# New Formulations for Price and Ticket Availability Decisions in Choice-based Network Revenue Management

Michael C. Clough[a], Timothy L. Jacobs[b], Esma S. Gel[a]

[a] *School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, AZ 85287*
[b] *Crew Systems and Technologies, American Airlines, Tempe, AZ 85281*

**Abstract**

We consider the network revenue management (RM) problem under a MultiNomial Logit (MNL) discrete choice model. We present two mathematical formulations that build on previous research, focusing on reducing the size and complexity of the problem while maintaining an accurate representation of passenger choice and industry practices. The first formulation addresses the computational issues of traditional network revenue management by introducing a Mixed Integer Program (MIP) that reduces the number of variables needed to consider while maintaining revenue performance versus other models. The second formulation eliminates the need for static pricing of fare classes, as seen in most RM literature, and considers a continuous pricing decision linked with ticket availability, directly incorporating passenger sensitivities to price and other attributes. Both models perform well in example test cases, improving on popular leg- and network-based RM methods used in practice.

*Keywords:* Airline Revenue Management (RM), MultiNomial Logit (MNL) Choice Model, Mixed Integer Programming (MIP), Dynamic Pricing

## 1. Introduction

From its inception, airlines have used Revenue Management (RM) techniques to improve their revenue performance or yield by optimizing the passenger mix through fare class seat availability or bid price hurdle rates. Both leg-based and Origin-Destination (O&D)-based approaches have used a common assumption that the passenger demand associated with a given flight or O&D path and fare class are known and forecast independent of other options within the market. For example, demand forecasts for full fare passengers on BOS-PHX-LAX are based on historical traffic observations on that specific path and do not explicitly account for the passenger demand associated with other paths in the market like BOS-ORD-LAX. In addition, most RM approaches used in practice today assume that fare classes are mutually exclusive of one another when optimizing seat allocations or bid prices. These assumptions preclude the demand interactions between different routes, fare classes and competition from other carriers in the same markets, and limit the quality of the optimization results and controls. To remove the limitations of these assumptions, the demand forecasts and optimization must consider the interactions between the different fare classes and routes available to potential passengers at the point of sale.

In this paper we propose two mathematical modeling formulations that explicitly incorporate the fare class and routing interactions using a MultiNomial Logit (MNL) choice model. The first

*Email addresses:* `Michael.Clough@asu.edu` (Michael C. Clough), `Tim.Jacobs@aa.com` (Timothy L. Jacobs), `Esma.Gel@asu.edu` (Esma S. Gel)

formulation, which we refer to as the Choice-based Mixed Integer Program (CMIP), represents an alternative formulation to the Choice-based Deterministic Linear Program (CDLP), as proposed by Liu and van Ryzin (2008). The CMIP assumes independent O&D markets can be solved independently, greatly reducing the size of the problem. The second formulation, referred to as the Price-Dynamic Choice-based Mixed Integer Program (PCMIP), takes a different approach to RM by optimizing ticket availability and price. Most choice-based revenue management models assume static pricing when selecting the availability of fare classes while ignoring the effects pricing has on passenger preference. The PCMIP directly incorporates price-dependent sensitivities, as well as other passenger preferences, while simultaneously optimizing fare class availability and prices. The added flexibility of setting prices within the optimization model nets higher revenue at a cost of complexity.

The remainder of this paper is organized as follows. Section 2 reviews the recent literature and developments associated with incorporating passenger choice into the RM process, Section 3 presents the CMIP formulation, Section 4 solves an illustrative example and compares performance to other models, Section 5 solves larger examples and compares performance, Section 6 introduces the PCMIP formulation and examples, and Section 7 highlights the conclusions and potential future research directions.

## 2. Literature Review

Although many RM models have been developed over the past 30 years, the following literature review focuses on the choice-based demand approaches that aim to model consumer behavior more accurately for the network RM problem.

To provide a framework for choice-based modeling, we first present an overview of some of the leg-and O&D-based methods available. Two leg-based independent demand methods worth noting, however, are that of Littlewood's 1972 paper (which was later republished in 2005), and Belobaba (1989). Often, Littlewood (2005) is cited as being one of the first models to solve the RM problem. His model determined the necessary protection limits by comparing two products' expected demand and fares. Littlewood (2005) proposed a rule, termed Littlewood's Rule, which determines protection levels for the higher fare classes. Belobaba (1989) expanded on Littlewood's research, and created the Expected Marginal Seat Revenue (EMSR) model. His first model, EMSR-a, executed pair-wise comparisons to determine how many seats to reserve for higher fare classes. The EMSR-a model could compare any number of pairs, and would aggregate protection levels as it moved up the fare class buckets (Belobaba, 1989). Later, Belobaba expanded on his own model, creating the EMSR-b methodology. EMSR-b, instead of aggregating over protection levels, aggregates the demand for higher paying passengers, and calculates a weighted fare for them (Belobaba, 1992). This weighted fare is then used for a comparison, and protection levels are calculated. For a more complete history of independent demand RM models, we refer the reader to Weatherford and Ratliff (2010). These independent demand models, like that of Littlewood (2005) and Belobaba (1989), were computationally efficient, but they lack the network interactions and competitive effects present in today's complex airline markets. To this end, research moved towards choice-based modeling techniques for solving the network RM problem.

Two of the first, and possibly most influential, papers in choice-based modeling for network RM were Gallego, Iyengar, Phillips, and Dubey (2004), and Talluri and van Ryzin (2004). In Gallego et al. (2004), the authors propose a linear program that solves the network RM problem with a general discrete choice model. Using the probability of a purchase as a parameter, the model determines the amount of a time that each set of policies, defined by the itinerary and fare, is to be offered. This method maximizes the revenue across the entire network, by selecting a subset of available policies, constrained by the available space consumed on a leg (Gallego et al., 2004). This model is later developed into the Choice-based Deterministic Linear Program (CDLP), and has become a benchmark for the testing of newer models in the field of network RM. In particular, the CDLP determines the optimal amount of time to offer a set of policies, $S$. This set $S$ is comprised of open and closed policies for each O&D fare class combination within the network.

Talluri and van Ryzin (2004) formulated the problem as a dynamic program, which modeled the probabilities of different purchases using a general discrete choice model, and determined which policy sets to offer based on the available capacity, similar to the model in Gallego et al. (2004). Talluri and van Ryzin introduced the concept of efficient sets, which allowed for search techniques to manage the complex nature of the solution space. From this point on, research in choice-based RM has gone in one of three directions: solution methodologies for the CDLP, approaches to solve dynamic programming formulations, or formulations that are new altogether.

### *2.1. CDLP Solution Methodologies*

In Kunnumkal and Topaloglu (2008), the authors created an alternative form of the CDLP, and obtained better results through the solution of the primal. Liu and van Ryzin (2008) expanded on the original CDLP, and developed an iterative approach by applying the bid prices generated from the CDLP to a leg-level decomposition approach to Talluri and van Ryzin (2004)'s dynamic program. The results of their method provided capacity and time-dependent bid prices, which are useful for industry application. The authors also expanded on the notion of efficient sets, and applied them to the CDLP, generating methods for solving this complex problem (Liu and van Ryzin, 2008). Shortly thereafter, Bront, Mendez-Diaz, and Vulcano (2009) developed a column generation algorithm to solve the CDLP, in the special case of having non-disjoint markets. Their model considered situations where market demand can overlap, and competition can arise between O&D's as well as pricing options. The authors provide details on how to solve the column generation algorithm for the CDLP, as well as provide two methods for solving the subproblem of determining which set to introduce into the reduced primal problem (Bront et al., 2009). Talluri (2011) relaxed the CDLP, and solved a Segment-based Deterministic Concave-Program (SDCP), which provided looser upper bounds to the original problem. Following this relaxation, Meissner, Strauss, and Talluri (2013) expanded on the model to include constraints on the product selections, creating the extended-SDCP.

### *2.2. Approaches to Solve Dynamic Programming Formulations*

Although Talluri and van Ryzin (2004) provided one of the first models utilizing dynamic programming formulations for network RM, others also explored this approach. Zhang and Cooper (2005) offered a different perspective, and created a Markov Decision Process (MDP) formulation for cases where multiple flights are being offered between O&D's in short time spans. Later, **?** developed an MDP model that allowed for substitution to take place between flights. Although both of these MDP formulations could be solved via dynamic programming, more efficient methods were found in the form of inventory-pooling (Zhang and Cooper (2005)) and heuristics (Zhang and Cooper (2009)). Some models were developed in conjunction with dynamic programs, like in Farias and van Roy (2007) and Adelman (2007). In Farias and van Roy (2007), the authors model the network RM problem as a dynamic program, and then solve it with a linear programming approximation. Their model is unique, as it solves for the bid prices directly, rather than producing policy-based decisions. Adelman (2007) utilized an affine approximation for the value function of his dynamic program. Similar to Farias and van Roy (2007), his model determines the bid prices for the network, and generates a dynamic set of bid prices. Kunnumkal and Topaloglu (2010) developed their own dynamic programming decomposition method, which solves the single-leg decomposition with revenue estimates for each leg in an itinerary. The revenue estimates were generated ahead of time through an optimization model utilizing the choice-based modeling schema.

Some models shift the focus from policy decisions and generate solutions that determine seat allocation policies. Huang and Liang (2011) developed a dynamic programming formulation, which they solve by estimating the value function of the dynamic program (DP) with a sampling technique. Their model solves for the seat control policies, rather than open or closed fare class decisions. In Zhang (2011), the authors proposed an alternative way to solve the dynamic programming formulation of Talluri and van Ryzin (2004), and provided better bounds on the optimal solution for the original problem. Kunnumkal (2011) took a different approach to solving the dynamic program, and offered two approximation models for solving the choice-based network RM. Lagrangian relaxations were done for both methods, one based on relaxing the flight leg capacities and the other based on perfect demand information. His model generates capacity-dependent policies, similar to that of the original dynamic programming formulation (Kunnumkal, 2011). Another unique formulation is found in Meissner and Strauss (2012), in which they develop a dynamic programming formulation that takes into account inventory sensitive bid prices. Their model estimates the value function of an MDP to determine capacity-dependent bid prices.

### *2.3. Alternative Formulations*

Other models different from the typical dynamic programming formulations and CDLP were also developed. van Ryzin and Vulcano (2008) developed an optimization model that solves the choice model independently from the optimization model itself, creating an easier and quicker solution methodology to the problem. Their model solves for nested protection levels, rather than policy or bid price optimization. Chaneton and Vulcano (2011) sought to simplify the problem by changing the formulation of the choice-based demand models. They estimate the choice-based demand by applying a linear approximation to the demand, creating a continuous function with a stationary point found using a sub-gradient algorithm. Their model allows for partially accepted itineraries in which the passenger requests can be accepted on legs within the itinerary, but not the entire itinerary itself (Chaneton and Vulcano, 2011). This approach is similar to that of Topaloglu (2009), in which a bid price solution methodology is developed by applying a leg-level decomposition approach. Chen and de Mello (2010) developed a formulation that modeled the buy-up behavior directly. Their model allows for passengers to step up in fare classes if their desired fare class is unavailable. From this buy up pattern, the authors were able to determine a demand stream, which then was used to solve a set of optimization problems.

Gallego, Ratliff, and Shebalov (2011) introduced the generalized attraction model, which can be applied to any demand input. The independent demand, as well as basic attraction models, were found to be special cases of this generalized attraction model. They develop their model to combat the complexity of the CDLP, resulting in a new formulation known as the Sales Based Linear Program (SBLP). Another mathematical model, in the form of a mixed integer program, was developed by Meissner and Strauss (2010). Their model solved for both policy decisions on restricted fare classes (i.e., fare classes in which discrete fares are determined in advance), as well as pricing decisions on unrestricted fare classes (i.e., fare classes in which a continuous range of available prices exist). Kunnumkal (2011) developed a two-step method for solving the network RM problem. His method first determines which policies are optimal through a choice-based mixed integer program, followed by a linear program that determines the marginal value of seats. He argues that the linear program can be randomized, and provides good solutions compared to those obtained by the CDLP (Kunnumkal, 2011). Meissner and Strauss (2011) also developed a mixed integer program, under the assumption that market segmentation is weak. This creates ambiguity in the demand stream, and their model proved to be computationally intractable. The authors provided alternative solution methods for solving their mixed integer program, citing cases where shorter run times were more advantageous than computational accuracy.

