathcal{M}$. Note that continuity is not restrictive as sales on null market alternatives do not correspond to real physical resources.

Market alternative allocation should be consistent with the available capacity. For each flight f and for each cabin b, total number of seats allocated can not exceed the capacity of cabin b.

$$\sum_{m \in \mathcal{M}_{fb}} \sum_{a \in \mathcal{A}_m} x_a \leq c_b, \quad \forall f \in \mathcal{F}, \forall b \in \mathcal{B}_f$$

Given a market $m \in M$, BAM spill and recapture effects require the introduction of total demand constraints and relative demand constraints.

Total market demand constraints guarantee that the market demand is allocated over all market alternatives including the null one, namely

$$\sum_{a \in \mathcal{A}_m} x_a + z_m = d_m \qquad \forall\, m \in \mathcal{M} \tag{2}$$

Relative demand constraints allow market alternative sales to exceed their corresponding demand to let market alternatives recapture part of the total market demand's spill in proportion to their attractiveness, namely

$$\bar{\zeta}_m x_a - \zeta_a z_m \le 0, \quad \forall\, a \in \mathcal{A}_m \quad \forall\, m \in \mathcal{M} \tag{3}$$

We easily show that the variables are bounded,

$$x_a \le \frac{\zeta_a}{\bar{\zeta}_m}\, z_m \le \frac{\zeta_a}{\bar{\zeta}_m}\, d_m, \quad \forall\, a \in \mathcal{A}_m \quad \forall\, m \in \mathcal{M}$$

We also introduce a lower bound on z_m

$$z_m \ge \bar{d}_m, \quad \forall\, m \in \mathcal{M} \tag{4}$$

The role of this bound is to force the lowest value of the null alternative to match the actual null alternative realization, so that the model reflects better the actual demand. Without loss of generality $\bar{d}_m$ can be replaced with any positive value and the following results still hold.

The objective function maximizes the total revenue

$$\max \sum_{m \in \mathcal{M}} \sum_{a \in \mathcal{A}_m} p_a\, x_a$$

Finally, SBIP corresponds to the following mixed integer linear problem

$$\begin{cases} \max\limits_{\mathbf{x} \in \mathbb{N}^{|\mathcal{A}|}, \mathbf{z} \in \mathbb{R}^{|\mathcal{M}|}} & \sum\limits_{m \in \mathcal{M}} \sum\limits_{a \in \mathcal{A}_m} p_a\, x_a & \\ & \sum\limits_{m \in \mathcal{M}_{fb}} \sum\limits_{a \in \mathcal{A}_m} x_a \le c_b, & \forall f \in \mathcal{F}, \forall b \in \mathcal{B}_f \\ & \zeta_{\bar{a}_m} x_a - \zeta_a z_m \le 0, & \forall\, m \in \mathcal{M}, \forall a \in \mathcal{A}_m \\ & \sum\limits_{a \in \mathcal{A}_m} x_a + z_m = d_m, & \forall\, m \in \mathcal{M} \\ & z_m \ge \bar{d}_m, & \forall\, m \in \mathcal{M} \end{cases} \tag{5}$$

where

$$\mathbf{x}_{\mathcal{A}_m} = \begin{pmatrix} x_{a_1} \\ \vdots \\ x_{a_{|\mathcal{A}_m|}} \end{pmatrix} \in \mathbb{N}^{|\mathcal{A}_m|} \qquad \mathbf{x} = \begin{pmatrix} \mathbf{x}_{\mathcal{A}_{m_1}} \\ \vdots \\ \mathbf{x}_{\mathcal{A}_{m_{|\mathcal{M}|}}} \end{pmatrix} \in \mathbb{N}^{|\mathcal{A}|} \qquad \mathbf{z} = \begin{pmatrix} z_{m_1} \\ \vdots \\ z_{m_{|\mathcal{M}|}} \end{pmatrix} \in \mathbb{R}^{|\mathcal{M}|}$$

The SBIP is the integer version of the BAM version of SBLP proposed by Gallego et al. (2015).

Our goal is to exploit the hidden structure and properties of SBIP to define decomposition and relaxation techniques, which maintain integrality on sales variables and solve large-scale network RM problems effectively. In the next section we explain on a toy example the role of integrality in airline network RM problems especially in a post-departure analysis scenario.

3.1. SBLP versus SBIP

As mentioned in section 1 we consider that the SBIP solution is used in a post-departure analysis scenario to identify what could have been done better in the past to improve the revenue. Such an analysis takes as input data the inventories, the fares and the actual reservations in a certain time window and uses these values to define a deterministic sales-based model to obtain the optimal revenue corresponding to the effective realization of the past demand over the observed periods. The optimal solution represents an optimal sale RM strategy with the given realization of the demand. Then the revenue results that could have been obtained by using this optimal strategy is compared (with full awareness of the demand) with the one actually realized by the carrier. The gap in the performance is used to gain insights about how much the future revenue can be improved and as a matter of fact it is used to adjust strategic and tactical decision of the airline.

We explain the main difference between SBLP and SBIP models by using a toy example. Let us assume that airline XX operates in a one directional market, say A-B, and sells only two alternative services from A to B, say x_1 and x_2. Let z denote the lost demand. In the upcoming sales period, say the next month, Airline XX wants to maximize its expected revenue. To this aim Airline XX builds a model that returns the probability of selling the two alternatives by using a SBLP model where the coefficients are retrieved from the demand forecast over the upcoming sales period. Now, let us assume that one month is over, and Airline XX wants to benchmark its RM strategy. In other words, XX wants to estimate the theoretical maximum revenue that it would have earned if the actual demand of the last period had been known in advance. The airline can estimate the attractivity of its two services by using the actual reservations achieved over the last month and a deterministic model can be built. Hence, XX solves an integer model SBIP (sales volume is integer for both services) by leveraging the attractivity coming from the actual reservations obtained in the last period. It is easy to check that the SBIP application naturally leads to a deterministic problem. In other words we are answering the question: which booking would you accept if you knew the exact values of the service demands? For example assume that the model is given by

$$\begin{cases} \max & x_1 + 10x_2 \\ & x_1 + x_2 + z = 40 \\ & x_1 \leq 2.1z \\ & x_2 \leq 0.9z \\ & z \geq 10 \\ & x_1 \geq 0,\ x_2 \geq 0 \end{cases}$$

Solving to optimality the linear relaxation (SBLP) gives the optimal solution $(\hat{x_1}, \hat{x_2}, \hat{z}) = (0,\ 18.94,\ 21.05)$ with maximum expected revenue $\hat{x_1} + 10\hat{x_2} \simeq 189.47$. However this solution is not giving the correct information about the true value of the optimal revenue. Approximating to the nearest integer will give $(0,\ 18,\ 22)$ with a revenue of 180. An important metric of a post-departure analysis is the difference between the actual sales and revenue obtained by the actual RM system in the observed period and the optimum revenue given by the sales-based models. This difference is used to gain insights into how much the future revenue can be improved and as a matter of fact it is used to adjust the strategic and tactical decisions of the airline. For example, suggesting fractional seat sales (e.g. sell 1.6 or 0.3 seats) may not be realistic. In fact solving the integer version will lead to the choice $(x_1^*, x_2^*, z^*) = (2, 18, 20)$, with objective value $x_1^* + 10x_2^* = 182$. We aim to reach realistic integral sales to enable airlines implement better tactical adjustments, strategic counter measures, etc. to improve its revenue and RM performance.

4. Decomposition Approach

Problem (5) can not be solved exactly for medium and large network RM problems. In order to cope with size limits, several decomposition approaches have been proposed in the literature, e.g. Birbil et al. (2013). The general idea is to decompose by flight or by O&D and then to optimize each single sub-problem using standard solution algorithms. Nevertheless, integrality is relaxed in favor of computational efficiency.

In our work, decomposition will be defined at market-service level. In this way, integrality will be preserved at a more granular level. The new solution approach is a master-slave decomposition, where all the slave problems can be solved a priori before the master.

In Subsection 4.1 the approach will be explained under the assumption that each *O&D* is served by at most one service itinerary, that leads to a market decomposition. The restrictive assumption is then removed in Subsection 4.2 where the market-service decomposition is finally described with all the theoretical and practical aspects.

4.1. Market Decomposition

In this section, market decomposition will be presented under the following assumption.

Assumption 4.1. *For each $m \in \mathcal{M}$ there exists a unique $s \in \mathcal{S}$ such that $s_a = s$ for all $a \in \mathcal{A}_m$.*

In other words, for a given market m, there is only one specific service itinerary connecting the specific *O&D* associated to the market.

Under this setting, main idea is to move from a market alternative allocation to a market allocation, so that all complexity related to the alternatives is hidden in the network problem. The corresponding problem is to allocate the available capacity to the markets generating more revenue. Of course, in order to maximize the network revenue it is important to find the revenue resulting from a given market allocation. The overall problem can be decomposed in a master-slave framework.

- *Master Problem* - Allocate seats to markets such that the overall network revenue is maximized.
- *Slave Problem* - Distribute allocated seats within the different alternatives such that the overall market profit is maximized.

Decomposition requires market allocation variables, namely $v_m \in \mathbb{N}, \quad \forall\, m \in \mathcal{M}$, and market allocation constraints

$$\sum_{a \in \mathcal{A}_m} x_a = v_m \in \mathbb{N}, \quad \forall\, m \in \mathcal{M}$$

Introducing such variables and constraints, problem (5) can be equivalently rewritten as

$$\left\{\begin{array}{lll}
\max\limits_{\mathbf{x}\in\mathbb{N}^{|\mathcal{A}|}, \mathbf{z}\in\mathbb{R}^{|\mathcal{M}|}, \mathbf{v}_{\mathcal{M}}\in\mathbb{N}^{|\mathcal{M}|}} & \sum\limits_{m\in\mathcal{M}} \sum\limits_{a\in\mathcal{A}_m} p_a\; x_a & \\
& \sum\limits_{a\in\mathcal{A}_m} x_a = v_m, & \forall\, m \in \mathcal{M} \\
& \sum\limits_{m\in\mathcal{M}_{fb}} v_m \leq c_b, & \forall f \in \mathcal{F}, \forall b \in \mathcal{B}_f \\
& \sum\limits_{a\in\mathcal{A}_m} x_a + z_m = d_m, & \forall\, m \in \mathcal{M} \\
& \zeta_{\bar{a}_m} x_a - \zeta_a z_m \leq 0, & \forall\, m \in \mathcal{M}, \forall a \in \mathcal{A}_m \\
& z_m \geq \bar{d}_m, & \forall\, m \in \mathcal{M} \\
& v_m \leq d_m - \bar{d}_m, & \forall\, m \in \mathcal{M}
\end{array}\right.$$

where

$$\mathbf{v}_{\mathcal{M}} = \left(v_{m_1}, \;\cdots\;, v_{m_{|\mathcal{M}|}}\right)^{\intercal}.$$

The last constraints that set an upper bound on the market allocation capacity v_m are redundant. We leave them in the formulation to maintain the fact that the host airline can never recapture from the null alternatives.

