Theses and Dissertations

2-1-2012

Modeling and optimization of single-leg multi-fare class overbooking problem: the case of Ethiopian Airlines

Getachew Basa Bonsa

Follow this and additional works at: https://fount.aucegypt.edu/etds

www.manaraa.com

The American University in Cairo

School of Sciences and Engineering

Modeling and Optimization of Single-Leg Multi-Fare Class Overbooking problem: the case of Ethiopian Airlines

A Thesis Submitted to

The Department of Mechanical Engineering in partial fulfillment of the requirements for

The degree of **Master of Science in Engineering**

With specializations in

Industrial Engineering

By Getachew Basa Bonsa

Under the supervision of Dr. Hatem Elayat

Professor of Industrial Engineering

Mechanical Engineering Department, American University in Cairo

Fall 2011

1

The American University in Cairo

School of Sciences and Engineering

Modeling and Optimization of Single-Leg Multi-Fare Class Overbooking problem: the case of Ethiopian Airlines

A Thesis Submitted to

The Department of Mechanical Engineering in partial fulfillment of the requirements for

The degree of **Master of Science in Engineering**

with specializations in

Industrial Engineering

Has been approved by

Dr. Hatem Elayat (Advisor)

Professor, Department of Mechanical Engineering

The American university in Cairo

Dr. Ashraf Nassef

Professor, Department of Mechanical Engineering

The American university in Cairo

Dr. Abdelghani Elimam

Professor, Department of Mechanical Engineering

The American university in Cairo

_____   _____   _____   _____

Program Director   **Date**   **Dean**   **Date**

2

Dedication

To my Mother,

Medhin Assfaw

Acknowledgment

I would like to extend my heartfelt gratitude to my Advisor Dr. Hatem Elayat for his academic guidance throughout this work. If not for his guidance and encouraging words, I hardly imagine the success of completing this work in its current form and the chance of learning that I have had.

I am also thankful to Dr. Ashraf who supported me much in bringing this work to light. It would be wrong of me if I write this section without mentioning his contribution while I was writing the Mat Lab codes, thank you Dr. Ashraf. I would also like to take this opportunity to thank to all of my AUC professors, especially Dr. Abdelghani for his effort to get me on the right track during the initial phase of this work.

Last but not least, I would like to thank my family and friends, who have been on my side throughout the study. Efa and Abi, you are my dear.

Abstract

The American University in Cairo

Modeling and Optimization of a single-leg multi-fare class overbooking problem: the case of Ethiopian airlines

By: Getachew Basa Bonsa

Supervisor: Dr. Hatem Elayat

Revenue Management, also known as yield management, is a technique used by airline industries to maximize revenue by allocating the available seats to the right customers at the right price. Overbooking is an airline revenue management technique that enables airlines to sell more seats than available in order to account for the fact that some of the passengers may not show-up or cancel their flights on the departure day. The objective of this thesis is to develop an overbooking model for a single-leg multi-fare class flight considering a realistic distribution of no-show data collected from the Ethiopian airlines. The overbooking model developed considers the interaction (i.e. the transfer of an extra passenger in a lower fare classes to higher fare class empty seat) between classes that may exist during boarding time. Moreover, this work investigates the economic rationale behind the no overbooking policy used by Ethiopian airlines for some of its flights. The overbooking model developed was solved using both a closed form approach using derivatives and Monte Carlo simulation with a derivative free optimization algorithm. A comparison of the revenue generated from no-overbooking policy, the closed form solution, and the Monte Carlo simulation solution approach shows that the Monte Carlo simulation solution approach performs well. Generally, the numerical results show that the overbooking model is effective in determining the optimal number of overbooking for a number of classes and a variety of compensation cost plans.