*2.4. Dynamic Pricing*

Dynamic pricing in a choice-based environment is a relatively new area of research within RM. Prior to choice-based demand models being the in the spotlight, some research had focused on the importance of pricing, as seen in Jacobs, Ratliff, and Smith (2010), where the authors investigate the relationship between pricing and revenue management controls. Jacobs et al. (2010) consider the impact capacity has on the dynamic pricing problem, creating a "price balance statistic" used for evaluating the quality of a strategy and finding the optimal mix of pricing, scheduled capacity, and RM controls.

Choice-based dynamic pricing research has taken a different approach. Aydin and Ryan (2000) consider a retail setting where consumers choose products based on a pre-defined selection and pricing setup. Zhang and Cooper (2009) introduce a Markov decision process to address dynamic pricing in an MNL environment. They consider multiple substitutable flights in a single O&D market, and show their model is intractable for realistic settings. In a similar approach, Dong, Kouvelis, and Tian (2009) develop a dynamic programming formulation in a retail environment. They consider an environment with a long lead time and short selling environment, where the retailer must determine both inventory and pricing. More recently, Zhang and Lu (2013) introduce a dynamic program that addressed pricing in airline RM. The authors introduce a dynamic program for dynamic pricing and offer a non-linear programming approximation approach. They compare their methods against both static pricing models and other choice-based demand models, concluding that dynamic pricing could have substantial gains versus their static counterparts. In the following sections we present two new formulations that improve on the research reviewed here.

## 3. Mathematical Model

We consider a network with legs $l \in L$, and containing multiple markets defined by set $J$. A market $j$ represents an O&D pair. There exists a set of policies, $I_j$, defined for each market $j \in J$, where a policy $i \in I_j$ is defined as a pair of itinerary and fare class assignments. Since each market can have multiple itineraries (i.e., paths) and fare classes, we define the set $K_{ji}$ that contains all defined fare class-itinerary pairs for market $j \in J$ and policy $i \in I_j$. The planning horizon associated with this model can be viewed as the time to departure. We discretize time into periods and denote the index set of time periods by $T$. Having defined the network parameters, the parameters and decision variables for our mathematical formulation are formally defined in Table 1.

| | |
|---|---|
| $\lambda_t$ | Number of customer requests for flights to the network in period $t$, $t \in T$ |
| $P_j$ | Probability of an arrival for market $j$, $j \in J$ |
| $S_{ji}$ | Probability that a purchase is made for market $j$ under policy $i \in I_j$ |
| $P_{k\|j,i}$ | Probability of a purchase on fare class-itinerary $k \in K_{ji}$, given purchase is made for market $j$ under policy $i \in I_j$ |
| $R_k$ | Revenue for a purchase on fare class-itin. $k$ given policy $i$ is used |
| $A_{kl}$ | Binary parameter representing consumption of leg $l$ for fare class-itin. $k \in K_{ji}$ |
| $c_l$ | Capacity of leg $l \in L$ available at the beginning of the planning horizon |
| $Z_{jit}$ | Fraction of period $t$ demand for market $j$ served under policy $i \in I_j$ |
| $X_{jit}$ | Binary decision variable to use policy $i \in I_j$ for market $j$ in period $t \in T$ |

Table 1: Table of notations used in CMIP

As an example, consider the network given in Figure 1, which is the same example used in Bront et al. (2009) and Liu and van Ryzin (2008). The network contains three nodes and leg capacities of 10, 5 and 5 seats for legs 1, 2, and 3, respectively. Each leg represents a single flight, thus there are no parallel flights for this network. Table 2 includes further data on this example; eight products were defined by what the authors refer to as "O&D path" and fare class combinations.

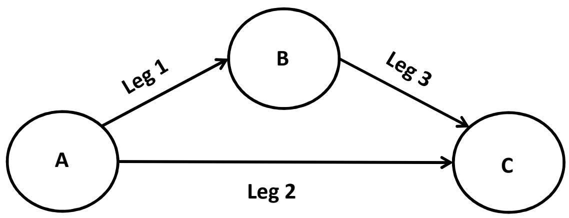


Figure 1: Illustrative example: Three-leg network

Table 2: O&D paths and fare classes for the illustrative network example (Liu and van Ryzin, 2008)

| Product | Origin-Dest. Path | Class | Fare |
|---|---|---|---|
| 1 | A - C | High | \$1200 |
| 2 | A - B - C | High | \$800 |
| 3 | A - B | High | \$500 |
| 4 | B - C | High | \$500 |
| 5 | A - C | Low | \$800 |
| 6 | A - B - C | Low | \$500 |
| 7 | A - B | Low | \$300 |
| 8 | B - C | Low | \$300 |

| Segment | Arrival Rate | Consideration Set | Preference Vector | Utility of No Purchase |
|---|---|---|---|---|
| 1 | 0.15 | $\{1, 5\}$ | (5, 8) | 2 |
| 2 | 0.15 | $\{1, 2\}$ | (10, 6) | 5 |
| 3 | 0.20 | $\{5, 6\}$ | (8, 5) | 2 |
| 4 | 0.25 | $\{3, 7\}$ | (4, 8) | 2 |
| 5 | 0.25 | $\{4, 8\}$ | (6, 8) | 2 |

Table 3: Data on demand and customer preferences for the illustrative network example (Liu and van Ryzin, 2008)

Table 3 includes data on customer preferences and utilities of the different products for each of the five segments. The preference vectors represent the utility that the products in the consideration set provides for the segment. For example, segment 1 has a consideration set of $\{1, 5\}$, and a preference vector of $(5, 8)$. This means that the first segment has a utility of 5 for product 1 (i.e., the A-C itinerary with a cost of \$1200) and a utility of 8 for product 5 (i.e., the A-C itinerary with a cost of \$800). The larger utility value implies that customer segment 1 prefers the A-C itinerary with a cost of \$800 over the A-C itinerary with a cost of \$1200. The no-purchase utility corresponds to the option of not purchasing either product.

For our formulation, the network illustrated in Figure 1 would result in three markets, i.e., $J = \{\text{AB}, \text{BC}, \text{AC}\}$. Note that market AC contains two different itineraries, the direct path from A to C as well as the path containing the connection A-B-C. For each market, $j \in J$, there is a set of available policies, $I_j$. Each policy lists the available options, each defined by an itinerary and the fare classes. For instance, the AC market has two competing paths, AC and ABC, thus the available policies for the AC market would be all combinations of high and low fare class options, as well as the possibility of closed itineraries.

When implementing policies in real life airline RM systems, certain fare classes are nested within their lower fare class counterparts. For instance, the policy containing AC high and AC low open simultaneously would be equivalent to opening only AC low, since policy implementation is generally based on bid prices. That is not to say the higher fare class is closed, just that there is no situation where an airline would refuse a higher paying passenger just because the policy only defines AC low as being open. This natural nesting among the fare classes eliminates the need to separately define policies in which AC high and AC low are open simultaneously. The elimination of these simultaneous policies, however, removes any buy up potential, thus the model makes a conservative assumption that buy up is negligible.

Finally, the set $K_{ji}$ includes all fare class-itinerary pairs defined for market $j$ under policy $i \in I_j$. The first component, market, is defined by the available market set $J$. The policy component is defined by the set of available policies $I_j$. The itinerary path is determined by the structure of the network itself. The combination of appropriate market-policy-itinerary path groupings generates the set $K$, which is referred to as the fare class-itinerary. For the illustrative example, the values of these sets can be seen in Table 4, in the fourth column.

We can now determine the values of our parameters $\lambda_t$, $P_j$, $A_{kl}$, $R_k$, $P_{k|j,i}$ and $S_{ji}$ for the illustrative network example. For this example, we assume that the arrival rate, or the number of unit demand arrivals per period, stays constant at the values listed under the column labeled "Arrival Rate" in Table 3 for each customer segment. We actually use these values for $\lambda_t P_j$ for each market. From Table 3 the arrival rate for market AC, for example, will be equal to the sum of the arrival rate values listed for customer segments 1, 2, and 3, i.e., $\lambda_t P_{\text{AC}} = 0.50$ customers for each time period. Similarly, $\lambda_t P_{\text{AB}} = 0.25$ and $\lambda_t P_{\text{BC}} = 0.25$ for all $t \in T$. Note that in the example, the time is scaled so that the total arrival rate, $\lambda_t = 1$.

The values of $A_{kl}$ can be determined by examining the network and fare class-itineraries. If fare class-itinerary $k \in K_{ji}$ consumes space (i.e., 1 unit of capacity) on leg $l$, then $A_{kl}$ is assigned a value of 1. Otherwise, $A_{kl}$ is assigned a value of zero. Since $R_k$ represents the revenue earned for fare class-itinerary $k$ being purchased, those values can be read directly from the pricing table.

The values for $P_{k|j,i}$ were determined through conditioning. For instance, if the policy available was AC High/ABC High, then a fraction of the purchases would purchase the AC itinerary while others would purchase the ABC itinerary. Since we are conditioning on the fact that a purchase was made, we merely need to determine what fraction of passengers purchased the AC High option (or, "fare class-itinerary") and what fraction of passengers purchased the ABC High option. To do this, we must first determine which customer segments, as defined by Table 3, prefer each of the options. For the AC High option (defined as product 1), we can see that customer segments 1 and 2 have utility values for this product (as defined by their consideration sets). Likewise, for the ABC High option (defined as product 2), we can see that only customer segment 2 has preference for this product.

We first calculate $S_{ji}$, which denotes the probability of a purchase by an arriving market $j$ customer under policy $i \in I_j$. For the above example of policy AC High/ABC High for the AC market, the probability of purchase by a market AC customer under policy $i \in I_j$ can be calculated as a weighted average of the purchase probabilities that can be calculated from the given utilities of the products available under the AC High/ABC High policy, and that of no purchase. That is,

$$S_{\text{AC,AC High/ABC High}} = \frac{5}{5+2}\left(\frac{0.15}{0.50}\right) + \frac{10+6}{10+5+6}\left(\frac{0.15}{0.50}\right) + 0\left(\frac{0.20}{0.50}\right) = 0.443 \; ,$$

by using the given utilities and conditioning on the event that the AC customer is from segment 1, 2 or 3, respectively.

Note that the utility values used above represent the respective $e^{u_l}$ terms used in the MNL model. Essentially, the ratio $5/(5+2)$ can be rewritten in traditional MNL format as $e^{1.609}/(e^{1.609} + e^{0.693})$, where the values of 1.609 and 0.639 would represent the MNL utilities for the AC market for AC High and the no purchase option, respectively. We use this simplified notation for ease of representation, following the tradition set by previous papers in this area.