In order to get a more compact representation, the following sets and function are introduced

- $\mathcal{G} = \left\{ \begin{array}{lll} (\mathbf{x}, \mathbf{z}) \in \mathbb{N}^{|\mathcal{A}|} \times \mathbb{R}^{|\mathcal{M}|} & : & \sum\limits_{a\in\mathcal{A}_m} x_a + z_m = d_m,\ \forall\, m \in \mathcal{M} \\ & & \zeta_{\bar{a}_m} x_a - \zeta_a z_m \leq 0,\ \forall\, m \in \mathcal{M},\ \forall\, a \in \mathcal{A}_m \\ & & z_m \geq \bar{d}_m,\ \forall\, m \in \mathcal{M} \end{array} \right\}$
- $\mathcal{V} = \left\{ \begin{array}{lll} \mathbf{v}_{\mathcal{M}} \in \mathbb{N}^{|\mathcal{M}|} & : & \sum\limits_{m\in\mathcal{M}_{fb}} v_m \leq c_b,\ \forall\, f \in \mathcal{F},\ \forall\, b \in \mathcal{B}_f \\ & & v_m \leq\ d_m - \bar{d}_m,\ \forall\, m \in \mathcal{M} \end{array} \right\}$
- $\mathcal{S}(\mathbf{v}_{\mathcal{M}}) = \left\{\ \mathbf{x} \in \mathbb{N}^{|\mathcal{A}|}\ :\ \sum\limits_{a\in\mathcal{A}_m} x_a = v_m,\ \forall\, m \in \mathcal{M}\ \right\}$
- Given $m \in \mathcal{M},\ f_m(\mathbf{x}) = \sum\limits_{a\in\mathcal{A}_m} p_a\; x_a$

The problem can now be rewritten in compact form as follows

$$\begin{cases} \max\limits_{\mathbf{x},\mathbf{z},\mathbf{v}_{\mathcal{M}}} & \sum\limits_{m\in\mathcal{M}} f_m(\mathbf{x}) \\ & \mathbf{v}_{\mathcal{M}} \in \mathcal{V} \\ & (\mathbf{x},\mathbf{z}) \in \mathcal{G} \\ & \mathbf{x} \in \mathcal{S}(\mathbf{v}_{\mathcal{M}}) \end{cases} \tag{6}$$

We have the following result that allows a two-stage decomposition.

Proposition 4.2. *Any optimal solution of the two stage problem*

$$\max_{\mathbf{v}_{\mathcal{M}}\in\mathcal{V}} \max_{(\mathbf{x},\mathbf{z})\in\mathcal{G}\cap\mathcal{S}(\mathbf{v}_{\mathcal{M}})} \sum_{m\in\mathcal{M}} f_m(\mathbf{x}) \tag{7}$$

is optimal for (6).

The proof is in Appendix 7.
Next step is to move from a two stage problem to Master - multi Slaves decomposition. The decomposition at market level requires a finer setting. Given market $m \in \mathcal{M}$, the following specific sets can be defined

- $\mathcal{G}_m = \left\{ \begin{array}{ll} (\mathbf{x}_{\mathcal{A}_m}, z_m) \in \mathbb{N}^{|\mathcal{A}_m|} \times \mathbb{R} \ : & \sum\limits_{a\in\mathcal{A}_m} x_a + z_m = d_m, \\ & \zeta_{\bar{a}_m} x_a - \zeta_a z_m \leq 0, \ \forall \ a \in \mathcal{A}_m \\ & z_m \geq \bar{d_m} \end{array} \right\}$
- $\mathcal{S}_m(v_m) = \left\{ \ \mathbf{x}_{\mathcal{A}_m} \in \mathbb{N}^{|\mathcal{A}_m|} \ : \ \sum\limits_{a\in\mathcal{A}_m} x_a = v_m \ \right\}$

Note that it holds that $\mathcal{G} = \prod\limits_{m\in\mathcal{M}} \mathcal{G}_m$.
At this point, slave problem in (7) can be viewed as the combination of separable subproblems defined at market level. Therefore, problem (7) can be finally rewritten as Master - multi Slave problem.

$$\max_{\mathbf{v}_{\mathcal{M}}\in\mathcal{V}} \sum_{m\in\mathcal{M}} \max_{(\mathbf{x}_{\mathcal{A}_m}, z_m)\in\mathcal{G}_{\Updownarrow}\cap\mathcal{S}_m(v_m)} f_m(\mathbf{x}) \tag{8}$$

For a given market and allocation, the inner problem leads to the following definition.

Definition 4.3. *Given a market $m \in \mathcal{M}$ and a feasible allocation $v_m \in \mathbb{N}$, the* ***Market Revenue Function*** *$p_m(v_m)$ is defined as*

$$p_m(v_m) := \max_{(\mathbf{x}_{\mathcal{A}_m}, z_m)\in\mathcal{G}_{\Updownarrow}\cap\mathcal{S}_m(v_m)} f_m(\mathbf{x}) \tag{9}$$

Under this notation, the *Master problem* is

$$\max_{\mathbf{v}_{\mathcal{M}} \in \mathcal{V}} \sum_{m \in \mathcal{M}} p_m(v_m) \tag{10}$$

Master problem is just defined over the market allocation decisions restricted by the available demand and seat flight capacities. The complexity of the original problem is hidden in the objective function, defined as the sum of the market revenue functions. The difficulty of solving problem (5) is moved to the Master problem (10), as the latter problem is a mixed integer nonlinear with nonlinear and non-differentiable objective function. Hence, it is convenient to slightly relax the problem in order to define a fast solution approach that can find a feasible solution of the Master problem (10) with a good optimality gap.

Given a market $m \in \mathcal{M}$ and a feasible allocation $v_m \in \mathbb{N}$, the value of $p_m(v_m)$ is the optimal value of the *Slave problem*

$$\max_{(\mathbf{x}_{\mathcal{A}_m}, z_m) \in \mathcal{G}_m \cap \mathcal{S}_m(v_m)} f_m(\mathbf{x}) \tag{11}$$

Although the Slave problem is an ILP, the problem is well-structured and easily solvable. Slave problem (11) is equivalent to a simple budgeting problem whose solution can be found by a greedy-type algorithm. Actually, z_m can be removed from the problem setting

$$z_m = d_m - \sum_{a \in \mathcal{A}_m} x_a = d_m - v_m,$$

so that the feasible set reduces to

$$\mathcal{G}_m \cap \mathcal{S}_m(v_m) = \left\{ \mathbf{x}_{\mathcal{A}_m} \in \mathbb{N}^{|\mathcal{A}_m|} : \sum_{a \in \mathcal{A}_m} x_a = v_m,\ x_a \leq \frac{\zeta_a}{\zeta_{\bar{a}_m}} (d_m - v_m), \forall a \in \mathcal{A}_m \right\} \tag{12}$$

Note that the formulation (11) can be improved by forcing the integrality $x_a \in \mathbb{N}$ in (12), namely by strengthening the upper bound as

$$\mathcal{G}_m \cap \mathcal{S}_m(v_m) = \left\{ \mathbf{x}_{\mathcal{A}_m} \in \mathbb{N}^{|\mathcal{A}_m|} : \sum_{a \in \mathcal{A}_m} x_a = v_m,\ x_a \leq \left\lfloor \frac{\zeta_a}{\zeta_{\bar{a}_m}} (d_m - v_m) \right\rfloor, \forall a \in \mathcal{A}_m \right\} \tag{13}$$

At this point, it is possible the describe the Greedy-type algorithm scheme for the computation of the solution for the Slave problem (11) using (13).

Slave solution Greedy Algorithm

Initialization. Reorder the alternatives as $\mathcal{A}_m = \left\{a_{j_1}, a_{j_2}, \ldots, a_{j_{|\mathcal{A}_m|}}\right\}$ so that

$$p_{a_{j_1}} \geq p_{a_{j_2}} \geq \ldots \geq p_{a_{j_{|\mathcal{A}_m|}}}.$$

Let be $x_{a_{j_i}} = 0$ for all $i = 1 \ldots, |\mathcal{A}_m|$.

for $h = 1, 2, \ldots, |\mathcal{A}_m|$ **do**

$$x_{a_{j_h}} = \min\left\{ \left\lfloor \frac{\zeta_{a_{j_h}}}{\zeta_{\bar{a}_m}} \ (d_m - v_m) \right\rfloor, \ v_m - \sum_{i<h} x_{a_{j_i}} \right\}$$

If $v_m - \sum_{i \leq h} x_{a_{j_i}} = 0$ then **Exit**

end

Stop. Return $x_{a_{j_i}}$ for all $i = 1 \ldots, |\mathcal{A}_m|$.

Finally, Market Revenue function presents some interesting properties.

Proposition 4.4. *Given a market $m \in \mathcal{M}$, for any $v, v+1 \in \mathcal{V}$, it holds*

$$\frac{p_m(v+1)}{v+1} \leq \frac{p_m(v)}{v}$$

The proof is in Appendix 8.

Roughly speaking the Marginal Market Revenue function is non-increasing. An immediate consequence is the fact that the Market Revenue Function is Lipschitz continuous.

Proposition 4.5. *Given a market $m \in \mathcal{M}$, the Market Revenue function $p_m(v_m)$ is Lipschitz within feasible market allocation v_m.*

The proof is in Appendix 9. The two propositions describe the fact that we can find concave approximations of the Market Revenue function.

4.2. Market-Service Decomposition

In this section we neglect assumption (4.1). Indeed, we discuss how the model can be modified to solve multiple services over market m.