List of Tables

List of Figures

Table of Contents

List of Acronyms and Abbreviations

ADD	Addis Ababa Bole International Airport
ARM	Airline Revenue Management
BLF	Break Even Load factor
CAI	Cairo International Airport
cdf	Cumulative distribution function
CRS	Computer reservation system
DBX	Dubai International Airport
ET	Ethiopian Airlines
GED	Generalized extreme value distribution
LF	Load Factor
MatLab	Matrix Laboratory
OB	Overbooking
PARM	Perishable asset revenue management
pdf	Probability density function
PNR	Passenger name record

1.   Chapter One

1.1. Motivation for the Thesis

After deregulation in 1979 airline carriers face a fierce competition due to the ever increasing introduction of low fare carriers in to the airline market [1]. Prior to deregulation airline companies have no power over setting or controlling the price of a ticket and their routes. It was the Civil Aeronautics Board (CAB) that sets the routes and the corresponding fare ticket that an airline has to operate by. As such, carriers simply accept passengers on first come first served policy, since carriers has no control over managing their revenues, until all the available seats are sold. Following deregulation, airline carriers start looking for ways of managing their revenues in order to compete in the market, and this leads to the evolution of different techniques of Airline Revenue Management (ARM). Revenue Management (RM), also called yield management or perishable asset management, is "Selling the right seat at the right time to the right passenger for the right price" [2]. That is, RM seeks to develop effective methodologies to allocate different seats at different prices in order to maximize revenue. As a result of the fierce competition in the market almost all established airlines has less control over the fare structure, but exploit the opportunity of using capacity control and overbooking models in order to compete. Accordingly, RM concentrates on two core types of problems that exist in the airline industry, namely, capacity control and overbooking. Capacity control methods and models primarily provide a decision support for making decision of whether to accept or deny a seat reservation request by a customer at a given time. The second type of problem exists because of the probabilistic nature of show-ups during time of departure. That is, booked customers sometimes may not show-up for different reasons with or without cancellation. In order to avoid spoilage cost (that is lost revenue due to flying with empty seats) airlines are advised to use the methods and models of overbooking. However, sometimes the number of show-ups during time of departure may be higher than the available physical capacity, and in such cases there will be compensation and loss of good will cost incurred by the airline. Hence, the overbooking model will consider all the costs involved in order to determine the optimal number of overbooking to be made in each class for a specific flight so that revenue is maximized. In short, while capacity control allocates seats

for fare-classes, overbooking reserves seats beyond the physical capacity of the aircraft both with the objective of maximizing the revenue.

The overbooking model considers the lost revenue due to cancellations and no-shows that will result in flying with empty seats, and the loss of good will and compensation cost due to excess number of show-ups than the available capacity. The overbooking problem is generally classified as static and dynamic based on the assumptions. Though the dynamic overbooking problem treats the overbooking problem in its realistic nature by taking into consideration the dynamic nature of cancellation over a period of time, it is not used by many airlines due to its mathematical intractability for a real world data. The static overbooking, which many airline use, simplifies the nature of the problem to make it mathematically tractable for real world data and daily use. However, many of the static overbooking models are modeled for a single class problem and did not include the loss of good will cost, and uses simplified form of the compensation cost in the development of the model. Furthermore, the commercial RM models are constructed based on the assumption that the demand distribution is simply the product of the show-up rate and overbooking limit, which is not the case when evaluated both on a theoretical and practical basis [3].

1.2. The Booking System Environment in the Airline Industry

The booking process in an airline reservation system begins with a request by a customer for a particular itinerary. Then the customer will be presented with alternative routes and their corresponding prices for the requested itinerary. An itinerary may involve a single origin and single or multiple destinations. A single origin destination flight is known as single-leg flight, and a flight that involves one or more legs is called multiple-leg flight. If the customers' bid price is greater than or equal to the threshold value of the ticket price for that particular itinerary, the booking operator accepts the request; otherwise he rejects the request. The demand at the point of opening of the booking process, which usually starts three months earlier than the flight date, is low and will increase gradually, and then when the departure date approaches demand falls down. Experience shows that low fare class passengers book earlier than high fare class passengers. In light of this pattern, the

task of the booking operator with the help of the RM tools, which are integrated in the computer reservation system (CRS), is to determine the best policy in determining whether to accept or reject a reservation request made by a customer at period T. This decision is crucial, since selling more seats to low fare-class passengers will lead to the loss of revenue that would have been generated from potential high fare class customers. On the other hand, rejecting many reservation request of a low fare class customer hoping for future high fare class passenger booking request will also increase the risk of flying with empty seats resulting in loss of revenue. In addition to the capacity control problem, the booking operator also has to decide the overbooking pad in each class. The overbooking pad is the number of extra bookings that the airline would like to make in excess of the physical capacity of the aircraft in order to account for the fact that some of the booked customers may not show or cancel during the day of departure. Even with the application of overbooking aircraft may fly with empty seats, in such cases the lost revenue is called spoilage cost.