Then, the probability that the customer bought fare class-itinerary $k \in K_{\text{AC,AC High/ABC High}}$ given that a purchase was made by an AC customer under policy AC High/ABC High can be calculated by

$$P_{\text{AC High}|\text{purchase under AC High/ABC High}} = \frac{\frac{5}{5+2}\left(\frac{0.15}{0.50}\right) + \frac{10}{10+5+6}\left(\frac{0.15}{0.50}\right)}{0.443} = 0.806 \ .$$

Similarly, we can calculate

$$P_{\text{ABC High}|\text{purchase under AC High/ABC High}} = \frac{\frac{6}{10+5+6}\left(\frac{0.15}{0.50}\right)}{0.443} = 0.194 \ ,$$

or, simply by observing that for this policy with two fare class-itinerary options,

$$P_{\text{ABC High}|\text{purchase under AC High/ABC High}} = 1 - P_{\text{AC High}|\text{purchase under AC High/ABC High}} \ .$$

Continuing similarly, we obtain the values in Table 4 for this illustrative example. Note that in columns one through five, the table provides the markets (i.e., $j \in J$), the policies defined for each market, (i.e., $i \in I_j$), the purchase probability for market $j$ under each policy $i \in I_j$, (i.e., $S_{ji}$ values), the defined fare class-itineraries (i.e., set $K_{ji}$) for each market $j$ under policy $i \in I_j$, and finally, the conditional probability that the option given by a particular fare class-itinerary $j \in K_{ji}$ will be selected, given that a purchase for market $j$ was made under policy $i \in I_j$.

| Market (Set $J$) | Policies (Set $I_j$) | $S_{ji}$ | Fare Class-Itineraries (Set $K_{ji}$) | $P_{k\|ji}$ |
|---|---|---|---|---|
| AB | AB High | 0.667 | AB High | 1 |
| | AB Low | 0.857 | AB Low | 1 |
| BC | BC High | 0.750 | BC High | 1 |
| | BC Low | 0.875 | BC Low | 1 |
| AC | AC High/ABC High | 0.443 | AC High<br>ABC High | 0.806<br>0.194 |
| | AC High/ABC Low | 0.761 | AC High<br>ABC Low | 0.556<br>0.444 |
| | AC Low/ABC High | 0.809 | AC Low<br>ABC High | 0.681<br>0.319 |
| | AC Low/ABC Low | 0.607 | AC Low<br>ABC Low | 0.773<br>0.227 |
| | AC High/ABC Closed | 0.414 | AC High<br>ABC Closed | 1<br>0 |
| | AC Low/ABC Closed | 0.580 | AC Low<br>ABC Closed | 1<br>0 |
| | AC Closed/ABC High | 0.279 | AC Closed<br>ABC High | 0<br>1 |
| | AC Closed/ABC Low | 0.347 | AC Closed<br>ABC Low | 0<br>1 |

Table 4: Set definitions and calculated parameters for the illustrative example adapted from Liu and van Ryzin (2008)

Finally, our formulation uses the following decision variables. $X_{jit}$ represents the binary variable that takes on a value of 1 if the decision is to use policy $i$ for market $j$ in period $t$, and $Z_{jit}$ represents the fraction of market $j$ demand served under policy $i$ in period $t$. To expand on the variable $Z_{jit}$, consider an example where, for a given market policy $i$ and five time periods, $Z_{jit}$ takes on values of $(0, 0, 1, 1, 0.172)$. This vector would represent the following set of decisions. For time periods 1 and 2, policy $i$ is not available and no demand for market $j$ would be served under this policy. During time periods 3 and 4, policy $i$ is available, and any arriving demand for market $j$ would be served. Finally, during time period 5, policy $i$ is available, but only 17.2% of the potential demand should be served. Note that the term "served" here does not necessarily mean that they will be purchasing a ticket; it basically means that they get to consider the various options available to them for market $j$, under policy $i \in I_j$. As a result of this consideration, they may or may not purchase a ticket on market $j$. Using these two decisions variables, and the parameters previously defined, we formulate the *Choice-based Mixed Integer Program (CMIP)* as follows.

$$\text{Maximize} \qquad \sum_{t \in T} \sum_{j \in J} \lambda_t P_j \sum_{i \in I_j} Z_{jit} \, S_{ji} \sum_{k \in K_{ji}} R_k P_{k|j,i} \tag{1}$$

Subject to:

$$\sum_{t \in T} \sum_{j \in J} \sum_{i \in I_j} \sum_{k \in K_{ji}} \lambda_t P_j Z_{jit} S_{ji} P_{k|j,i} A_{kl} \leq c_l, \qquad \text{for all } l \in L, \tag{2}$$

$$\sum_{i \in I_j} X_{jit} \leq 1 \ , \qquad \text{for all } j \in J, t \in T, \tag{3}$$

$$Z_{jit} \leq X_{jit} \ , \qquad \text{for all } j \in J, i \in I_j, t \in T, \tag{4}$$

$$Z_{jit} \in R^+, X_{jit} \in \{0, 1\}, \qquad \text{for all } j \in J, i \in I_j, t \in T. \tag{5}$$

The objective function (1) represents the total expected revenue across all time periods for the decision variable $Z_{jit}$. The objective function can be broken into two main components: the arrival rate of demand per market and the expected revenue for a given policy. The first component, the arrival rate of demand per market, is the product of the expected number of customer requests in period $t$, (i.e., $\lambda_t$) and the probability that an arrival demands a ticket for market $j$, (i.e., $P_j$). This product, $\lambda_t P_j$, represents the expected number of arrivals in time period $t \in T$ for market $j$. The second component, the expected revenue obtained from market $j$ under policy $i \in I_j$, $E[R_j(i)]$, assuming that customer preferences remain unchanged throughout the planning horizon, can be calculated as follows.

$$\begin{aligned} E[R_j(i)] &= (1 - S_{ji})\, 0 + S_{ji} \, E[\text{Revenue} \mid \text{purchase on market } j \text{ under policy } i \in I_j] \\ &= S_{ji} \sum_{k \in K_{ji}} R_k P_{k|j,i} \ , \end{aligned} \tag{6}$$

where $R_k$ denotes the revenue from a sale on fare class-itinerary $k \in K_{ji}$, $P_{k|j,i}$ denotes the conditional probability of a purchase on fare class-itinerary $k \in K_{ji}$, given a purchase for market $j$ is made under policy $i \in I_j$, and finally, $S_{ji}$ is the probability that a purchase for market $j$ is made under policy $i \in I_j$.

Constraint set (2) ensures that the capacity constraints on the legs are not violated. The continuous decision variable, $Z_{jit}$, allows for the partial accommodation of demand, and facilitates the determination of bid price values for the flight leg, $l \in L$. The purpose of a bid price is to represent the marginal value of an extra seat on a given leg. In the event constraint set (2) is binding, we can increase the capacity of a leg to determine what impact this increase would have on the objective function. The value of variable $Z_{jit}$ could be increased a marginal amount, no greater than one, if the current value is less than one. In the event $Z_{jit}$ was already at a value of one, then the model could select a different $Z_{jit}$ to improve the objective function. This would force the constraint to be binding, again, and the objective function, which also contains the $Z_{jit}$ variable, would increase appropriately. This increase would be analogous to the shadow price of a linear program, thus it can be used as the marginal value of a seat on a given leg. The marginal value of a seat is then translated into the bid price for the leg, and could be used for bid price based control policies.

One difference between our formulation and other formulations stems from the fact that the other models account for all market combinations in the form of sets, whereas CMIP combines policies for O&D markets to determine the overall policy for the network. For instance, the CDLP selects which sets are optimal, while the CMIP model selects, individually, which O&D fare class combinations optimize our revenue. Since the CMIP focuses on a market-by-market level, the total number of variables for the problem is greatly reduced, which results in reasonable solution times for larger networks, as we show for a large network instance. As mentioned above, the complexity of the network greatly impedes the quality of the solutions that one can obtain from the CDLP formulation within a reasonable run time. Hence, having fewer variables in the CMIP formulation allows for the modeling of larger networks with solution times that are implementable for industry use.

To illustrate the magnitude of the difference in variables, consider the small network that we considered earlier, depicted in Figure 1. In this network, for a single time period, the CMIP has a total of 24 variables and 18 constraints. The CDLP, for the same scenario, has a total of 255 variables and 4 constraints. As we increase the number of time periods, the CMIP increases in both variable and constraint totals, while the CDLP does not. The advantage of the CMIP, however, is when the complexity of the network is increased. Adding just one more node with respective high and low fare classes and connections (assuming this node is independent of the markets currently in the network), would only increase the CMIP to 28 variables and 22 constraints. This same network for the CDLP formulation would have 16,383 variables and 5 constraints. The CDLP has a smaller constraint set, yet the variable space is exponentially increasing as the complexity of the network is increased. The CMIP has a much smaller variable space, and a reasonably sized constraint space. As the complexity of the network is increased, the variable and constraint space does not increase in an exponential fashion; the total number of variables for the CMIP, however, would increase multiplicatively for each additional time period.

## 4. Solution of the Illustrative Example

We solved the CDLP and CMIP formulations for the example presented above (see Figure 1), and implemented the obtained policy decisions in a simulation to compare the performance of the two approaches. We used AMPL and Gurobi 5.0.1 to build and solve the formulations. We assumed that buy-up did not happen, as there only is a very small probability that a passenger will purchase a higher priced ticket if a lower priced ticket is available.

We programmed the simulation in MATLAB following a traditional Monte Carlo simulation approach. First, the simulation takes bid prices as the control, and generates the available fare classes in each time period. We assume a stationary arrival rate of customers (I.e., $\lambda_t = \lambda$ for every time period $t \in T$) and generate exponential customer interarrival times with this rate. For each customer, the simulation model generates the identity of the market that the customer is interested in purchasing, as well as what, if anything, the customer purchases using $P_j$, $S_{ji}$ and $P_{k|ji}$ in a relatively standard random number generation scheme. In case of a purchase, the capacity of the legs for the requested itinerary is reduced, and the total revenue is updated. We ran the simulation for 2000 iterations for each of the network instances tested. This simulation is used for all of the results following the illustrative example.

The solutions from the two formulations generated similar bid prices across the majority of the tests. The average total revenue values obtained with the two approaches were also comparable. As seen from Table 5, the CDLP and CMIP reach identical bid prices in every case except for $T = 5$ and $\lambda = 5$. Due to the nature of this illustrative example, the leg AC only has pricing options of \$800 and \$1200. Hence, having a bid price of \$750 implies both pricing options are to be open. Note that obtaining a bid price of \$0 (which can be observed in the case of the CMIP for $T = 5$ and $\lambda = 5$), would have the same effect as having a bid price of \$750. This implies the CDLP and the CMIP generate identical bid price control strategies across all combinations for this example.

| | | CMIP Bid Prices | | | CDLP Bid Prices | | | Obj. Fn. | Obj. Fn. |
|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | T | AB | AC | BC | AB | AC | BC | CMIP | CDLP |
| 1 | 1 | \$0 | \$0 | \$0 | \$0 | \$0 | \$0 | \$515 | \$497 |
| | 5 | \$0 | \$0 | \$0 | \$0 | \$0 | \$0 | \$2,577 | \$2,485 |
| | 10 | \$0 | \$0 | \$0 | \$0 | \$0 | \$0 | \$5,155 | \$4,971 |
| 5 | 1 | \$0 | \$0 | \$0 | \$0 | \$0 | \$0 | \$2,577 | \$2,485 |
| | 5 | \$0 | \$0 | \$500 | \$0 | \$750 | \$500 | \$10,664 | \$10,064 |
| | 10 | \$300 | \$1200 | \$500 | \$300 | \$1,200 | \$500 | \$13,168 | \$13,167 |
| 10 | 1 | \$0 | \$0 | \$0 | \$0 | \$0 | \$0 | \$5,155 | \$4,971 |
| | 5 | \$300 | \$1,200 | \$500 | \$300 | \$1,200 | \$500 | \$13,168 | \$13,167 |
| | 10 | \$500 | \$1,200 | \$500 | \$500 | \$1,200 | \$500 | \$13,500 | \$13,500 |

Table 5: Results from CMIP and CDLP formulations

For the results presented in the last two columns of Table 5, the CMIP increased expected revenue by an average of 2.72% when compared to the CDLP. Although these models are similar, the way they handle expected traffic is different. The CDLP uses the direct probability of purchases generated from the MNL choice model. The CMIP uses the probability of purchases for a given fare class-itinerary, conditioned on a purchase being made. These differences are subtle, yet impact the CDLP and CMIP objective functions, so comparison on these values alone is insufficient. To this end, we simulated the policy for the instance of $\lambda = 5$ and $T = 5$, to see whether the differences between revenues would continue to hold. This problem instance was chosen since this was the only case where the bid prices differed between the two models. We determined the 95% confidence intervals around the expected revenue for both simulations. The CMIP resulted in an interval of $(\$6,959, \$10,071)$, while the CDLP resulted in an interval of $(\$6,950, \$10,088)$. As expected, the results we observed for the CMIP and CDLP were very close to one another. Based on the results of the simulation, we conclude that, in this set of problem instances, the CDLP and CMIP provide similar solutions, and can be used interchangeably.