First of all, we have to redefine the Slave Problems to take into account multiple services. Let $\mathcal{S}_m := \{s_1, \ldots, s_{q_m}\}$ be the set of services associated with market m for all $m \in \mathcal{M}$ and $\mathcal{S}_{fm} \subseteq \mathcal{S}_m$ be the subset of services that use flight $f \in \mathcal{F}$.

In the following, for the sake of simplicity we remove the dependence of q from the market m. Let $\mathcal{P}_m := \{\mathcal{P}_{s_1}, \ldots \mathcal{P}_{s_q}\}$ be a service based partition of alternatives $\mathcal{A}_m$ associated to market m namely $\mathcal{P}_{s_i} \subseteq \mathcal{A}_m$, $i = 1, \ldots, q$, $\mathcal{P}_{s_i} \cap \mathcal{P}_{s_h} = \emptyset$, for all $i, h \in \{1, \ldots, q\}$ $\ i \neq h$ and $\cup_{i=1}^{q} \mathcal{P}_{s_i} = \mathcal{A}_m$. For any fixed market $m \in \mathcal{M}$ we have $\mathcal{P}_{s_i} = \{a \in \mathcal{A}_m \ : \ s_a = s_i\}$. For ease of exposition we have removed the dependency of $\mathcal{P}_{s_i}$ from m.

Following the same line as in Section 4.1 we define the Multi-Service Market Revenue Function. The capacity v_m allocated to market m is split among the services. We denote with $w_{s_i m}$ the capacity allocated to each service $s_i \in \mathcal{S}_m$.
By definition we have

$$v_m = \sum_{s_i \in \mathcal{S}_m} w_{s_i m} \tag{14}$$

We need to modify the constraints defining sets $\mathcal{V}$ and $\mathcal{S}$ in the single service case.
To this aim, let us introduce the following vectorial notation

$$\mathbf{w}_{\mathcal{S}_m} = \begin{pmatrix} w_{s_1 m} \\ \vdots \\ w_{s_q m} \end{pmatrix} \qquad \mathbf{w} = \begin{pmatrix} w_{\mathcal{S}_{m_1}} \\ \vdots \\ w_{\mathcal{S}_{|\mathcal{M}|} m} \end{pmatrix} \qquad \eta = \sum_{m \in \mathcal{M}} q_m$$

Introducing such variables and constraints and taking into account the fact that $\mathcal{M}_b \subseteq \mathcal{M}$ is the subset of markets referring to cabin $b \in \mathcal{B}_f$ for any $f \in \mathcal{F}$, problem (5) can be equivalently rewritten as

$$\begin{cases} \max\limits_{\mathbf{x} \in \mathbb{N}^{|\mathcal{A}|}, \mathbf{z} \in \mathbb{R}^{|\mathcal{M}|}, \mathbf{v}_{\mathcal{M}} \in \mathbb{N}^{|\mathcal{M}|}} & \sum\limits_{m \in \mathcal{M}} \sum\limits_{a \in \mathcal{A}_m} p_a \; x_a \\ & \sum\limits_{a \in \mathcal{A}_m} x_a = v_m, & \forall \; m \in \mathcal{M} \\ & \sum\limits_{m \in \mathcal{M}_b} \sum\limits_{s_i \in \mathcal{S}_{fm}} w_{s_i m} \le c_b, & \forall \; f \in \mathcal{F}, \; \forall \; b \in B_f \\ & v_m = \sum\limits_{s_i \in \mathcal{S}_m} w_{s_i m}, & \forall \; m \in \mathcal{M} \\ & \sum\limits_{a \in \mathcal{A}_m} x_a + z_m = d_m, & \forall \; m \in \mathcal{M} \\ & \zeta_{\bar{a}_m} x_a - \zeta_a z_m \le 0, & \forall \; m \in \mathcal{M}, \forall a \in \mathcal{A}_m \\ & z_m \ge \overline{d}_m, & \forall \; m \in \mathcal{M} \\ & v_m \le d_m - \overline{d}_m, & \forall \; m \in \mathcal{M} \end{cases}$$

We can consider the following sets

- $\mathcal{V} = \left\{ \begin{array}{ll} (\mathbf{v}, \mathbf{w}) \in \mathbb{N}^{|\mathcal{M}|} \times \mathbb{N}^{\eta} \; : & v_m = \sum\limits_{s_i \in \mathcal{S}_m} w_{s_i m}, \quad \forall \; m \in \mathcal{M}, \\ & \sum\limits_{m \in \mathcal{M}_b} \sum\limits_{s_i \in \mathcal{S}_{fm}} w_{s_i m} \le c_b, \quad \forall \; f \in \mathcal{F} \; \forall \; b \in B_f, \\ & v_m \; \le d_m - \overline{d}_m, \quad \forall m \in \mathcal{M} \end{array} \right\}$
- $\mathcal{S}(\mathbf{w}) = \left\{ \; \mathbf{x} \in \mathbb{N}^{|\mathcal{A}|} \; : \; \sum_{a \in \mathcal{P}_{s_i}} x_a = w_{s_i m}, \quad \forall \; m \in \mathcal{M}, \; \forall \; s_i \in \mathcal{S}_m \; \right\}$
- $\mathcal{L}(\mathbf{v}) = \left\{ \mathbf{x} \in \mathbb{N}^{|\mathcal{A}|} \; : \; \sum\limits_{a \in \mathcal{A}_m} x_a = v_m, \quad \forall \; m \in \mathcal{M} \right\}$

We introduce the functions

$$f_{s_i m}(\mathbf{x}) = \sum_{a \in \mathcal{P}_{s_i}} p_a \; x_a, \qquad m \in \mathcal{M}, \;\; s_i \in \mathcal{S}_{\Updownarrow}.$$

The multi service-market problem can now be rewritten in compact form as

$$\begin{cases} \max\limits_{\mathbf{x},\mathbf{w},\mathbf{v}_{\mathcal{M}}} & \sum\limits_{m \in \mathcal{M}} \sum\limits_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x}) \\ & (\mathbf{v}_{\mathcal{M}}, \mathbf{w}) \in \mathcal{V} \\ & (\mathbf{v}_{\mathcal{M}}, \mathbf{z}) \in \mathcal{G} \\ & \mathbf{x} \in \mathcal{S}(\mathbf{w}) \end{cases}$$

where set $\mathcal{G}$ is defined in Section 4.1.
As in Section 4.1 we can eliminate variables $\mathbf{z}$ from the formulation defining

$$\Gamma(\mathbf{v}) = \mathcal{G} \cap \mathcal{L}(\mathbf{v}) = \left\{ \mathbf{x} \in \mathbb{N}^{|\mathcal{A}|} \; : \; x_a \leq \left\lfloor \frac{\zeta_a}{\zeta_{\bar{a}_m}} \; (d_m - v_m) \right\rfloor \quad \forall m \in \mathcal{M}, \; \forall a \in \mathcal{A}_m \right\}$$

Since we assume $(\mathbf{v}, \mathbf{w}) \in \mathcal{V}$, condition $v_m \leq \sum\limits_{a \in \mathcal{A}_m} d_a \;\; \forall m \in \mathcal{M}$ is always satisfied.
Hence, this constraint is redundant and we remove it.
The problem of allocating capacities to the service on the markets can be formulated as

$$\begin{aligned} \max_{\mathbf{x},\mathbf{v},\mathbf{w}} \quad & \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x}) \\ & (\mathbf{v}, \mathbf{w}) \in \mathcal{V} \\ & \mathbf{x} \in \Gamma(\mathbf{v}) \\ & \mathbf{x} \in \mathcal{S}(\mathbf{w}) \end{aligned} \tag{15}$$

We bring back the same reasoning made in the previous sections to get equivalence with the two-stage problem.

Proposition 4.6. *Any optimal solution of the two stage problem*

$$\max_{(\mathbf{v},\mathbf{w}) \in \mathcal{V}} \; \max_{\mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x}) \tag{16}$$

is optimal for (15).

The proof is in Appendix 10.
We can further decompose problem (16). Indeed for each $m \in \mathcal{M}$, and $s_i \in \mathcal{S}_m$ consider the sets

- $\mathcal{H}_m(v_m) = \left\{ \mathbf{x}_{\mathcal{A}_m} \in \mathbb{N}^{|\mathcal{A}_m|} \; : \; x_a \leq \left\lfloor \frac{\zeta_a}{\zeta_{\bar{a}_m}} \; (d_m - v_m) \right\rfloor \quad \forall a \in \mathcal{A}_m \right\}, \quad m \in \mathcal{M}$
- $\mathcal{S}_m(\mathbf{w}_{\mathcal{S}_m}) = \left\{ \mathbf{x}_{\mathcal{A}_m} \in \mathbb{N}^{|\mathcal{A}_m|} \; : \; \sum\limits_{a \in \mathcal{P}_{s_i}} x_a = w_{s_i m} \quad \forall s_i \in \mathcal{S}_m \right\}, \quad m \in \mathcal{M}$

- $\mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m}) = \mathcal{H}_m(v_m) \cap \mathcal{S}_m(\mathbf{w}_{\mathcal{S}_m}), \quad m \in \mathcal{M}$

with the following properties

$$\prod_{m\in\mathcal{M}} \mathcal{H}_m(v_m) = \Gamma(\mathbf{v}) \qquad \prod_{m\in\mathcal{M}} \mathcal{S}_m(\mathbf{w}_{\mathcal{S}_m}) = \mathcal{S}(\mathbf{w}) \qquad \prod_{m\in\mathcal{M}} \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m}) = \Gamma(\mathbf{v})\cap\mathcal{S}(\mathbf{w})$$

We have the following equivalence result.

Proposition 4.7. *Any optimal solution of the problem*

$$\max_{(\mathbf{v},\mathbf{w})\in\mathcal{V}} \sum_{m\in\mathcal{M}} \max_{\mathbf{x}_{\mathcal{A}_m}\in\mathcal{T}_m(v_m,\mathbf{w}_{\mathcal{S}_m})} \sum_{s_i\in\mathcal{S}_m} f_{s_i m}(\mathbf{x}) \tag{17}$$

is optimal for (16).

The proof is in Appendix 11.