1.3. RM problems

Revenue management problems occur in almost all service industries where reservation is part of the business process. Transportation sectors such as Airlines, auto rental, railway, tour operators, cargo, and cruises are few examples that use RM tools. In addition, hotels and lodging facilities, healthcare industries, apparel industries, and telecommunication companies are also areas where RM has found application [4]. A comprehensive study on the aspects of revenue management in the airline industry can be found in [5]. [6] has also outlined the characteristics of revenue management looking at it from a general perspective in addition to introducing a new term; Perishable Asset Revenue management (PARM). A comprehensive survey of RM problems can be found in [7] and [8]. The following characteristics are some of the common denominator found in the problem of RM in all the service industries [9].

1. Capacity is fixed,
2. Capacity is a perishable asset,
3. Available seat or asset can be reserved or sold in advance,
4. Demand is very erratic,

5. Demand can be partitioned in accordance with the associated price of the seat

As pointed out in the earlier section, the two most important questions that the RM system tries to answer are the seat inventory control and the overbooking level. Since there is little control over price because of the competition in the airline industry to improve revenue, the airline industries focus has shifted in trying to find the optimal number of mixes of discount fares and full fares in order to maximize their revenue. According to [10] the profit contribution of a full fare sale over a discounted fare was estimated to be around $50 million annually, making the consideration of the seat inventory control an important aspect in RM. In addition, booking level control has also produced a substantial profit margin that it has to be considered as an important aspect of RM. For example, Lufthansa, the German airline, credits a revenue increase of € 150 million in 2005 for the practice of overbooking, which makes overbooking one of the prominent airline revenue management techniques used at Lufthansa [11]. However, though capacity control and overbooking are the predominantly considered RM problems, McGill and van Ryzin has also identified pricing and forecasting as additional RM problems [7]. As outlined above, since the fare structure is predominantly affected by the fare structure offered by competitors of the same flight in the market, it makes the control over price extremely difficult for the airline industries. Consequently, pricing has drawn little interest and attention for researchers. In comparison, forecasting, which is at the heart of all RM tools since the inputs in working with capacity control and overbooking are drawn from the forecasts, has drawn some attention recently [12]. Of all the four mentioned RM problems, this thesis considers the overbooking problem and a comprehensive definition and nature of the problem is presented below.

1.4. *Overbooking*

Overbooking is the process of selling more flight tickets than the available physical seats [13]. Overbooking is practiced in order to compensate the number of cancellations and no-shows that occur during the departure time. According to [11], on average 15% of American airline seats could have been spoiled if overbooking were not practiced. Though overbooking could save or generate a significant amount

of the revenue, it could also result in compensation cost or payouts and loss of customer goodwill cost when the number of show-ups exceeds the available capacity. Especially, when there are a number of competitors in the market, loss of customer goodwill should not be tolerated, as such this paper tries to model the cost of loss of customer goodwill and incorporate it in the cost function model. Therefore, the aim of solving the overbooking problem is to calculate the optimal number of excess seat pads (ticket to be sold) while maximizing the net profit, which is the revenue produced from selling tickets minus the compensation, loss of customer goodwill, and expected lost revenue. The specific problem to be studied in this thesis is explained below, and will also be elaborated in the subsequent sections. Before introducing the overbooking problem, it is crucial to describe the nature and characteristics of overbooking as it exists in the airline industry.

✓ The booking operator must decide the optimal number of overbooking during the opening time of the booking process.
✓ The booking operator cannot observe the no-show during making the decision of making the overbooking
✓ No opportunity to recover the cost of flying with empty seats. That is, the seat of an airline is perishable (has specific time of use).
✓ The show-up of passengers is very erratic
✓ If show-up exceeds the seat capacity, denied boarding passengers should be compensated
✓ If capacity exceeds the show-up, cost of lost opportunity will be incurred by the airline

1.5.   Problem Definition

Overbooking is an airline revenue management (ARM) technique which seeks to account for the no-shows and cancellations by making more reservations than the available capacity in order to maximize revenue. The approaches for the overbooking problem can be broadly categorized as static and dynamic models. In the static model, the dynamic nature of reservation (cancellations over a period of

time) is ignored, and the concern is to find the optimal number of overbooking at the opening period of the reservation that minimizes cost or maximizes revenue. The dynamic model considers the dynamic nature of reservation, and seeks to find a policy by which the booking operator decides whether to accept or reject a request made by a customer for a reservation of a certain class at time T. Although dynamic overbooking models treat the overbooking problem in its realistic state, generally the models are mathematically intractable for a real world problem. As such, many of the commercial RM systems used by the airlines are static models [3]. Therefore, this thesis seeks to extend the static overbooking model by incorporating a realistic cost function of overbooking and relaxing some of the assumptions made in prior studies.