*4.1. Implementing the Solution*

One advantage of the CMIP formulation is the fact that both the policy and bid price controls are useful for industry application. The solution to the CMIP indicates which policy should be offered for each market during a particular time period. In the three-leg network problem instance, the solution would instruct, for instance, to open the high fare class for itinerary ABC during time periods 1, 2 and 3. Additionally, it would indicate to open the low fare class for itinerary AB during time periods 1 and 2, while opening the high fare class for time period 3. A reservation system could directly interpret this decision to open and close these particular fare classes, following the guidance of the CMIP solution. The reservation system could then generate cut offs for certain fare classes and calculate protection levels based on the airline's current system, if necessary.

An effective bid price control can be derived from the solution to this formulation as well. Effectively, the lowest open fare class-itinerary on a leg, at optimality, reflects the marginal value of that leg. In terms of right-hand side sensitivity, adding an additional seat to this leg would increase the overall network revenue by the value of that itinerary as long as the optimal basis of the integer solution does not change. This marginal value is analogous to a bid price that can be used for inventory control. Since both the policy and bid price controls from the CMIP are implementable, this model could be utilized for either reservation system, as well as a reservation system that utilizes both solutions for pricing and capacity controls.

*4.2. Comparison to Other Network RM Methods*

Two common models currently being used in the industry include a stochastic network flow formulation and the EMSR-b model, discussed in the literature review. The network flow formulation represents one approach commonly used and is a popular O&D RM strategy. The network formulation, as seen in Appendix A of Jacobs, Smith, and Johnson (2008), represents a stochastic passenger flow model, solved using a Lagrangian relaxation approach with a sub-gradient algorithm. The network flow formulation solves for the bid prices associated with each leg, and calculates the protection limit for each fare class on each leg using Littlewood's Rule.

The EMSR-b model represents a leg-based control strategy which estimates the bid prices using protection limits based on Littlewood's Rule. To account for the connecting traffic between O&D pairs, the revenue of the connecting fare was prorated and allocated to each leg in the O&D. The industry uses various versions of the EMSR-b and network formulation, making these industry models a good technique to compare against.

The EMSR-b model and the network formulation were solved for multiple passenger demand scenarios. The results show, using simulations of nine combinations of $\lambda$ and $T$, that the CMIP performs better than both models. The CMIP outperforms the EMSR-b by 11.82% in mean revenue over all of the nine cases given in Table 6. A standard z-test on the difference of each set of means for the results shown in Table 6 illustrated that the differences in the expected revenue of all nine combinations are statistically significant, with p-values less than 0.001.

In some situations the CMIP simulation resulted in larger confidence intervals, but this is due to the highly segmented nature of fares. The difference between a sale and no sale is at least $300 (in the case of the lowest fare for legs AB and BC), creating a large gap between revenues among simulations, yielding large standard deviations in relation to overall revenue. It is important to note, however, that as a larger amount of demand enters the network, the width of the confidence intervals for the CMIP reduces drastically; this is the exact opposite of the EMSR-b simulation, where the confidence intervals become wider as more demand enters the network.

For the network formulation, the gains were slightly less since the network formulation tends to perform better than the EMSR-b. The expected revenue showed an average increase of 9.60% over the nine cases presented in Table 7. Similar to the statistical tests of the EMSR-b, these nine cases show that the expected revenues are significantly different between the CMIP and network formulation, with p-values less than 0.001. All three model simulations were run together, yet we chose to display the ESMR-b and network formulation results separately for easier comparison. Similar to the EMSR-b, the network formulation uses a segmented demand model for predicting the expected number of passengers. This causes the network formulation to open up the lower fare classes earlier than the CMIP does. Since the CMIP keeps the lower fare classes closed for a longer period of time, a higher overall revenue is earned.

The results show a significant performance difference between the four tested network RM models. Figure 2 illustrates the gains by each of the models as one increases the total number of passengers introduced into the system over the entire time horizon. At the lowest level of passenger

| λ | T | CMIP<br>95% Confidence Interval | EMSR-b<br>95% Confidence Interval | % Increase<br>Over Mean |
|---|---|---|---|---|
| | 1 | (\$0, \$1,414) | (\$0, \$903) | 10.8 |
| 1 | 5 | (\$228, \$4,468) | (\$1,175, \$3,191) | 7.0 |
| | 10 | (\$1,891, \$7,213) | (\$2,692, \$5,430) | 10.8 |
| | 1 | (\$261, \$4,465) | (\$1,175, \$3,191) | 7.7 |
| 5 | 5 | (\$6,959, \$10,071) | (\$3,915, \$9,663) | 20.3 |
| | 10 | (\$10,229, \$12,199) | (\$7,834, \$11,488) | 13.8 |
| | 1 | (\$1,891, \$7,221) | (\$2,692, \$5,430) | 10.9 |
| 10 | 5 | (\$10,186, \$12,220) | (\$5,164, \$8,818) | 13.8 |
| | 10 | (\$13,174, \$13,772) | (\$10,069, \$13,799) | 11.4 |

Table 6: Expected revenue confidence intervals by bid pricing controls simulation (CMIP vs. EMSR-b)

| λ | T | CMIP<br>95% Confidence Interval | Network Formulation<br>95% Confidence Interval | % Increase<br>Over Mean |
|---|---|---|---|---|
| | 1 | (\$0, \$1,414) | (\$0, \$903) | 10.8 |
| 1 | 5 | (\$228, \$4,468) | (\$1,169, \$3,203) | 6.9 |
| | 10 | (\$1,891, \$7,213) | (\$2,659, \$5,521) | 10.2 |
| | 1 | (\$261, \$4,465) | (\$1,169, \$3,203) | 7.5 |
| 5 | 5 | (\$6,959, \$10,071) | (\$3,949, \$9,971) | 19.3 |
| | 10 | (\$10,229, \$12,199) | (\$8,771, \$12,517) | 5.1 |
| | 1 | (\$1,891, \$7,221) | (\$2,659, \$5,521) | 10.2 |
| 10 | 5 | (\$10,186, \$12,220) | (\$8,791, \$12,537) | 5.0 |
| | 10 | (\$13,174, \$13,772) | (\$10,069, \$13,799) | 11.4 |

Table 7: Expected revenue confidence intervals by bid pricing controls simulation (CMIP vs. Network Formulation)

demand, all four models behave similarly: they sell to any passenger that shows up. At the highest level of passenger demand, all the models again behave similarly: only sell to the highest paying passengers. However, as one moves from zero demand to a higher demand, the models begin to deviate from one another. The two dominating curves, the CMIP and CDLP, produces higher revenues compared to the EMSR-b and the network formulation. In fact, as seen in Figure 2, the CMIP and CDLP perform quite similarly.

## 5. Additional Examples

### *5.1. Small Network Instance*

In addition to running the model on the three leg network seen above, we also tested it on another network given in Liu and van Ryzin (2008). This network, depicted in Figure 3, is a small 22 product network, consisting of 7 legs.

The network contains a direct flight from A to B, with competition from A to B through the hub, H. There are two flights from each of the direct legs between A, H, B, and C: an early flight and a later flight. The network instance data, including all of the MNL choice parameters, can be found in Tables 18 and 19 in the Appendix for the reader's convenience. The small network instance was ran for 1,000 time periods, with $\lambda_t$ equal to 0.91 for each time period. This would represent a total of 910 customers introduced into the network.

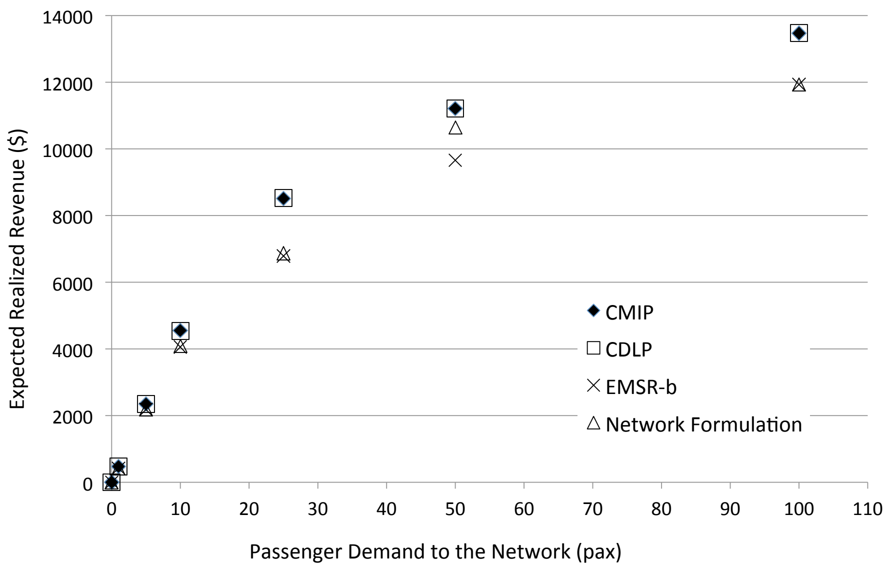


Figure 2: Passenger demand to the network versus expected revenue under different RM methods

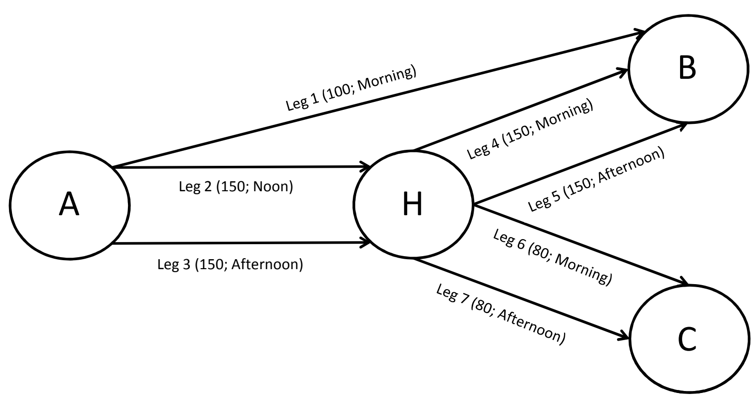


Figure 3: Small network instance - Adapted from Liu and van Ryzin (2008)

We ran our model, and compared the results to those given in Bront et al. (2009) for the example in Liu and van Ryzin (2008), and saw that our model performed similarly to the CDLP. Across the five tests, each for a different fraction of total network capacity, our model performed 9.0% better, on average, than the CDLP, while still maintaining similar levels of network Load Factor (LF). The LF, defined as the average across all legs of the ratio of seats taken to total capacity, represents how many seats, on average, are consumed across the entire network. In addition to comparing it to the CDLP, we also compared it to the model solved using the independent demand assumption (referred to as the INDEP model) found in Bront et al. (2009), in which a deterministic linear program is solved with demand values generated under the assumption that all of the products are simultaneously open.