We can write $\mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m}) = \prod_{s_i\in\mathcal{S}_m} \mathcal{T}_{s_i m}(v_m, w_{s_i m}), \forall\ m \in \mathcal{M}$, where

$$\mathcal{T}_{s_i m}(v_m, w_{s_i m}) = \left\{ \begin{array}{lll} \mathbf{x}_{\mathcal{P}_{s_i}} \in \mathbb{N}^{|\mathcal{P}_{s_i}|} : & \sum_{a\in\mathcal{P}_{s_i}} x_a = w_{s_i m}, & \\ & x_a \le \left\lfloor \frac{\zeta_a}{\zeta_{\bar{a}_m}}\ (d_m - v_m) \right\rfloor & \forall a \in \mathcal{P}_{s_i} \end{array} \right\} \quad m \in \mathcal{M},\ s_i \in \mathcal{S}_m$$

We then can further decompose problem (17) as follow

$$\max_{(\mathbf{v},\mathbf{w})\in\mathcal{V}} \sum_{m\in\mathcal{M}} \sum_{s_i\in\mathcal{S}_m} \max_{\mathbf{x}_{\mathcal{P}_{s_i}}\in\mathcal{T}_{s_i m}(v_m,w_{s_i m})} f_{s_i m}(\mathbf{x}) \tag{18}$$

and we finally get

Proposition 4.8. *Any optimal solution of the problem*

$$\max_{(\mathbf{v},\mathbf{w})\in\mathcal{V}} \sum_{m\in\mathcal{M}} \sum_{s_i\in\mathcal{S}_m} \max_{\mathbf{x}_{\mathcal{P}_{s_i}}\in\mathcal{T}_{s_i m}(v_m,w_{s_i m})} f_{s_i m}(\mathbf{x}) \tag{19}$$

is optimal for (17).

The proof is in Appendix 12.
We are now ready to define the **Market Revenue Function** which generalizes Definition 4.3 for a given service.

Definition 4.9 (Market Revenue Function)**.** *For a given $m \in \mathcal{M}$ and assigned values $w_{s_i} \in \mathbb{N}$ for each $s_i \in \mathcal{S}_m$ such that $v_m \le \sum_{a\in\mathcal{A}_m} d_a$, the **Market Revenue Function** is*

$$p^{\mathcal{S}}_m(w_{s_1 m}, w_{s_2 m}, \dots, w_{s_q m}, v_m) := \max_{\mathbf{x}_{\mathcal{A}_m}\in\ \mathcal{T}_m(\mathbf{w})} \sum_{s_i\in\mathcal{S}_m} f_{sm}(\mathbf{x})$$

We note that the problem used in Definition 4.9 can be split into q smaller problem with the same structure. Indeed let us introduce the following definition.

Definition 4.10 (Market-Service Revenue Function)**.** *For a given $m \in \mathcal{M}$ such that $v_m \leq \sum_{a \in \mathcal{A}_m} d_a$, $s_i \in \mathcal{S}_m$ and an assigned value $w_{s_i m} \in \mathbb{N}$, we define the **Market-Service Revenue Function** as follows*

$$p_m^{s_i}(w_{s_i m}, v_m) := \max_{\mathbf{x}_{\mathcal{P}_{s_i}} \in \mathcal{T}_{s_i m}(v_m, w_{s_i m})} f_{s_i m}(\mathbf{x})$$

where $\mathbf{x}_{\mathcal{P}_{s_i}} = \left(x_{a_1}, \ \cdots, \ x_{a_{|\mathcal{P}_{s_i}|}} \right)^\intercal$.

Then we can easily write

$$p_m^{\mathcal{S}}(w_{s_1 m}, w_{s_2 m}, \dots, w_{s_q m}, v_m) = \sum_{s_i \in \mathcal{S}_m} p_m^{s_i}(w_{s_i m}, v_m)$$

and we define the Market-Service Slave problems for each $m \in \mathcal{M}$, $s_i \in \mathcal{S}_m$, and a feasible pair $v_m, w_{s_i m}$ as

$$\max_{\mathbf{x}_{\mathcal{P}_{s_i}} \in \mathcal{T}_{s_i m}(v_m, w_{s_i m})} f_{s_i m}(\mathbf{x})$$

Then the Master problem takes the form

$$\max_{(\mathbf{v}, \mathbf{w}) \in \mathcal{V}} \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} p_m^{s_i}(w_{s_i m}, v_m) \tag{20}$$

The size of the Master problem is large but not exponential. In fact, for each value of v_m the number of possibilities is less than quadratic for each market-service pair.
After performing experiments on different data sets, we observe that for each market $m \in \mathcal{M}$ and each service $s_i \in \mathcal{S}_m$ the structure of the Market Revenue Function $p_m^{s_i}(w_{s_i m}, v_m)$ is basically concave with some local oscillation. Hence we can construct a lower and an upper concave approximation, respectively $Lp_m^{s_i}(w_{s_i m}, v_m)$ and $Up_m^{s_i}(w_{s_i}, v_m)$, such that

$$Lp_m^{s_i}(w_{s_i m}, v_m) \leq p_m^{s_i}(w_{s_i m}, v_m) \leq Up_m^{s_i}(w_{s_i m}, v_m).$$

We report details of the construction and in the error made by using lower and upper concave approximation in Appendix **??**. Below we just report a simple example, using a one-market-one-service case, to show the behaviour of the function on a small real problem.

Example 4.11. *Consider a one-market-one-service example. In the table below, we report the set of alternatives with the corresponding attractiveness and fare. The total demand $d_m = 131.6$.*

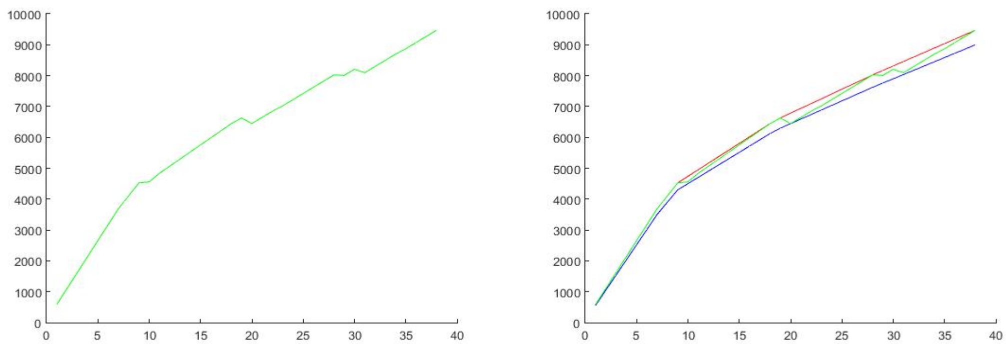


Figure 1: Left: Market Revenue Function. Right: Market Revenue Function, upper and lower approximation.

Alternative	*Attractiveness*	*Price*
a_1	*0*	*985*
a_2	*0.004*	*777*
a_3	*0.808*	*586*
a_4	*4.454*	*517*
a_5	*0.026*	*464*
a_6	*1.758*	*418*
a_7	*0.402*	*339*
a_8	*2.399*	*276*
a_9	*7.457*	*230*
a_{10}	*23.426*	*201*
a_{11}	*0.487*	*171*
a_0	*90.399*	

In figure 1 we plot the Market Revenue Function, lower and upper concave approximations.

Once we construct a concave upper approximation, we can substitute in problem (20) the function $p_m^{s_i}(w_{s_im}, v_m)$ with its piecewise linear concave approximation $Up_m^{s_i}(w_{s_im}, v_m)$. As it's well known, the sum of piecewise linear concave functions is still piecewise linear and concave. Hence we obtain a relaxed problem which has the same feasible region as in (20) and below objective function.

$$\sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} Up_m^{s_i}(w_{s_im}, v_m) = \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} \min_{(h,k) \in I \times J} \left\{ \gamma_{hk}^{s_im} + \alpha_{hk}^{s_im} v_m + \beta_{hk}^{s_im} w_{s_im} \right\}$$

where $I = \{l_m, \ldots, u_m\}$ and $J = \{l_{s_i}, \ldots, u_{s_i}\}$. It is well known that such an approximated problem can be reformulated as a linear programming problem by introducing auxiliary variables. By doing so we get the following Relaxed Master Problem.

$$\max_{(\mathbf{v},\mathbf{w},\mathbf{r})\in\mathcal{F}\times\mathbb{R}^\eta} \quad \sum_{m\in\mathcal{M}}\sum_{s_i\in\mathcal{S}_m} r_{s_im}$$

$$r_{s_im} \leq \gamma_{ij}^{s_im} + \alpha_{ij}^{s_im} v_m + \beta_{ij}^{s_im} w_{s_im}, \quad \forall m\in\mathcal{M}, \forall s_i\in\mathcal{S}_m, \forall h\in I, \forall k\in J \tag{21}$$

where $\mathbf{r}$ is the vector with components r_{s_im}, $m\in\mathcal{M}$ and $s_i\in\mathcal{S}_m$.

We compare the two problems: SBIP formulation (5) and Concave Approximation (21). In the standard SBIP formulation the number of variables is given by the total number of market-service alternatives plus the total number of markets. In the Concave Approximation of SBIP the a number of variables is given by the number of markets plus twice the number of market-services. Note that the size of the Concave Approximation problem does not depend on the number of alternatives.

Furthermore, we remove difficult total demand and relative demand constraints, and we introduce simplified linearized constraints that depend on only three variables (z_m, v_m and w_{s_im}).

The Concave Approximation always provides a measure of the solution quality. In fact, a given solution $(\hat{\mathbf{v}}, \hat{\mathbf{w}}, \hat{\mathbf{r}})$ of (21) is also a feasible point for (20) and the objective function is higher than the one computed in the master. So we have a messure between the concave approximated function and the true one as

$$\delta = \sum_{m\in\mathcal{M}}\sum_{s_i\in\mathcal{S}_m} \hat{r}_{s_im} - \sum_{m\in\mathcal{M}}\sum_{s_i\in\mathcal{S}_m} p_m^{s_i}(\hat{w}_{s_im}, \hat{v}_m)$$

In the next section we will show how to improve the value of δ.