The first column of Table 8 indicates the percentage of the base capacity used for both the model solution as well as the simulation. This value represents an increase or decrease in the amount of capacity available, while maintaining the demand over the time horizon. The table includes the expected revenues obtained by the CDLP and CMIP solutions, as well as the percent increase they offer over the INDEP model. The table entries for CDLP and INDEP come from the simulation results reported in Bront et al. (2009). As the capacity in the network increases, the network load factor should decrease, as the demand introduced into the network does not increase, although the available space does. As seen in Table 8, the load factors decrease as the amount of capacity increases for all of the models, as expected.

| Percent of | CMIP | | | CDLP | | | INDEP | |
|---|---|---|---|---|---|---|---|---|
| Base Cap. | Revenue | Inc. (%) | LF (%) | Revenue | Inc. (%) | LF (%) | Revenue | LF (%) |
| 60 | $224,114 | 30.0 | 98.5 | $207,890 | 20.6 | 91.3 | $172,362 | 97.7 |
| 80 | $278,241 | 36.0 | 92.1 | $261,264 | 27.7 | 85.6 | $204,572 | 94.6 |
| 100 | $297,752 | 31.7 | 83.6 | $277,738 | 22.9 | 80.8 | $226,002 | 87.7 |
| 120 | $315,832 | 29.5 | 77.0 | $282,842 | 16.0 | 71.6 | $243,930 | 82.5 |
| 140 | $318,153 | 22.8 | 70.2 | $285,417 | 10.2 | 62.0 | $259,039 | 77.0 |

Table 8: Expected revenues and percent increase over INDEP (Small Network instance)

*5.2. Large Network Instance*

We finally use a large network instance with realistic aspects to further test the performance of the CMIP. The network structure, as well as the revenue values associated with each itinerary can be found in Jacobs et al. (2008). This network contains 48 legs, each with an initial capacity of 200 seats, joining 10 cities (see Figure 4). Each leg represents a single flight. There are no parallel flights available in this network. There are 178 itineraries, with three fare classes each, denoted as Y, M, and Q, for the given itinerary. There are a total of 90 O&D markets, with markets containing either one, two, or three possible paths between O&D. For a single time period, the CMIP model has a total of 3,598 variables. Note that for this example, the CDLP would result in a total of $2^{534} - 1$ (or, about 5.6E160) variables. The total number of arrivals to the system was set to 9,750 for a single time period. A single time period was used to see a representation of a solution for an entire booking horizon. The MNL values associated with each choice were arbitrarily generated, as well as the individual arrival probabilities per market. These values can be found in three tables (Tables 20, 21, and 22) given in the Appendix, and are separated by how many competing itineraries existed between O&D, for easier classification. The preference vectors in Table 20 represent the utility of the three fare classes, Y, M, and Q. The preference vectors in Table 21 represent the utility of the three fare classes for each of the two itineraries available for that market. The first three utility values refer to the Y, M, and Q fare classes of the first itinerary, while the last three utility values refer to the Y, M, and Q fare classes of the second itinerary. Similarly, Table 22 displays the preference vectors for the three itineraries with respect to the Y, M, and Q fare classes.

As indicated by the results given in Table 9, the time required to solve this model is quite reasonable. As the capacity becomes more constraining, the model needs more time to find an optimal solution. However, it can still be solved in a reasonable amount of time. Since the CDLP is expected to have a total of 5.6E120 variables, we did not program the decomposition approach presented in Bront et al. (2009), as it would have proved to be computationally prohibitive.

In addition to the revenue and network LF, the available seat mile (ASM) and revenue per available seat mile (RASM) are also reported. The ASM is calculated by the number of seats available on a leg, multiplied by the distance traveled by the flight on that leg, then summed for all flight legs. The RASM is the total expected revenue divided by the ASM. This value is often used in the industry, and although the values seen in the table come from a fabricated instance, the RASMs are in-line with what is seen in industry practice.

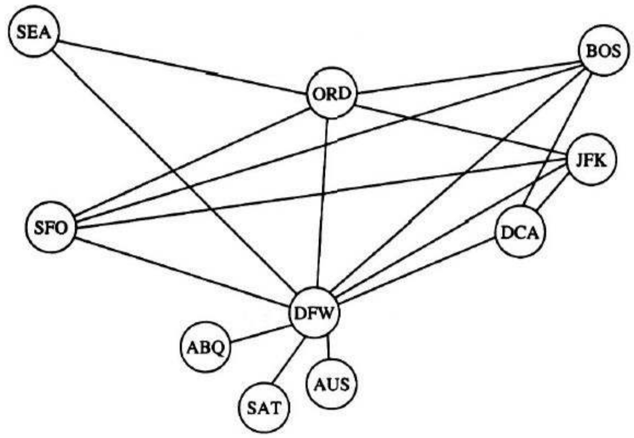


Figure 4: Large network instance, adapted from Jacobs et al. (2008)

| Percent of Base Capacity | CMIP | | | | |
|---|---|---|---|---|---|
| | Revenue | LF (%) | ASM | RASM | Elapsed Time (sec.) |
| 60 | $1,018,043 | 82.2 | 6,380,160 | $0.16 | 142.87 |
| 80 | $1,146,655 | 78.5 | 8,506,880 | $0.13 | 66.04 |
| 100 | $1,231,055 | 70.6 | 10,633,600 | $0.12 | 3.07 |
| 120 | $1,283,682 | 62.6 | 12,760,320 | $0.10 | 3.65 |
| 140 | $1,322,589 | 55.2 | 14,887,040 | $0.09 | 1.59 |

Table 9: Expected revenues for the large network instance

### *5.3. Computational Complexity*

Many of the existing RM approaches, such as the dynamic program proposed in Talluri and van Ryzin (2004) generate an exponential number of solutions by explicitly combining market policy strategies together across the network. This is problematic for industry use, as networks and fare class buckets have grown to create many itineraries. One example of this complexity issue is apparent in the CDLP's solution set, $S$. Since set $S$ contains all the fare class and O&D controls for the network, $S$ is dependent on the network definition, including prices and control strategies. For example, a small three node, four itinerary, two fare class network yields 255 decision variables for the CDLP formulation. However, if we increase the complexity of the network to four nodes, seven itineraries, and keep the two fare classes, the model has a total of 16,383 decision variables. This number continues to grow exponentially when any parameter of the network is increased, which can be seen in Table 10. This table illustrates the growth in complexity of the CDLP versus that of the CMIP for the previous examples. As one can see, the CMIP doesn't increase in size as fast, which would allow for consideration of being tractable for industry use.

| Network Instance | Number of Products | Number of Variables | |
|---|---|---|---|
| | | CMIP | CDLP |
| Three Leg Example | 8 | 24 | 255 |
| Small Network Example | 22 | 116 | $4.2 \times 10^{6}$ |
| Large Network Example | 534 | 3598 | $5.6 \times 10^{160}$ |

Table 10: Variable complexity of CMIP and CDLP for a single time period

## 6. Dynamic Pricing

In traditional choice-based revenue management the attributes, such as price, used to generate the utilities are selected ahead of time and assumed to be static. Although the customer only sees one price for a ticket, the airline has a range of prices for each type of ticket, thus traditional choice-based revenue management models fall short of modeling the true environment of customer purchasing behavior by excluding these price ranges and assuming a static price is available. In this section, we introduce an extension to the work in the previous sections that models the environment in which customers are purchasing tickets by incorporating the structure of an MNL directly into the mathematical model. With this approach we are able to determine which tickets to offer a market as well as an optimal price for the preferences of that market. The added flexibility allows the model, the Price-dynamic Choice-based Mixed Integer Program (PCMIP), to yield higher revenue than the static counterpart, and eliminates the two-step process of offering tickets followed by pricing them. This creates a feasible model for industry use since it capitalizes on the complexity reduction of the CMIP while incorporating complex pricing decisions.

### *6.1. Model Formulation*

To introduce pricing into the Choice-based Mixed Integer Program (CMIP) formulation we must go back to the formulation of the MNL itself. MNL models determine the probability of selecting a particular option out of a group of alternative options by a ratio of predicted values of a regression model of the form

$$P(y_n = j|x_n) = \frac{e^{x'_n \beta_j}}{\sum_l e^{x'_n \beta_j}}. \tag{7}$$

In this equation $y_n$ is a dependent variable referencing person $n$, $j$ represents an option indexed 1 to $J$, and $x_n$ is a vector of preferences unique to person $n$. $P(y_n = j|x_n)$ is the probability, given the preferences of $n$, that person $n$ selects option $j$. The $\beta$s for this MNL model are determined from a regression equation containing the attributes associated with the preferences of the sample population. The previous chapter simplified this equation by converting $e^{x'_n \beta_j}$ into utilities, $u_k$, provided in the data sets. In revenue management this regression model is characterized by many attributes including price, time of day, path from origin to destination, number of connections, and others. Different segments of an origin-destination market, such as business and leisure, would potentially have different sensitivities to these preferences but could easily be considered in a single MNL model.

Under this context, we can incorporate pricing into the CMIP by characterizing a regression equation as

$$\alpha_{kt} = Y_{kt}\beta_k + \boldsymbol{\omega_k}\boldsymbol{\delta_k}, \tag{8}$$

where $\alpha_{kt}$ is the response for ticket $k$ in time unit $t$, $Y_{kt}$ is the price for ticket $k$ in time unit $t$, $\beta_k$ is the sensitivity to price for ticket $k$ in the customer population. $\boldsymbol{\omega_k}$ is the vector associated with the attributes of ticket $k$ and $\boldsymbol{\delta_k}$ is the sensitivity (fitted regression values) mapping for ticket $k$. In a standard MNL format, we then know that the utilities of a particular ticket, $k$, would be

$$u_{kt} = e^{\alpha_{kt}}, \tag{9}$$

where $u_{kt}$ would have the same interpretation as the utilities seen in the previous sections.

Since the utilities have been determined directly from a regression equation, we can evaluate the probability a purchase is made. Given an origin-destination market $j$ we can denote a super set of all possible ticket offerings as $I_j$, with $i \in I_j$ representing a set of all available tickets for this origin-destination. $K_{ji}$ then, denotes the set of tickets available to purchase, where $K_{ji} \subseteq I_j$.

We can then calculate the probability a customer from market $j$ will purchase a ticket in offer set $K_{ji} \subseteq I_j$ in time period $t$, $P_{jit}$, as

$$P_{jit} = \frac{\sum_{k \in K_{ji}} u_{kt}}{\sum_{k \in K_{ji}} u_{kt} + v_j}, \tag{10}$$

where $v_j$ is the no-purchase utility for origin-destination market $j$. Likewise, we can then calculate the probability a customer purchased ticket $k \in K_{ji}$, given that a purchase was made, as

$$Q_{kt|ji} = \frac{u_{kt}}{\sum_{k \in K_{ji}} u_{kt} + v_j} P_{jit}^{-1}. \tag{11}$$

With these values we can now extend the model introduced previously, the Choice-based Mixed Integer Program, to develop the *Price-dynamic Choice-based Mixed Integer Program* (PCMIP):

$$\text{Maximize} \quad \sum_{t \in T} \gamma_t \sum_{j \in J} \lambda_j \sum_{i: K_{ji} \subseteq I_j} P_{jit} \sum_{k \in K_{ji}} Y_{kt} Q_{kt|ji} \tag{12}$$

Subject to:

$$\sum_{t \in T} \gamma_t \sum_{j \in J} \lambda_j \sum_{i: K_{ji} \subseteq I_j} P_{jit} \sum_{k \in K_{ji}} Q_{kt|ji} A_{kl} \leq c_l, \qquad \text{for all } l \in L, \tag{13}$$

$$\sum_{i: K_{ji} \subseteq I_j} X_{jit} \leq 1 \ , \qquad \text{for all } j \in J, t \in T, \tag{14}$$

$$u_{kt} = e^{Y_{kt}\beta_k + \boldsymbol{\omega_k \delta_k}} \qquad \text{for all } j \in J, k \in K_{ji}, \tag{15}$$

$$P_{jit} = \frac{\sum_{k \in K_{ji}} u_{kt}}{\sum_{k \in K_{ji}} u_{kt} + v_j} \qquad \text{for all } j \in J, i : K_{ji} \subseteq I_j, \tag{16}$$

$$Q_{kt|ji} = \frac{u_{kt}}{\sum_{k \in K_{ji}} u_{kt} + v_j} P_{jit}^{-1} \qquad \text{for all } j \in J, i : K_{ji} \subseteq I_j, k \in K_{ji}, \tag{17}$$

$$\text{LB}_k \leq Y_{kt} \leq \text{UB}_k \qquad \text{for all } j \in J, k \in K_{ji}, t \in T, \tag{18}$$

$$X_{jit} \in \{0, 1\}, \qquad \text{for all } j \in J, i : K_{ji} \subseteq I_j, t \in T. \tag{19}$$

There are two decision variables for the PCMIP. $X_{jit}$ represents the binary decision to select offer set $K_{ji}$ for origin-destination market $j$ in time unit $t$. $Y_{kt}$ represents the price for ticket $k$ in time unit $t$, where $k \in K_{ji}$. The previous decision variable in the CMIP, $Z_{jit}$, is no longer needed as we are setting prices directly. Previously, $Z_{jit}$ was used to generate complementary slackness for the capacity constraint in an effort to utilize a bid price. Since $Y_{kt}$ represents the price to sell a ticket for, the model can manipulate $Y_{kt}$ directly to reach capacity as oppose to setting $Z_{jit}$. With these decision variables the PCMIP will select an offer set as well as assign the prices to each ticket available in that offer set.

The objective function (12) is similar to that of the CMIP, except $Y_{kt}$ replaces the revenue parameter, $R_k$, since price is now a decision. Constraint set (13) prevents the expected demand from exceeding the capacity on a leg, $c_l$. Constraint set (14) forces the model to select only one ticket offer set of the super set $I_j$. Constraint set (15) generates the utilities for ticket $k$ while constraint sets (16) and (17) define $P_{jit}$ and $Q_{kt|ji}$ based on equations 10 and 11, respectively. Constraint set (18) establishes lower and upper bounds on the pricing decision for each ticket, while constraint set (19) defines the binary restriction for $X_{jit}$.

The PCMIP is unique when compared to other choice-based revenue management models previously developed. First, it determines the ticket offer set at any point in time while simultaneously setting ticket prices. In previous revenue management models prices are assumed to be static and subsequent algorithms adjust prices as necessary based on capacity, given a solution from the revenue management model. The PCMIP eliminates the two-step optimization process by combining the decisions into a single, concise model. Secondly, the PCMIP reduces the number of offer sets to consider, when compared to other choice-based revenue management models, by separating origin-destination markets that have no interaction with other origin-destination markets, much like the CMIP does. Although airline networks are large due to a highly connected network, the PCMIP separates the problem into independently solvable offer sets making it manageable for the airline industry to implement.

### *6.2. Problem Instances and Results*

Two problem instances were constructed to illustrate the effectiveness of optimizing price along with availability. These problem instances were solved with both the PCMIP and CMIP, representing the cases of dynamic-pricing and static-pricing, respectively. The network for the problem instances, seen in Figure 5, consisted of a single O&D pair with three flight options representing early, mid-day, and late flights, with their capacities presented in parentheses. This network was adapted from Bront et al. (2009) and provided a simple network to analyze the impact pricing could have on revenue.

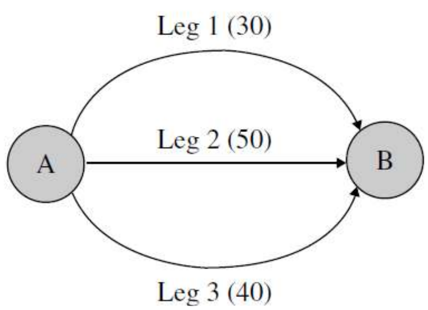


Figure 5: Parallel problem instance (Bront et al., 2009)

Table 11: Products and fare classes for parallel problem instance (Bront et al., 2009)

| Product | Leg | Class | Fare |
|---|---|---|---|
| 1 | 1 | Low | \$400 |
| 2 | 1 | High | \$800 |
| 3 | 2 | Low | \$500 |
| 4 | 2 | High | \$1000 |
| 5 | 3 | Low | \$300 |
| 6 | 3 | High | \$600 |

Two regression equations were developed for this network to highlight the advantages of the PCMIP. The first regression equation only accounted for sensitivity to price, setting $\beta_1 = -0.002$, yielding a regression equation of the form $\alpha_{kt} = -0.002Y_{kt}$. The second regression equation took a more realistic approach, incorporating both price and time of day (ToD) preference, yielding a regression equation of the form $\alpha_{kt} = -0.002Y_{kt} + \text{AB}_1 + 0.5\text{AB}_2$, where $\text{AB}_i$ equals 1 if the leg $i$ is selected and 0 otherwise. The two problem instances for the parallel network were solved by both models (PCMIP and CMIP) for five time periods ($T = 5$), with an increasing arrival rate for subsequent time periods, where $\gamma = (6, 12, 24, 48, 96)$, and their solutions were simulated. The bounds on the prices within the PCMIP were generated by finding the midpoint between the fare classes in Table 11. For instance, Leg 1's High fare class has a lower bound of \$600 and no upper bound, whereas Leg 1's Low fare class has a lower bound of \$0 and an upper bound of \$599.99.

The first problem instance, with price being the only factor, converged to the solution shown in Table 12 for each of the models. The prices of each ticket on each leg are given, as these represent both the availability of a ticket and the price at which it's sold. The PCMIP is able to select its own pricing policy, whereas the CMIP must select the prices displayed in Table 11. A fare of \$0 represent a closed option, and thus no purchase can be made on that leg during that time period.

The simulation of the solutions from Tables 12 and 13 resulted in the PCMIP consistently outperforming the CMIP, as displayed in Table 14 and Figure 6. The PCMIP yielded a 6.4% increase in revenue overall, while maintaining a similar load factor (PCMIP 78% versus CMIP 82%) and traffic. Since the PCMIP is free to select price, and thus less constrained than the CMIP, the higher

Table 12: PCMIP Solution - No ToD pref.

| Time Period | Leg 1 | Leg 2 | Leg 3 |
|---|---|---|---|
| 1 | $945.24 | $945.24 | $800.00 |
| 2 | $945.24 | $945.24 | $800.00 |
| 3 | $945.24 | $945.24 | $800.00 |
| 4 | $945.24 | $945.24 | $800.00 |
| 5 | $945.24 | $945.24 | $800.00 |

Table 13: CMIP Solution - No ToD pref.

| Time Period | Leg 1 | Leg 2 | Leg 3 |
|---|---|---|---|
| 1 | $0.00 | $1000.00 | $600.00 |
| 2 | $0.00 | $1000.00 | $0.00 |
| 3 | $800.00 | $500.00 | $0.00 |
| 4 | $800.00 | $1000.00 | $600.00 |
| 5 | $800.00 | $1000.00 | $600.00 |

revenue gains were expected, and the PCMIP dominated the CMIP in cumulative revenue over the five time periods.

Table 14: Revenue and traffic per time period - No ToD pref.

| Time Period | PCMIP | | CMIP | |
|---|---|---|---|---|
| | Rev. | Traffic | Rev. | Traffic |
| 1 | $2,627 | 2.96 | $1,997 | 2.76 |
| 2 | $4,896 | 5.52 | $2,346 | 2.35 |
| 3 | $10,236 | 11.53 | $7,459 | 12.30 |
| 4 | $20,883 | 23.55 | $19,623 | 26.24 |
| 5 | $41,375 | 46.54 | $38,902 | 51.05 |
| Total | $80,016 | 90.10 | $70,327 | 94.70 |

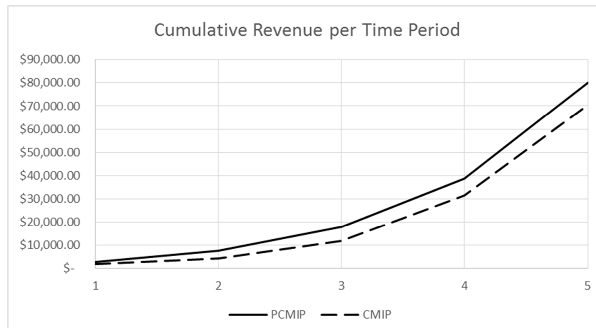


Figure 6: Cumulative revenue per time period - No ToD pref.

The second problem instance for this network, where price and time of day was considered, yielded a considerably more dynamic solution, as seen in Tables 15 and 16. Now that price isn't the only factor to consider, the PCMIP must balance the sensitivity to price along with the time of day preferences, resulting in a more complex pricing structure. These results were simulated, and the PCMIP outperformed the CMIP by 21%, while managing to achieve a higher load factor of 94%, versus that of the CMIP's 86% (see Table 17 and Figure 7). It is worth noting that the "bump" in the cumulative revenue (Figure 7) at time period 3 stems from the difference in solutions. The PCMIP favors shutting down Legs 1 and 2 early in the time horizon in anticipation of future demand, whereas the CMIP only closes Leg 1. Ultimately, though, the PCMIP's solution yields higher revenue and dominates the CMIP solution in cumulative revenue for all time periods, similar to when price was the only consideration.

Although these problem instances are small in size, they highlight the importance price has on the decision of availability. Setting only availability, much in the way traditional revenue management models approach the problem, fails to account for the relationship between passenger preference and pricing. The PCMIP provides an alternative solution methodology to a two-step revenue management/pricing system, incorporating choice-based demand and passenger preference in a manageable model.

## 7. Conclusions and Future Work

The proposed CMIP and PCMIP formulations use an MNL model to explicitly model the impact of network-wide offerings and passenger preference on the probability of purchase to better reflect customer behavior. Using problem instances of varying size, we have shown that CMIP

Table 15: PCMIP Solution - With ToD pref.

| Time Period | Leg 1 | Leg 2 | Leg 3 |
|---|---|---|---|
| 1 | $1000.00 | $608.07 | $0.00 |
| 2 | $1000.00 | $588.62 | $800.00 |
| 3 | $0.00 | $700.00 | $0.00 |
| 4 | $0.00 | $666.42 | $800.00 |
| 5 | $1000.00 | $800.00 | $800.00 |

Table 16: CMIP Solution - With ToD pref.

| Time Period | Leg 1 | Leg 2 | Leg 3 |
|---|---|---|---|
| 1 | $800.00 | $500.00 | $0.00 |
| 2 | $0.00 | $500.00 | $600.00 |
| 3 | $0.00 | $1000.00 | $600.00 |
| 4 | $0.00 | $1000.00 | $0.00 |
| 5 | $800.00 | $500.00 | $600.00 |

Table 17: Revenue per time period - With ToD pref.

| Time Period | PCMIP | | CMIP | |
|---|---|---|---|---|
| | Rev. | Traffic | Rev. | Traffic |
| 1 | $2,782 | 3.36 | $2,601 | 3.83 |
| 2 | $6,491 | 8.08 | $3,884 | 6.97 |
| 3 | $6,816 | 9.74 | $8,809 | 12.85 |
| 4 | $22,001 | 29.70 | $10,059 | 10.06 |
| 5 | $53,266 | 60.65 | $43,895 | 68.09 |
| Total | $91,356 | 111.52 | $69,248 | 101.79 |

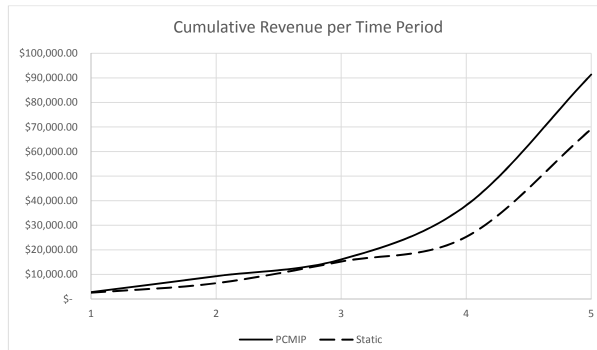


Figure 7: Cumulative revenue per time period - With ToD pref.

outperforms both the EMSR-b and a basic network formulation, as well as the benchmark CDLP, which utilizes the same MNL model to develop its probability of purchase but yields a solution that is somewhat difficult to decipher and implement. The PCMIP introduces more flexibility to the problem of revenue management by allowing the model to dynamically set prices while managing ticket availability. The PCMIP consistently outperformed the CMIP, as expected with a less constrained, more flexible model.

From a pragmatic perspective, the CMIP and PCMIP approaches build on the advantages of previous models by addressing passenger choice in a computationally more efficient manner. Future work includes full scale tests of these approaches and calibration of the passenger choice model needed to drive the optimization. Other areas include expanding the model to handle bookings of multiple passengers at once, and time dependent demand utilities and pricing sensitivities.