Example 4.12. *As an example, we consider a dataset from a real network airline. In the table below we report the number of market-services (MS) and the corresponding number of alternatives (A per MS)*

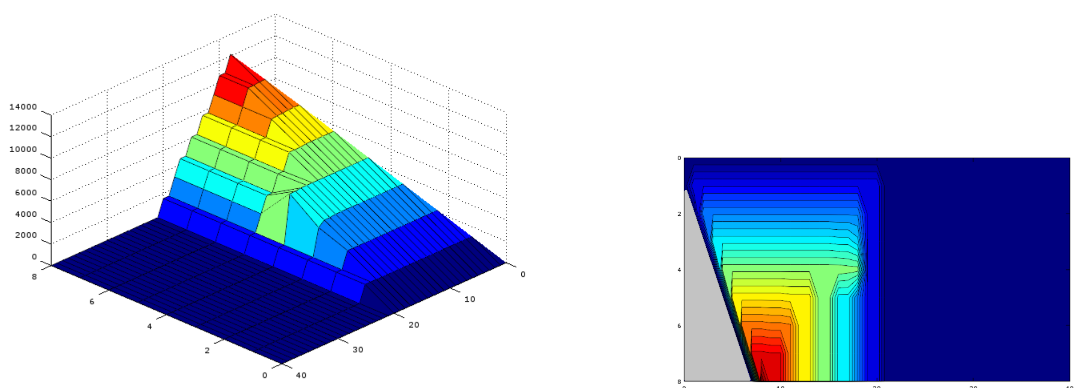
	MS	*A per MS*
	2 060	*1*
	99	*2*
	228	*3*
	3 293	*5*
	13 904	*11*
MS total	*19 584*	
Average A per MS		*8.8*

The standard SBIP for this problem has potentially 184 701 variables (172 351 discrete, 12 350 continuous), the decomposed one has only 51 518 variables (31 934 discrete, 19 584 continuous). We say "potentially" because in real applications it is common to have fixed variables.

4.3. Flat Valley disjunctive cuts

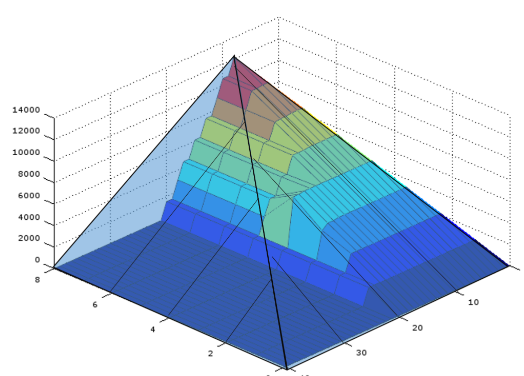
During the analysis of Concave Approximation (21) solutions, we observe a large flat valley in the shape of the Market-Service Revenue Function (MSRF). We called this particular phenomenon as *Flat Valley*.
Indeed, when MSRF has a high peak with multiple services for a given market, the upper concave relaxation Up_m^s of the MSRF often contains a point such that $p_m^s(v_m, w_{sm}) = 0$. This fact generates a flat area in the level curves of the MSRF. If this area is considerably far from the highest peak, then the approximation gives a significant value of δ. As an example, see Figure 2 where such a case is plotted and the corresponding Concave piecewise linear approximation is reported in Figure 3 where the large values of δ are shown.

Figure 2: MRF: level curves and plot



In order to handle these cases we adopt an effective idea. When the current solution stays in the Flat Valley then its value is considered zero, otherwise we build the concave approximation by using only the points outside of the Flat Valley. It's important to underline that the presence of a Flat Valley appears when the value of v_m exceeds

Figure 3: Concave piecewise linear approximation of the MRF



a certain level. Indeed due to the constraints in the Slave problem, the value of v_m limits the quantity of possible resources allocatable. If v_m is too high then there are no resources to allocate and the objective function is zero. Let $m \in \mathcal{M}$ and $s \in \mathcal{S}_m$ be a Market-Service which MSRF has a significant δ due to the Flat Valley and let g_{sm} be the value of v_m beyond which the Flat Valley appears. We add constraints to problem (21) modelling the following conditions:

$$\begin{array}{lll} \text{If } v_m \geq g_{sm} & \Rightarrow & r_{sm} = 0 \\ \text{If } v_m \leq g_{sm} - 1 & \Rightarrow & r_{sm} \leq \gamma_{ij}^{sm} + \alpha_{ij}^{sm} v_m + \beta_{ij}^{sm} w_{sm} \qquad i \in \hat{I}, j \in \hat{J} \end{array}$$

where $\hat{I} = \{l_m, \ldots, g_{sm}\}$ and $\hat{J} = \{l_s, \ldots, min\{g_{sm}, u_s\}\}$. These disjunctive constraints can be easily imposed by using standard Big-M constraints as follows

$$\begin{array}{ll} r_{sm} \leq \gamma_{ij}^{sm} + \alpha_{ij}^{sm} v_m + \beta_{ij}^{sm} w_{sm} + M(1 - y_m) & s \in \mathcal{S}_m, i \in I(g_{sm}), j \in J(g_{sm}) \\ r_{sm} \leq M y_m & \\ v_m + \tilde{M} y_m \geq g_{sm} & \\ v_m \leq g_{sm} - 1 + \tilde{M}(1 - y_m) & \end{array}$$

where $y_m \in \{0, 1\}$ is a new boolean variable. Upper bound values for M and $\tilde{M}$ can be found analytically. We used $M = \max_{v_m, w_{sm}} \{p_m^s(v_m, w_{sm})\}$ and $\tilde{M} = u_m$.

The overall algorithm computes a solution to the Concave Approximation, then suitable Big-M constraints are added to the Master problem if the approximation gap is too large due to Flat Valley effect. The scheme is detailed below.

Approximation and Flat Valley Algorithm

Preprocessing.
For each $m \in \mathcal{M}$, $s \in \mathcal{S}_m$ and feasible pair (w, v) in the integer lattice, compute $p_m^s(w, v)$.
Concave Approximation
Solve the Relaxed Master Problem $\mathcal{RMP}$ as formulated in (21).
Evaluation of δ
Given a solution to $\mathcal{RMP}$, let $\mathcal{H} \subseteq \mathcal{S}$ be the set of Market-Services with $\delta > 0$.
Flat Valley Management
For each Market-Service in $\mathcal{H}$, add and remove constraints as reported in section (4.3), then solve the Master Problem with Flat Valley constraints.

5. Computational Experience

Computational analysis considered 293 real-world data instances, each one representing one-day bookings of a medium airline. We used IBM ILOG CPLEX 12.6.0.1 as solver for mixed integer programs. The code has been written in Java and experiments have

been carried out on a machine with 4 (single-core) CPUs Intel Xeon E5-2670 at 2.60 GHz and 62 GB of RAM hosting Red Hat Enterprise Linux Server 7.2.
Initially we want to compare the SBIP formulation and the Concave Approximation. Then we will make some analysis with respect to the Flat Valley framework.
Each dataset is characterized by a variety of parameters, each with its own influence on the performance of the formulation. We summarize the information in Table 1.

Table 1:

	Markets	Legs	Services	MarketAlternatives
MAX	15,756.00	291.00	4,029.00	225,661.00
AVERAGE	12,160.43	279.17	3,378.78	168,208.17
MIN	8,829.00	258.00	2,726.00	121,216.00
STD DEV.	1,440.65	6.06	255.33	21,875.17

We consider δ as the gap between the approximating function (Concave Approximation or Flat Valley) and the real objective function. We will speak about MIP GAP as the gap between the value of the estimated optimum and the best integer solution found by the solver. MIP GAP is used in CPLEX as a stopping rule, once a solution with this MIP GAP or lower is reached then the algorithm stops. We measure the performance using both the computational time and deterministic time measured in ticks. results are in Table 2.
The experiments were performed as follows: first we run the Concave Approximation and we find the δ of the Concave Approximation; then we run the Standard SBIP until the MIP GAP of the Standard SBIP is equal to δ.

Table 2: Computational effort of Concave Approximation algorithm versus SBIP

	Time Conc. Approx.	Equiv. Time SBIP	Ticks Conc. Approx.	Equiv. Ticks SBIP
MAX	82.12	90.01	59,437.87	46093.67
AVERAGE	14.51	31.02	8,239.96	16903.63
MIN	2.91	6.11	639.71	3607.12
STD DEV.	10.00	17.89	7,229.62	9377.09

On average our approach performs 50% better in terms of time and ticks. In addition, the standard deviation is remarkably smaller. Figure 4 compares computational times required by SBIP and Concave Approximation to achieve equivalent solutions for each instance.
The most relevant aspect of our approach is the significant reduction in discrete variables. In particular, our method leads to a reduction in the number of discrete variables by 60% on average, with significantly small standard deviation, at the cost of an increase in the number of continuous variables. An important benefit of proposed decomposition is that the number of sub-problems in Branch-and-Bound enumeration scheme turns out to be considerably reduced, in favor of the acceptable increased size of each sub-problem. Table 3 reports the number of discrete variables, as well as the

Figure 4: Comparing computational performance of SBIP and Concave Approximation

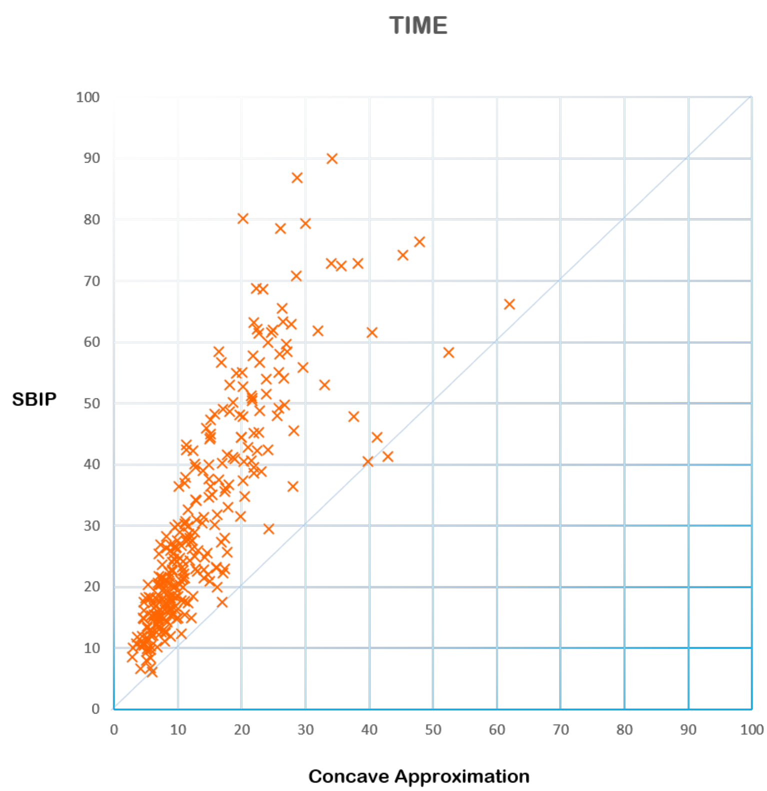
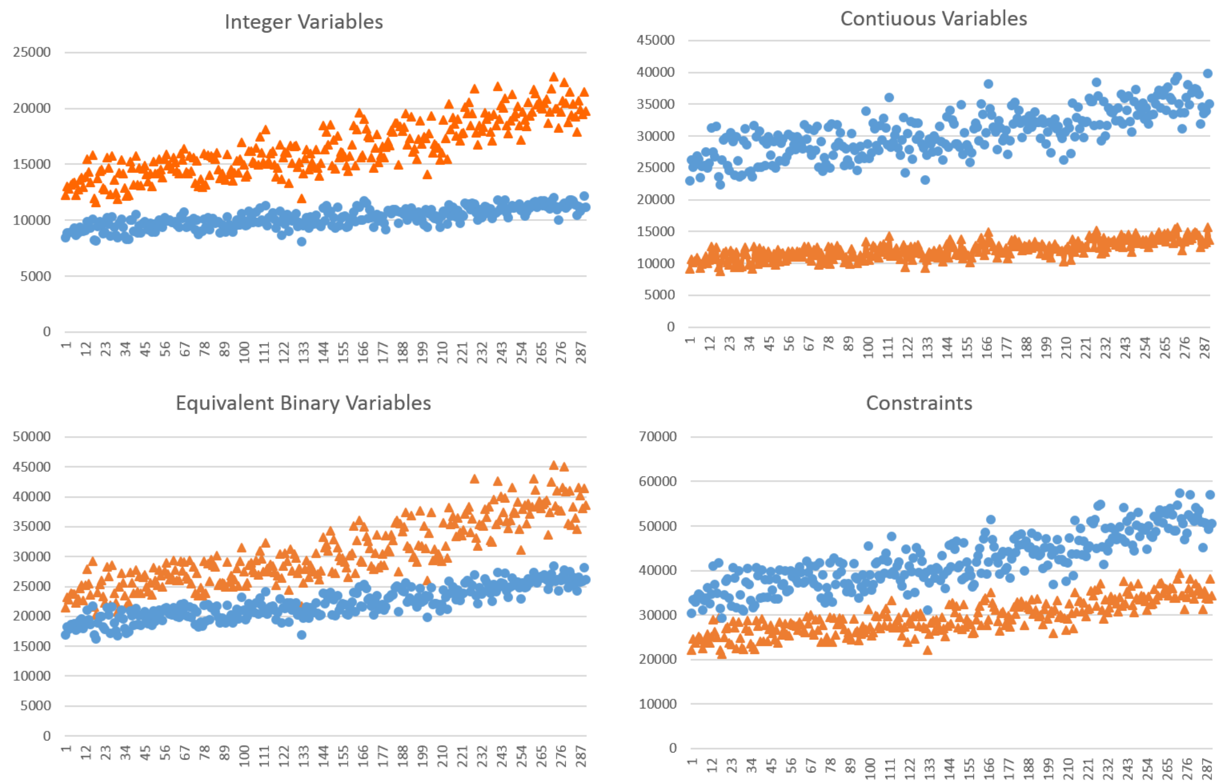


Figure 5: Variables and constraints ordered by time. Standard SBIP (orange triangles) and Concave Approximation (blue circles).



corresponding number of equivalent binary variables in relation to their bounds. Table 4 reports the number of continuous variables and the total number of variables in the addressed programs. As Figure 5 shows, the trend is an increment of these increases, in fact despite continuous variables seem to have the same trend, discrete and equivalent binaries grow more in the SBIP than in the Concave Approximation.

Table 3: Discrete Variables (DV) and Equivalent Binary Variables (EBV) for Concave Approximation and SBIP.

	SBIP DV	Conc. Approx. DV	SBIP EBV	Conc. Approx. EBV
MAX	22,850.00	12,146.00	45,246.00	28,477.00
AVERAGE	16,445.00	10,184.64	30,620.11	22,128.28
MIN	11,599.00	8,111.00	20,247.00	16,156.00
STD.DEV.	2,464.68	856.63	5,435.03	2,647.40

The linearization of the objective function implies that several new constraints have to be added. Although the number of such constraints is relatively big, they have a simple structure; in fact, they are composed by only three variables, two continuous and one discrete (the Market allocation variable can be considered as a continuous

Table 4: Continuous Variables (CV) and Total Variables (Tot) for Concave Approximation and SBIP.

	SBIP CV	Conc. Approx. CV	Sbip Tot	Conc. Approx. Tot
MAX	15,655.00	39,807.00	38,505.00	51,953.00
AVERAGE	12,078.06	30,506.91	28,523.06	40,691.56
MIN	8,778.00	22,297.00	20,377.00	30,453.00
STD.DEV.	1,423.07	3,638.40	3,810.90	4,463.07

variable, since its integrality is maintained by the Market-Service allocation variables, see problem (21)). On the other hand, only a few constraints are active, so only a relatively small portion of them are needed at a time. Please refer to Table 5 to see total number of constraints in both programs.

Table 5: Constraints (Cons) for Concave Approximation and SBIP.

	SBIP Cons	Conc. Approx. Cons
MAX	39,464.00	57,344.00
AVERAGE	29,442.95	42,209.10
MIN	21,226.00	29,303.00
STD.DEV.	3,837.24	6,098.33

Finally, Table 6 confirms that the number of nodes is reduced on average in Concave Approximation. There are some outliers making the two results comparable in the worst case.

We observe that Flat Valley approach is less performing than Concave Approximation in terms of computational cost, but it still outperforms the standard SBIP in most of the instances. We also observe that Flat Valley approach is more effective than the Concave Approximation for some specific instances. Main reason is that Flat Valley method solves a weaker convex relaxation of the problem, thus the resulting problem turns out to be more structured and then its resolution may take advantage of better computational performance. Please refer to Table 7 for more detailed computational results.

We may conclude that choosing between Concave Approximation and Flat Valley approaches strongly depends on the features of airline network. Flat Valley becomes more effective as certain Market-Service has high prices. As the MIP GAP value provides an objective stopping criterion to obtain equivalent solutions from discussed methods, we observe that the δ obtained with the Flat Valley is usually more accurate than the one

Table 6: Branch-and-Cut nodes to solve Concave Approximation and SBIP.

	Concave Approximation Nodes	SBIP Nodes
MAX	17320.00	11023.00
AVERAGE	2592.81	4609.72
MIN	1.00	1158.00
STD.DEV.	2411.30	2485.82

Table 7: Computational performance of Flat Valley

	Flat Valley Ticks	Flat Valley Time
MAX	59739.72	89.85
AVERAGE	13160.56	19.87
MIN	1941.90	3.54
STD.DEV.	10020.58	14.77

Table 8: Comparison between Concave Approximation and Flat Valley GAPs, taking into account their absolute value (Abs), the absolute difference (AbsD) and the percentage difference (PercD).

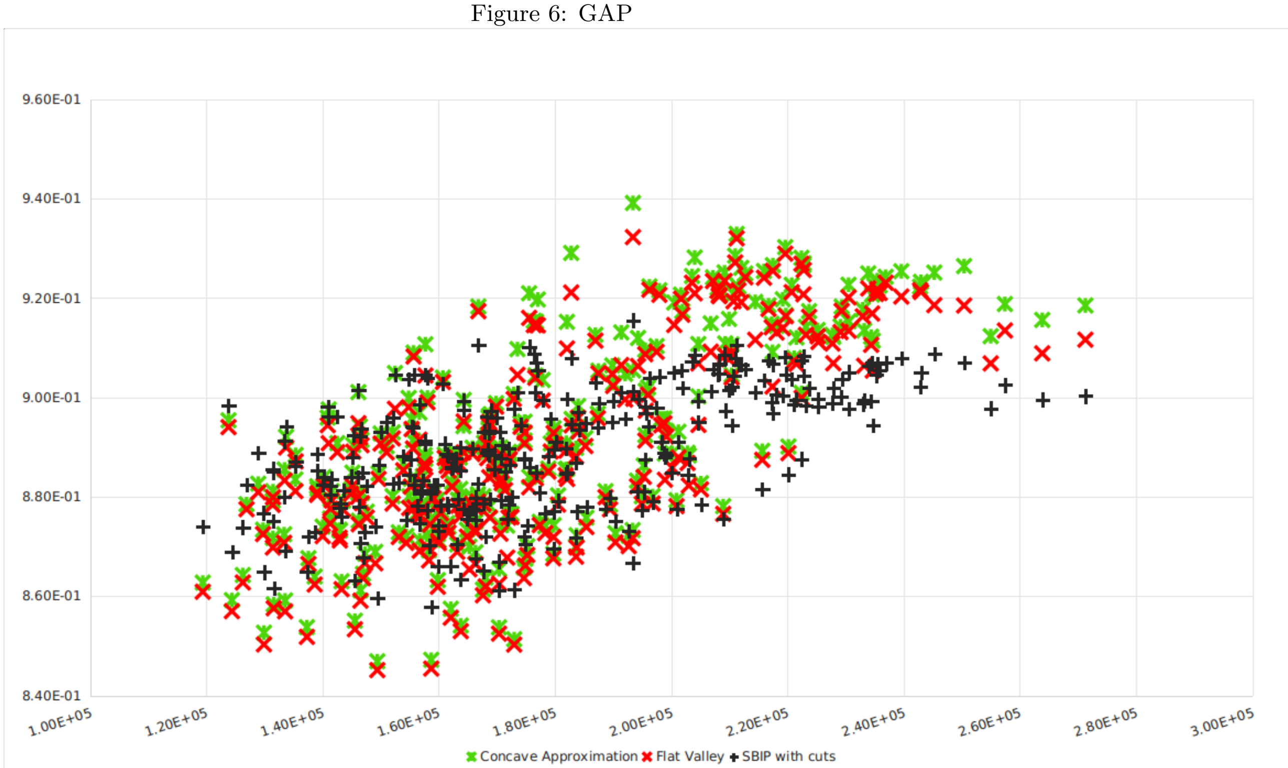
	Conc. Approx. Abs	Flat Valley Abs	Gap AbsD	Gap PercD
MAX	0.17	0.12	0.05	34.84
AVERAGE	0.10	0.09	0.01	11.14
MIN	0.06	0.05	0.001	1.76
STD.DEV.	0.024	0.017	0.012	7.76

obtained using a Concave approximation. On average the gain in percentage is about 10%, as shown in Table 8.
From another point of view, the relaxation of each formulation we propose (Concave Approximation, Flat Valley, Improved SBIP) provides better bounds than the linear relaxation of standard SBIP. They are always tighter than standard SBLP. This fact is apparently counter-intuitive, since optimal solution to each introduced approximation is an upper bound to the optimal solution of standard SBIP.
In particular, as we can see in Figure 6, Improved SBIP provides tighter bounds as the complexity of the problem grows. In the same picture, we plot the ratio between the bound provided by each introduced approximation and the bound of Standard SBIP.