With the results from the examples and the possibility of industry specific extensions, the CMIP and PCMIP look promising for future research. These models could improve on the methods currently being used by leading airline companies today as well as be the groundwork for further development in the area of RM. Utilizing choice modeling and mathematical programming, the choice-based mixed integer program and price-dynamic choice-based mixed integer program successfully optimizes the network RM problem for the airline industry. Further development of this model could assist in changing the way the industry solves their network RM problems.

## References

D. Adelman. Dynamic bid prices in revenue management. *Operations Research*, 55(4):647–661, 2007.

G. Aydin and J. K. Ryan. Product line selection and pricing under the multinomial logit choice model. Working paper, Purdue University, Indianapolis, 2000.

P. Belobaba. *Optimal versus heuristic methods for nested seat allocation.* AGIFORS Yield Management Study Group, 1992.

P. P. Belobaba. Application of a probabilistic decision model to airline seat inventory control. *Operations Research*, 37(2):183–197, 1989.

J. J. M. Bront, I. Mendez-Diaz, and G. Vulcano. A column generation algorithm for choice-based network revenue management. *Operations Research*, 57(3):769–784, 2009.

J. M. Chaneton and G. Vulcano. Computing bid-prices for revenue management under customer choice behavior. *Manufacturing and Service Operations Management*, 13(4):452–470, 2011.

L. Chen and T. H. de Mello. Mathematical programming models for revenue management under customer choice. *European Journal of Operational Research*, 203:294–305, 2010.

L. Dong, P. Kouvelis, and Z. Tian. Dynamic pricing and inventory control of substitute products. *Manufacturing & Service Operations Management*, 11(2):317–339, 2009.

V. F. Farias and B. van Roy. An approximate dynamic programming approach to network revenue management. Working paper (available at http://www.standford.edu/ bvr/psfiles/adp-rm.pdf), Stanford University, 2007.

G. Gallego, G. Iyengar, R. Phillips, and A. Dubey. Managing flexible products on a network. CORC technical report tr-2004-01, Department of Industrial Engineering and Operations Research, Columbia University, 2004. http://www.corc.ieor.columbia.edu/reports/techreports/.

G. Gallego, R. Ratliff, and S. Shebalov. A general attraction model and an efficient formulation for the network revenue management problem. Working Paper, 2011.

K. Huang and Y.-T. Liang. A dynamic programming algorithm based on expected revenue approximation for the network revenue management problem. *Transportation Research Part E*, 47:333–341, 2011.

T. L. Jacobs, B. C. Smith, and E. L. Johnson. Incorporating network flow effects into the airline fleet assignment process. *Transportation Science*, 42(4):514–519, 2008.

T. L. Jacobs, R. Ratliff, and B. C. Smith. Understanding the relationship between price, revenue management controls and scheduled capacity - a price balance statistic for optimizing pricing strategies. *Journal of Revenue and Pricing Management*, 9:356–373, 2010.

S. Kunnumkal. Randomization approaches for network revenue management with customer choice behavior. www.isb.edu/faculty/Sumit_Kunnumkal/papers/choice_random.pdf, 2011.

S. Kunnumkal and H. Topaloglu. A refined deterministic linear program for the network revenue management problem with customer choice behavior. *Wiley InterScience*, 2008. www.interscience.wiley.com.

S. Kunnumkal and H. Topaloglu. A new dynamic programming decomposition method for the network revenue management problem with customer choice behavior. *Prodcution and Operations Management*, 19(5):575–590, 2010.

K. Littlewood. Forecasting and control of passenger bookings. *Journal of Revenue and Pricing Management*, 4(2):111–123, 2005.

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

J. Meissner and A. Strauss. Network revenue management with inventory-sensitive bid prices and customer choice. *European Journal of Operational Research*, 216:459–468, 2012.

J. Meissner and A. K. Strauss. Pricing structure optimization in mixed restricted/unrestricted fare environments. *Journal of Revenue and Pricing Management*, 9:399–418, 2010.

J. Meissner and A. K. Strauss. Choice-based network revenue management under weak market segmentation. Technical report, Kuehne Logistics University and Lancaster University Management School, 2011.

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

K. Talluri. A randomized concave programming method for choice network revenue management. urlwww.econ.upf.edu/docs/papers/downloads/1215.pdf, 2011.

K. Talluri and G. van Ryzin. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1):15–33, 2004.

H. Topaloglu. Using lagrangian relaxation to computer capacity-dependent bid prices in network revenue management. *Operations Research*, 57(3):637–649, 2009.

G. van Ryzin and G. Vulcano. Computing virtual nesting controls for network revenue management under customer choice behavior. *Manufacturing and Service Operations Research*, 10(3):448–467, 2008.

L. R. Weatherford and R. M. Ratliff. Review of revenue management methods with dependent demands. *Journal of Pricing and Revenue Management*, 9:326–340, 2010.

D. Zhang. An improved dynamic programming decomposition approach for network revenue management. *Manufacturing and Service Operations Management*, 13(1):35–52, 2011.

D. Zhang and W. L. Cooper. Revenue management for parallel flights with customer-choice behavior. *Operations Research*, 53(3):415–573, 2005.

D. Zhang and W. L. Cooper. Pricing substitutable flights in airline revenue management. *European Journal of Operational Research*, 197:848–861, 2009.

D. Zhang and Z. Lu. Assessing the value of dynamic pricing in network revenue management. *INFORMS Journal on Computing*, 25(1):102–115, 2013.

## Appendix

| Product | Legs | Class | Fare | Product | Legs | Class | Fare |
|---|---|---|---|---|---|---|---|
| 1 | 1 | H | 1000 | 12 | 1 | L | 500 |
| 2 | 2 | H | 400 | 13 | 2 | L | 200 |
| 3 | 3 | H | 400 | 14 | 3 | L | 200 |
| 4 | 4 | H | 300 | 15 | 4 | L | 150 |
| 5 | 5 | H | 300 | 16 | 5 | L | 150 |
| 6 | 6 | H | 500 | 17 | 6 | L | 250 |
| 7 | 7 | H | 500 | 18 | 7 | L | 250 |
| 8 | {2,4} | H | 600 | 19 | {2,4} | L | 300 |
| 9 | {3,5} | H | 600 | 20 | {3,5} | L | 300 |
| 10 | {2,6} | H | 700 | 21 | {2,6} | L | 350 |
| 11 | {3,7} | H | 700 | 22 | {3,7} | L | 350 |

Table 18: Product definitions for the Small Network instance - Adapted from Liu and van Ryzin (2008)

| Segment | O-D | Consideration Set | Preference Vector | Utility of No Purchase | $\lambda_t$ |
|---|---|---|---|---|---|
| 1 | A-B | {1,8,9,12,19,20} | (10,8,8,6,4,4) | 1 | 0.08 |
| 2 | A-B | {1,8,9,12,19,20} | (1,2,2,8,10,10) | 5 | 0.20 |
| 3 | A-H | {2,3,13,14} | (10,10,5,5) | 1 | 0.05 |
| 4 | A-H | {2,3,13,14} | (2,2,10,10) | 5 | 0.20 |
| 5 | H-B | {4,5,15,16} | (10,10,5,5) | 1 | 0.10 |
| 6 | H-B | {4,5,15,16} | (2,2,10,8) | 5 | 0.15 |
| 7 | H-C | {6,7,17,18} | (10,8,5,5) | 1 | 0.02 |
| 8 | H-C | {6,7,17,18} | (2,2,10,8) | 5 | 0.05 |
| 9 | A-C | {10,11,21,22} | (10,8,5,5) | 1 | 0.02 |
| 10 | A-C | {10,11,21,22} | (2,2,10,10) | 5 | 0.04 |

Table 19: Segment definitions for the Small Network instance - Adapted from Liu and van Ryzin (2008)

| Single Itinerary Markets | | | | |
|---|---|---|---|---|
| Market | Itinerary | Arrival Rate | Preference Vector | No Purchase Utility |
| SATBOS | SATDFW1DFWBOS1 | 0.002 | (1, 2, 3) | 4 |
| SATSEA | SATDFW1DFWSEA1 | 0.005 | (2, 2, 4) | 5 |
| SEAABQ | SEADFW1DFWABQ1 | 0.002 | (3, 3, 3) | 8 |
| SEAAUS | SEADFW1DFWAUS1 | 0.002 | (4, 4, 5) | 10 |
| SEADCA | SEADFW1DFWDCA1 | 0.005 | (3, 5, 7) | 11 |
| SEAJFK | SEADFW1DFWJFK1 | 0.007 | (1, 2, 3) | 8 |
| SEASAT | SEADFW1DFWSAT1 | 0.002 | (2, 3, 8) | 10 |
| SEASFO | SEADFW1DFWSFO1 | 0.002 | (1, 1, 5) | 9 |
| SFOABQ | SFODFW1DFWABQ1 | 0.002 | (3, 4, 8) | 9 |
| SFOAUS | SFODFW1DFWAUS1 | 0.002 | (2, 2, 7) | 9 |
| SFODCA | SFODFW1DFWDCA1 | 0.002 | (1, 1, 4) | 11 |
| SFOORD | SFODFW1DFWORD1 | 0.005 | (1, 2, 3) | 5 |
| SFOSAT | SFODFW1DFWSAT1 | 0.005 | (1, 1, 2) | 4 |

Table 20: Consideration sets and utility values for the single itineraries in the large network example