6. Conclusions and future research

SBIP is one of the most effective revenue management models. In real applications, SBIP is a large-scale integer programming problem, that cannot be solved in practice. Even its linear relaxation, the SBLP, requires an excessive computational effort by state-of-the-art commercial solvers. The results in this paper investigate important properties of SBIP which lead to an improved formulation. In particular, we introduced simple cuts that significantly reduce the bound provided by SBLP. Moreover, we investigate a decomposition approach based on the market-service concept.
Flight-based or O&D decomposition approaches in the literature suffer from breaking the important relations among markets and flights availability. Our decomposition overcomes this drawback and reflects the real structure of the airline network.
This decomposition leads to a non-linear non-convex problem with discrete variables. In order to produce a solution in a reasonable time we derive a heuristic procedure, called Concave Approximation, using the concave hull of the objective function.
Furthermore, to improve the solution quality we develop a formulation, called Flat Valley, by incorporating disjunctive cuts.

Figure 6: GAP



9.60E-01
9.40E-01
9.20E-01
9.00E-01
8.80E-01
8.60E-01
8.40E-01
1.00E+05
1.20E+05
1.40E+05
1.60E+05
1.80E+05
2.00E+05
2.20E+05
2.40E+05
2.60E+05
2.80E+05
3.00E+05
Concave Approximation
Flat Valley
SBIP with cuts

Computational experience shows a globally good behaviour, not only for the integer model but also for its relaxation.
Finally, we present an extension of our problem with respect to the adoption of other customer choice models. Briefly, our decomposition covers all customer choice models, like BAM and GAM, and it is open to a variety of new possibilities (e.g. machine learning models).

APPENDICES

7. Proof of proposition (4.2)

Proof. Now let be $(\hat{\mathbf{x}}, \hat{\mathbf{z}}, \hat{\mathbf{v}}_{\mathcal{M}})$ an optimal solution for (7). Since it is optimal then it is also feasible and $(\hat{\mathbf{x}}, \hat{\mathbf{z}}) \in \mathcal{G}$, $\hat{\mathbf{x}} \in S(\hat{\mathbf{v}}_{\mathcal{M}})$ and $\hat{\mathbf{v}}_{\mathcal{M}} \in \mathcal{V}$. Thus it is feasible for (6). We call $h(\mathbf{v}) = \max_{(\mathbf{x},\mathbf{z}) \in \mathcal{G} \cap S(\mathbf{v})} \sum_{m \in \mathcal{M}} f_m(\mathbf{x})$.

$$h(\hat{\mathbf{v}}) \geq h(\mathbf{v}), \quad \forall\ \mathbf{v} \in \mathcal{V}$$

$$h(\mathbf{v}) \geq \sum_{m \in \mathcal{M}} f_m(\mathbf{x}), \quad \forall (\mathbf{x}, \mathbf{z}) \in \mathcal{G} \cap S(\mathbf{v})$$

then

$$\sum_{m \in \mathcal{M}} f_m(\hat{\mathbf{x}}) = h(\hat{\mathbf{v}}) \geq \sum_{m \in \mathcal{M}} f_m(\mathbf{x}), \quad \forall\ \mathbf{v} \in \mathcal{V},\ \forall\ (\mathbf{x}, \mathbf{z}) \in \mathcal{G} \cap S(\mathbf{v})$$

□

8. Proof of proposition (4.4)

Proof. By contradiction assume that there exists $m \in \mathcal{M}$ and allocations $v, v+1 \in \mathcal{V}$ such that

$$\frac{p_m(v+1)}{v+1} > \frac{p_m(v)}{v} \tag{22}$$

It is possible to write

$$p_m(v+1) = p_m(v) + \Delta p_m(v+1)$$

where $\Delta p_m(v+1)$ represents the variation. Replacing it in (22) results in

$$\frac{p_m(v)}{v+1} + \frac{\Delta p_m(v+1)}{v+1} > \frac{p_m(v)}{v}$$

$$\Downarrow$$

$$\Delta p_m(v+1)v > p_m(v) \tag{23}$$

Without loss of generality, it is possible to assume that the fares are ordered decreasingly

$$p_{a_1} \geq p_{a_2} \geq \ldots \geq p_{a_r} \geq \ldots \geq p_{a_{|\mathcal{A}_m|}},$$

so that by definition

$$p_m(v_m) = p_{a_1} x_{a_1} + p_{a_2} x_{a_2} + \ldots + p_{a_r} x_{a_r}$$

where $r \leq |\mathcal{A}_m|$ and $x_{a_i} = 0$ for $i > r$. Therefore the following condition holds

$$p_m(v_m) \geq p_{a_r} v_m$$

On the other hand we are increasing the allocation v of one unit in the Slave problem under (4.3) so that

$$\sum_{i=1}^{|\mathcal{A}_m|} (x_{a_i}^{v+1} - x_{a_i}^{v}) = \sum_{i=1,\ldots,r} (x_{a_i}^{v+1} - x_{a_i}^{v}) + \sum_{i>r} x_{a_i}^{v+1} = 1$$

where $x_{a_i}^{v+1} - x_{a_i}^{v} \leq 0$ (since $x_a^v \leq \frac{\zeta_a}{\zeta_{\bar{a}_m}} (d_m - v)$, $\forall\, a \in \mathcal{A}_m$ and $\forall\, v$ feasible for (10)). Hence, the following inequality holds

$$\begin{aligned}
\Delta p_m(v+1) &= p_m(v+1) - p_m(v) \\
&= \sum_{i=1,\ldots,r} p_{a_i}(x_{a_i}^{v+1} - x_{a_i}^{v}) + \sum_{i>r} p_{a_i} x_{a_i}^{v+1} \\
&\leq p_{a_r} \sum_{i=1,\ldots,r} (x_{a_i}^{v+1} - x_{a_i}^{v}) + p_{a_r} \sum_{i>r} x_{a_i}^{v+1} \\
&\leq p_{a_r}
\end{aligned}$$

Finally, this means that

$$p_m(v) \geq p_{a_r} v \geq \Delta p_m(v+1) v$$

which is in contradiction with (23). Hence the thesis follows. □

9. Proof of proposition (4.5)

Proof. Suppose to take two feasible allocation $\hat{v}$ and $\tilde{v}$ feasible for (10) such that $\hat{v} \geq \tilde{v}$. As result of property (4.4), applied recursively, it follows that

$$\frac{p_m(\hat{v})}{\hat{v}} \leq \frac{p_m(\tilde{v})}{\tilde{v}}$$

With simple passages, the following inequalities hold

$$p_m(\hat{v}) \leq \frac{\hat{v}}{\tilde{v}} p_m(\tilde{v})$$

$$p_m(\hat{v}) - p_m(\tilde{v}) \leq \left(\frac{\hat{v}}{\tilde{v}} - 1\right) p_m(\tilde{v})$$

$$p_m(\hat{v}) - p_m(\tilde{v}) \leq (\hat{v} - \tilde{v}) \frac{p_m(\tilde{v})}{\tilde{v}}$$

$$|p_m(\hat{v}) - p_m(\tilde{v})| \leq |\hat{v} - \tilde{v}| \frac{p_m(\tilde{v})}{\tilde{v}}$$

Therefore, it follows that

$$|p_m(\hat{v}) - p_m(\tilde{v})| \leq |\hat{v} - \tilde{v}| \ \max \left\{ \frac{p_m(\tilde{v})}{\tilde{v}} : \tilde{v} > 0 \text{ and Master-feasible} \right\}$$

Because of the maximum within $\frac{p_m(\tilde{v})}{\tilde{v}}$ and the absolute value function, the inequality is still valid even if $\hat{v} < \tilde{v}$. From the proof of proposition (4.4), it follows that $\frac{p_m(\tilde{v})}{\tilde{v}} \leq \max_{a \in \mathcal{A}_m} \{p_a\} = p^*$, so that

$$|p_m(\hat{v}) - p_m(\tilde{v})| \leq |\hat{v} - \tilde{v}| \ L$$

where $L = p^* + \epsilon$, with $\epsilon > 0$ to prevent hill formulated slaves. Note that the following expression is always valid for Master-Feasible values and even 0 is acceptable. □

10. Proof of proposition (4.6)

Proof. Now let be $(\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{w}})$ an optimal solution for (16). Since it is optimal then it is also feasible and $\hat{\mathbf{x}} \in \Gamma(\hat{\mathbf{v}})$, $\hat{\mathbf{x}} \in S(\hat{\mathbf{w}})$ and $(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \in \mathcal{V}$. Thus it is feasible for (15).
We call $h(\mathbf{v}, \mathbf{w}) = \max_{\mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \sum_{m \in \mathcal{M}} \sum_{s \in \mathcal{S}_m} f_{sm}(\mathbf{x})$

$$h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \geq h(\mathbf{v}, \mathbf{w}), \quad \forall (\mathbf{v}, \mathbf{w}) \in \mathcal{V}$$

$$h(\mathbf{v}, \mathbf{w}) \geq \sum_{m \in \mathcal{M}} \sum_{s \in \mathcal{S}_m} f_{sm}(\mathbf{x}), \quad \forall \mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})$$

then

$$\sum_{m \in \mathcal{M}} \sum_{s \in \mathcal{S}_m} f_{sm}(\hat{\mathbf{x}}) = h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \geq \sum_{m \in \mathcal{M}} \sum_{s \in \mathcal{S}_m} f_{sm}(\mathbf{x}), \quad \forall (\mathbf{v}, \mathbf{w}) \in \mathcal{V}, \ \forall \mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})$$

□

11. Proof of proposition (4.7)

Proof. Let be $(\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{w}})$ a solution for (17). We have

$$\hat{\mathbf{x}}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m}) \quad \forall m \in \mathcal{M} \quad \Rightarrow \quad \hat{\mathbf{x}} \in \prod_{\Updownarrow \in \mathcal{M}} \mathcal{T}_{\Updownarrow}(\sqsubseteq_{\Updownarrow}, \mathbf{w}_{\mathcal{S}_{\Updownarrow}}) = -(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})$$

and $(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \in \mathcal{V}$.
Let consider the functions

$$h(\mathbf{v}, \mathbf{w}) = \sum_{m \in \mathcal{M}} \max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x})$$

We have that

$$h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \geq h(\mathbf{v}, \mathbf{w}) \quad \forall (\mathbf{v}, \mathbf{w}) \in \mathcal{V}$$

$$\max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x}) \geq \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x}), \quad \forall\ \mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m}),\ \forall\ m \in \mathcal{M}$$

Summing by $m \in \mathcal{M}$

$$h(\mathbf{v}, \mathbf{w}) \geq \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x}), \quad \forall\ \mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})$$

$$\Rightarrow h(\mathbf{v}, \mathbf{w}) \geq \max_{\mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x})$$

We can now observe that

$$h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) = \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\hat{\mathbf{x}}) = \max_{\mathbf{x} \in \Gamma(\hat{\mathbf{v}}) \cap \mathcal{S}(\hat{\mathbf{v}}) \cap \mathcal{S}(\hat{\mathbf{w}})} \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x}))$$

Since $(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \in \mathcal{V}$, $\hat{\mathbf{x}} \in \Gamma(\hat{\mathbf{v}}) \cap \mathcal{S}(\hat{\mathbf{v}}) \cap \mathcal{S}(\hat{\mathbf{w}})$ and $h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \geq \max_{\mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x})$.

Finally

$$h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) = \max_{\mathbf{x} \in \Gamma(\hat{\mathbf{v}}) \cap \mathcal{S}(\hat{\mathbf{v}}) \cap \mathcal{S}(\hat{\mathbf{w}})} \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x})) \geq \max_{\mathbf{x} \in \Gamma(\mathbf{v}) \cap \mathcal{S}(\mathbf{v}) \cap \mathcal{S}(\mathbf{w})} \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_{\Updownarrow}} f_{s_i m}(\mathbf{x}), \quad \forall (\mathbf{v}, \mathbf{w}) \in \mathcal{V}$$

□

12. Proof of proposition (4.8)

Proof. Let be $(\hat{\mathbf{x}}, \hat{\mathbf{v}}, \hat{\mathbf{w}})$ a solution for (19). We have

$$\hat{\mathbf{x}}_{\mathcal{P}_{s_i}} \in \mathcal{T}_{s_i m}(v_m, w_{s_i m}) \quad \forall m \in \mathcal{M},\ \forall \int \in \mathcal{S}_{\Updownarrow} \quad \Rightarrow \quad \hat{\mathbf{x}}_{\mathcal{A}_{\Updownarrow}} \in \prod_{\int \in \mathcal{S}_{\Updownarrow}} \mathcal{T}_{\int\rangle\Updownarrow}(\sqsubseteq_{\Updownarrow}, \sqsupseteq_{\int\rangle\Updownarrow}) = \mathcal{T}_{\Updownarrow}(\sqsubseteq_{\Updownarrow}, \mathbf{w}_{\mathcal{S}_{\Updownarrow}}) \quad \forall \Updownarrow \in \mathcal{M}$$

and $(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \in \mathcal{V}$.
Let consider the functions

$$g_m(\mathbf{v}, \mathbf{w}) = \sum_{s_i \in \mathcal{S}_m} \max_{\mathbf{x}_{\mathcal{P}_{s_i}} \in \mathcal{T}_{s_i m}(v_m, w_{s_i m})} f_{s_i m}(\mathbf{x}), \quad m \in \mathcal{M}$$

and

$$h(\mathbf{v}, \mathbf{w}) = \sum_{m \in \mathcal{M}} g_m(\mathbf{v}, \mathbf{w}) = \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} \max_{\mathbf{x}_{\mathcal{P}_{s_i}} \in \mathcal{T}_{s_i m}(v_m, w_{s_i m})} f_{s_i m}(\mathbf{x})$$

We have that

$$h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \geq h(\mathbf{v}, \mathbf{w}) \quad \forall (\mathbf{v}, \mathbf{w}) \in \mathcal{F}$$

$$\max_{\mathbf{x}_{\mathcal{P}_{s_i}} \in \mathcal{T}_{s_i m}(v_m, w_{s_i m})} f_{s_i m}(\mathbf{x}) \geq f_{s_i m}(\mathbf{x}), \quad \forall\ \mathbf{x}_{\mathcal{P}_{s_i}} \in \mathcal{T}_{s_i m}(v_m, w_{s_i m}),\ \forall\ m \in \mathcal{M},\ \forall\ \int_\rangle \in \mathcal{S}_\Updownarrow$$

Summing by $s_i \in \mathcal{S}_\Updownarrow$ we have

$$g_m(\mathbf{v}, \mathbf{w}) \geq \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x}), \quad ,\ \forall\ \hat{\mathbf{x}}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m}),\ \forall\ m \in \mathcal{M}$$

$$\Rightarrow g_m(\mathbf{v}, \mathbf{w}) \geq \max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x}), \quad \forall\ m \in \mathcal{M}$$

Summing by $m \in \mathcal{M}$

$$h(\mathbf{v}, \mathbf{w}) \geq \sum_{m \in \mathcal{M}} \max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x})$$

We can now observe that

$$h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) = \sum_{m \in \mathcal{M}} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\hat{\mathbf{x}}) = \sum_{m \in \mathcal{M}} \max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(\hat{v}_m, \hat{\mathbf{w}}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x})$$

Since $(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \in \mathcal{V}$, $\hat{\mathbf{x}}_{\mathcal{A}_m} \in \mathcal{T}_m(\hat{v}_m, \hat{\mathbf{w}}_{\mathcal{S}_m})$ $m \in \mathcal{M}$ and $h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) \geq \sum_{m \in \mathcal{M}} \max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(\hat{v}_m, \hat{\mathbf{w}}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x})$.

Finally

$$\begin{aligned} h(\hat{\mathbf{v}}, \hat{\mathbf{w}}) = \sum_{m \in \mathcal{M}} \max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(\hat{v}_m, \hat{\mathbf{w}}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_m} \quad & f_{s_i m}(\mathbf{x}) \geq \\ & \geq h(\mathbf{v}, \mathbf{w}) \geq \\ & \geq \sum_{m \in \mathcal{M}} \max_{\mathbf{x}_{\mathcal{A}_m} \in \mathcal{T}_m(v_m, \mathbf{w}_{\mathcal{S}_m})} \sum_{s_i \in \mathcal{S}_m} f_{s_i m}(\mathbf{x}) \quad \forall (\mathbf{v}, \mathbf{w}) \in \mathcal{V} \end{aligned}$$

□

Birbil SI, Frenk J, Gromicho JA, Zhang S (2013) A network airline revenue management framework based on decomposition by origins and destinations. *Transportation Science* 48(3):313–333.

Bront JJM, Méndez-Díaz I, Vulcano G (2009) A column generation algorithm for choice-based network revenue management. *Operations Research* 57(3):769–784.

Gallego G, Phillips R (2004) Revenue management of flexible products. *Manufacturing & Service Operations Management* 6(4):321–337.

Gallego G, Ratliff R, Shebalov S (2010) A new formulation for choice based rm under a general attraction model. *AGIFORS 50th symposium, Nice, France* .

Gallego G, Ratliff R, Shebalov S (2015) A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research* 63(1):212–232.

Guillermo G, Richard R, Sergey S (2011) A general attraction model and an efficient formulation for the network revenue management problem. Technical report, Columbia University.

Kunnumkal S, Talluri K (2015) On a piecewise-linear approximation for network revenue management. *Mathematics of Operations Research* 41(1):72–91.

Kunnumkal S, Topaloglu H (2008) A refined deterministic linear program for the network revenue management problem with customer choice behavior. *Naval Research Logistics* 55(6):563–580.

Kunnumkal S, Topaloglu H (2010) Computing time-dependent bid prices in network revenue management problems. *Transportation Science* 44(1):38–62.

Liang D, Ratliff R, Remenyi N (2017) Robust revenue opportunity modeling with quadratic programming. *Journal of Revenue and Pricing Management* 1–11.

Liu Q, Van Ryzin G (2008) On the choice-based linear programming model for network revenue management. *Manufacturing & Service Operations Management* 10(2):288–310.

Luce RD (1959) *Individual Choice Behavior a Theoretical Analysis* (John Wiley and sons).

McGill JI, Van Ryzin GJ (1999) Revenue management: Research overview and prospects. *Transportation science* 33(2):233–256.

Meissner J, Strauss A, Talluri K (2013) An enhanced concave program relaxation for choice network revenue management. *Production and Operations Management* 22(1):71–87.

Smith B, Penn C (1988) Analysis of alternate origin-destination control strategies. *AGIFORS PROCEEDINGS*.

Smith BC, Leimkuhler JF, Darrow RM (1992) Yield management at american airlines. *interfaces* 22(1):8–31.

Strauss AK, Klein R, Steinhardt C (2018) A review of choice-based revenue management: Theory and methods. *European Journal of Operational Research* ISSN 0377-2217, URL `http://dx.doi.org/https://doi.org/10.1016/j.ejor.2018.01.011`.

Talluri KT, Van Ryzin GJ (2006) *The theory and practice of revenue management*, volume 68 (Springer Science & Business Media).

Van Ryzin G, Vulcano G (2008) Simulation-based optimization of virtual nesting controls for network revenue management. *Operations Research* 56(4):865–880.

Vossen TW, Zhang D (2015) Reductions of approximate linear programs for network revenue management. *Operations Research* 63(6):1352–1371.

Vulcano G, van Ryzin G, Chaar W (2010) Om practicechoice-based revenue management: An empirical study of estimation and optimization. *Manufacturing & Service Operations Management* 12(3):371–392.

Vulcano G, Van Ryzin G, Ratliff R (2012) Estimating primary demand for substitutable products from sales transaction data. *Operations Research* 60(2):313–334.

Walczak D, Boyd EA, Cramer R (2012) Revenue management. *Quantitative Problem Solving Methods in the Airline Industry*, 101–161 (Springer).