| Double Itinerary Markets | | | | | | |
|---|---|---|---|---|---|---|
| Market | Itinerary | | Arrival Rate | | Preference Vector | No Purchase Utility |
| ABQAUS | ABQDFW1DFWAUS1 | ABQDFW2DFWAUS2 | 0.037 | 0.012 | (1, 2, 3, 4, 5, 6) | 7 |
| ABQBOS | ABQDFW1DFWBOS1 | ABQDFW2DFWBOS2 | 0.002 | 0.049 | (2, 3, 4, 5, 6, 7) | 8 |
| ABQDCA | ABQDFW1DFWDCA1 | ABQDFW2DFWDCA2 | 0.002 | 0.007 | (1, 3, 4, 5, 5, 6) | 9 |
| ABQDFW | ABQDFW1 | ABQDFW2 | 0.005 | 0.005 | (2, 2, 3, 3, 4, 5) | 7 |
| ABQJFK | ABQDFW1DFWJFK1 | ABQDFW2DFWJFK2 | 0.01 | 0.01 | (1, 1, 2, 3, 4, 7) | 9 |
| ABQORD | ABQDFW1DFWORD1 | ABQDFW2DFWORD2 | 0.012 | 0.002 | (2, 2, 3, 5, 10, 11) | 12 |
| ABQSAT | ABQDFW1DFWSAT1 | ABQDFW2DFWSAT2 | 0.002 | 0.002 | (1, 1, 2, 2, 3, 3) | 4 |
| ABQSEA | ABQDFW1DFWSEA1 | ABQDFW2DFWSEA2 | 0.005 | 0.007 | (1, 1, 1, 2, 4, 5) | 9 |
| ABQSFO | ABQDFW1DFWSFO1 | ABQDFW2DFWSFO2 | 0.002 | 0.005 | (1, 3, 3, 4, 5, 6) | 6 |
| AUSABQ | AUSDFW1DFWABQ1 | AUSDFW2DFWABQ2 | 0.007 | 0.002 | (1, 1, 2, 3, 4, 5) | 5 |
| AUSBOS | AUSDFW1DFWBOS1 | AUSDFW2DFWBOS2 | 0.002 | 0.002 | (1, 2, 3, 4, 5, 6) | 7 |
| AUSDCA | AUSDFW1DFWDCA1 | AUSDFW2DFWDCA2 | 0.002 | 0.005 | (2, 3, 4, 5, 6, 7) | 8 |
| AUSDFW | AUSDFW1 | AUSDFW2 | 0.005 | 0.01 | (1, 3, 4, 5, 5, 6) | 9 |
| AUSJFK | AUSDFW1DFWJFK1 | AUSDFW2DFWJFK2 | 0.005 | 0.005 | (2, 2, 3, 3, 4, 5) | 7 |
| AUSORD | AUSDFW1DFWORD1 | AUSDFW2DFWORD2 | 0.002 | 0.002 | (1, 1, 2, 3, 4, 7) | 9 |
| AUSSAT | AUSDFW1DFWSAT1 | AUSDFW2DFWSAT2 | 0.01 | 0.005 | (2, 2, 3, 5, 10, 11) | 12 |
| AUSSEA | AUSDFW1DFWSEA1 | AUSDFW2DFWSEA2 | 0.002 | 0.005 | (1, 1, 2, 2, 3, 3) | 4 |
| AUSSFO | AUSDFW1DFWSFO1 | AUSDFW2DFWSFO2 | 0.007 | 0.002 | (1, 1, 1, 2, 4, 5) | 9 |
| BOSABQ | BOSDFW1DFWABQ1 | BOSDFW2DFWABQ2 | 0.005 | 0.002 | (1, 3, 3, 4, 5, 6) | 6 |
| BOSAUS | BOSDFW1DFWAUS1 | BOSDFW2DFWAUS2 | 0.002 | 0.002 | (1, 1, 2, 3, 4, 5) | 5 |
| BOSDFW | BOSDFW1 | BOSDFW2 | 0.002 | 0.005 | (1, 2, 3, 4, 5, 6) | 6 |
| BOSJFK | BOSDFW1DFWJFK1 | BOSDFW2DFWJFK2 | 0.01 | 0.002 | (1, 3, 4, 5, 5, 6) | 9 |
| BOSSAT | BOSDFW1DFWSAT1 | BOSDFW2DFWSAT2 | 0.002 | 0.002 | (1, 2, 3, 4, 4, 5) | 5 |
| BOSSEA | BOSDFW1DFWSEA1 | BOSDFW2DFWSEA2 | 0.005 | 0.002 | (1, 1, 2, 2, 3, 4) | 4 |
| DCAABQ | DCADFW1DFWABQ1 | DCADFW2DFWABQ2 | 0.002 | 0.005 | (1, 3, 3, 4, 5, 6) | 7 |
| DCAAUS | DCADFW1DFWAUS1 | DCADFW2DFWAUS2 | 0.002 | 0.01 | (1, 1, 2, 2, 3, 3) | 5 |
| DCAORD | DCADFW1DFWORD1 | DCADFW2DFWORD2 | 0.002 | 0.002 | (1, 1, 4, 4, 6, 7) | 8 |
| DCASAT | DCADFW1DFWSAT1 | DCADFW2DFWSAT2 | 0.005 | 0.002 | (1, 2, 3, 4, 4, 5) | 6 |
| DCASEA | DCADFW1DFWSEA1 | DCADFW2DFWSEA2 | 0.002 | 0.002 | (1, 1, 3, 3, 4, 4) | 5 |
| DCASFO | DCADFW1DFWSFO1 | DCADFW2DFWSFO2 | 0.005 | 0.002 | (1, 1, 1, 2, 4, 5) | 7 |
| DFWABQ | DFWABQ1 | DFWABQ2 | 0.002 | 0.005 | (1, 2, 3, 3, 4, 5) | 8 |
| DFWAUS | DFWAUS1 | DFWAUS2 | 0.002 | 0.002 | (1, 4, 4, 5, 8, 9) | 10 |
| DFWBOS | DFWBOS1 | DFWBOS2 | 0.005 | 0.005 | (2, 2, 4, 5, 6, 6) | 9 |
| DFWJFK | DFWJFK1 | DFWJFK2 | 0.002 | 0.012 | (1, 1, 3, 4, 4, 9) | 9 |
| DFWORD | DFWORD1 | DFWORD2 | 0.005 | 0.002 | (1, 1, 1, 2, 3, 4) | 5 |
| DFWSAT | DFWSAT1 | DFWSAT2 | 0.002 | 0.007 | (1, 2, 2, 3, 3, 4) | 5 |
| DFWSEA | DFWSEA1 | DFWSEA2 | 0.005 | 0.007 | (1, 2, 2, 3, 6, 6) | 10 |
| DFWSFO | DFWSFO1 | DFWSFO2 | 0.01 | 0.005 | (1, 3, 3, 5, 6, 6) | 9 |
| JFKABQ | JFKDFW1DFWABQ1 | JFKDFW2DFWABQ2 | 0.002 | 0.005 | (1, 1, 3, 4, 5, 5) | 9 |
| JFKAUS | JFKDFW1DFWAUS1 | JFKDFW2DFWAUS2 | 0.005 | 0.01 | (1, 3, 4, 7, 8, 9) | 10 |
| JFKBOS | JFKDFW1DFWBOS1 | JFKDFW2DFWBOS2 | 0.002 | 0.005 | (2, 2, 5, 6, 6, 7) | 9 |
| JFKORD | JFKDFW1DFWORD1 | JFKDFW2DFWORD2 | 0.005 | 0.005 | (1, 2, 3, 3, 4, 5) | 7 |
| JFKSAT | JFKDFW1DFWSAT1 | JFKDFW2DFWSAT2 | 0.002 | 0.005 | (1, 2, 3, 4, 5, 5) | 6 |
| JFKSEA | JFKDFW1DFWSEA1 | JFKDFW2DFWSEA2 | 0.005 | 0.01 | (1, 2, 2, 3, 4, 5) | 5 |
| ORDABQ | ORDDFW1DFWABQ1 | ORDDFW2DFWABQ2 | 0.002 | 0.005 | (1, 2, 5, 5, 8, 9) | 10 |
| ORDAUS | ORDDFW1DFWAUS1 | ORDDFW2DFWAUS2 | 0.005 | 0.01 | (1, 1, 2, 2, 3, 3) | 4 |
| ORDBOS | ORDBOS1 | ORDDFW1DFWBOS1 | 0.01 | 0.012 | (1, 1, 5, 5, 8, 9) | 10 |
| ORDDCA | ORDDFW1DFWDCA1 | ORDDFW2DFWDCA2 | 0.002 | 0.005 | (2, 2, 4, 4, 6, 6) | 8 |
| ORDDFW | ORDDFW1 | ORDDFW2 | 0.002 | 0.01 | (1, 1, 3, 4, 5, 7) | 8 |
| ORDJFK | ORDDFW1DFWJFK1 | ORDDFW2DFWJFK2 | 0.007 | 0.005 | (1, 4, 5, 5, 6, 6) | 10 |
| ORDSAT | ORDDFW1DFWSAT1 | ORDDFW2DFWSAT2 | 0.002 | 0.005 | (1, 4, 4, 6, 10, 11) | 12 |
| ORDSEA | ORDDFW1DFWSEA1 | ORDSEA1 | 0.01 | 0.005 | (1, 4, 3, 8, 5, 10) | 11 |
| ORDSFO | ORDDFW1DFWSFO1 | ORDDFW2DFWSFO2 | 0.01 | 0.002 | (1, 2, 2, 3, 4, 4) | 5 |
| SATABQ | SATDFW1DFWABQ1 | SATDFW2DFWABQ2 | 0.005 | 0.002 | (2, 2, 4, 5, 5, 6) | 9 |
| SATAUS | SATDFW1DFWAUS1 | SATDFW2DFWAUS2 | 0.002 | 0.01 | (1, 4, 4, 9, 10, 11) | 15 |
| SATDCA | SATDFW1DFWDCA1 | SATDFW2DFWDCA2 | 0.005 | 0.007 | (2, 3, 3, 4, 4, 9) | 9 |
| SATDFW | SATDFW1 | SATDFW2 | 0.002 | 0.005 | (1, 1, 2, 2, 3, 3) | 4 |
| SATJFK | SATDFW1DFWJFK1 | SATDFW2DFWJFK2 | 0.007 | 0.002 | (1, 1, 4, 4, 6, 7) | 8 |
| SATORD | SATDFW1DFWORD1 | SATDFW2DFWORD2 | 0.002 | 0.002 | (3, 4, 8, 8, 10, 10) | 15 |
| SATSFO | SATDFW1DFWSFO1 | SATDFW2DFWSFO2 | 0.005 | 0.01 | (1, 1, 4, 4, 9, 9) | 10 |
| SEABOS | SEADFW1DFWBOS1 | SEADFW2DFWBOS2 | 0.002 | 0.005 | (1, 2, 2, 5, 5, 9) | 10 |
| SEADFW | SEADFW1 | SEADFW2 | 0.002 | 0.01 | (1, 4, 4, 5, 9, 9) | 4 |
| SEAORD | SEAORD1 | SEADFW1DFWORD1 | 0.002 | 0.002 | (1, 1, 2, 2, 4, 4) | 3 |
| SFODFW | SFODFW1 | SFODFW2 | 0.002 | 0.005 | (1, 1, 3, 3, 4, 5) | 2 |
| SFOJFK | SFODFW1DFWJFK1 | SFOJFK1 | 0.005 | 0.002 | (1, 1, 4, 5, 6, 7) | 10 |
| SFOSEA | SFODFW1DFWSEA1 | SFODFW2DFWSEA2 | 0.002 | 0.002 | (1, 2, 3, 4, 5, 6) | 7 |

Table 21: Consideration sets and utility values for the double itineraries in the large network example

| Triple Itinerary Markets | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Market | Itinerary | | | Arrival Rate | | | Preference Vector | No Purchase Utility |
| BOSDCA | BOSDCA1 | BOSDFW1DFWDCA1 | BOSDFW2DFWDCA2 | 0.002 | 0.005 | 0.002 | (1, 1, 2, 2, 3, 3, 4, 4, 5) | 7 |
| BOSORD | BOSORD1 | BOSDFW1DFWORD1 | BOSDFW2DFWORD2 | 0.005 | 0.002 | 0.012 | (1, 2, 2, 3, 4, 5, 6, 6, 7) | 7 |
| BOSSFO | BOSDFW1DFWSFO1 | BOSDFW2DFWSFO2 | BOSSFO1 | 0.002 | 0.002 | 0.005 | (1, 2, 4, 4, 7, 8, 9, 9, 10) | 11 |
| DCABOS | DCADFW1DFWBOS1 | DCADFW2DFWBOS2 | DCABOS1 | 0.024 | 0.027 | 0.049 | (1, 1, 2, 4, 4, 5, 6, 6, 9) | 9 |
| DCADFW | DCADFW1 | DCADFW2 | DCABOS1BOSDFW2 | 0.002 | 0.005 | 0.01 | (2, 2, 2, 3, 4, 6, 7, 8, 9) | 9 |
| DCAJFK | DCADFW1DFWJFK1 | DCADFW2DFWJFK2 | DCAJFK1 | 0.005 | 0.005 | 0.002 | (1, 2, 3, 4, 4, 5, 6, 7, 8) | 8 |
| DFWDCA | DFWDCA1 | DFWDCA2 | DFWBOS1BOSDCA1 | 0.002 | 0.01 | 0.017 | (1, 3, 3, 6, 6, 7, 8, 9, 10) | 12 |
| JFKDCA | JFKDFW1DFWDCA1 | JFKDFW2DFWDCA2 | JFKDCA1 | 0.005 | 0.005 | 0.002 | (1, 2, 3, 4, 4, 5, 6, 7, 8) | 8 |
| JFKDFW | JFKDFW1 | JFKDFW2 | JFKDCA1DCADFW2 | 0.002 | 0.005 | 0.002 | (1, 2, 3, 4, 4, 5, 7, 7, 8) | 8 |
| JFKSFO | JFKDFW1DFWSFO1 | JFKDFW2DFWSFO2 | JFKSFO1 | 0.002 | 0.005 | 0.002 | (1, 3, 3, 4, 4, 6, 7, 8, 9) | 9 |
| SFOBOS | SFOBOS1 | SFODFW1DFWBOS1 | SFODFW2DFWBOS2 | 0.002 | 0.005 | 0.007 | (1, 1, 1, 5, 5, 5, 8, 8, 8) | 3 |

Table 22: Consideration sets and utility values for the triple itineraries in the large network example