1.6.   Objectives

The objectives of this study include:

- ✓ Model the overbooking problem for a single-leg multi-fare class problem as a cost minimization and a revenue maximization problem.
- ✓ Develop the overbooking problem in such a way that it could be constrained by a user defined probability of loss of the revenue.
- ✓ Compare the results of both models
- ✓ Model the cost function of the overbooking problem based on a realistic distribution of the no-show data.
- ✓ Propose a suitable optimization procedure for the overbooking model

1.7.   Thesis Overview

A literature review of the overbooking models is explained and presented in chapter 2. Chapter 3 presents the mathematical formulation of the overbooking problem using a realistic distribution of the no-show data, and a solution approach using both the closed form and a Monte Carlo simulation for the use of the derivative free Nelder Mead algorithm [14]. Chapter 4 presents a numerical analysis and evaluation of the proposed solution approaches in solving the overbooking model. Finally, Chapter 5 and 6 presents the conclusion and recommendation for future work respectively.

2. RM Methods and Literature Review

2.1. *Overbooking*

Overbooking is the practice of intentionally selling more seats than the available physical capacity of the plane in order to compensate the number of no-shows and cancellation, which can be as high as 15% [15], during the time of departure. A more recent study shows that the benefits obtained from using overbooking accounts for an average of $1 billion increase in revenue per year [16]. Though overbooking can improve the revenue of an airline it has also risks associated with it, when the number of show-ups is greater than the fixed capacity. That is, when the number of show-ups is greater than the available capacity, some of the passengers who already bought a ticket will be bumped (i.e. denied boarding) of the flight either voluntarily or involuntarily. In both case there is a financial loss that the airline should incur in the form of compensation cost to be paid toward the bumped passengers. In addition to the compensation cost, the bumped passengers will retain a bad image of the service that should be considered as loss of customer goodwill cost, which will have a massive long term impact on the business of the airline. However, it was estimated that financial loss due to overbooking is less when compared with not practicing overbooking [17]. Accordingly, the objective of the overbooking model is to find the optimal number of overbooking level that the airline should reserve in order to minimize the expected cost or maximize the expected revenue.

The history of overbooking goes back to the pioneering work of Beckmann and Bobkowski [18]. Their statistical modeling of the overbooking problem laid a foundation for today's revenue management in the airline industry. The first overbooking model proposed by Beckmann was a single leg single fare-class problem, which is a very simplified form of the actual overbooking problem that airline faces. His model tries to determine the optimal overbooking level by balancing the spoilage cost (lost revenue due to empty seats) with compensation cost (lost revenue due to bumping of passengers). Thompson [19] developed an overbooking model for a two fare class using the cancellation rates while ignoring the probability distribution of the demand and the no-show rates. His model determines the overbooking limit for a given probability of overbooking. Thompson's work has been

extended by Taylor as well as by Rothstein and Stone [20]. Taylor's overbooking model, though is a very simplified model, has been implemented and used by many airlines for their booking level control. It was also considered that Taylor's model was used as a basis for a family of subsequent overbooking models. Bodily and Pfeifer [21] also studied the static overbooking problem using the probability of customer cancellation and no-shows for a single fare-class problem, which is a highly simplified form of the actual scenario. All the above models deal either with a single fare-class or two fare-class overbooking model, which is not always the case for a real world problem. Latter researches, however, consider the multi fare-class overbooking problem [22],[23],[24]. Chi [22] considers the multi fare-class overbooking problem and develops a dynamic programming model. His model determines the maximum overbooking level that should be used in every fare-class for a known demand and show-up distribution of every class. He further assumed that cancellations can be made without any penalty cost, which made his model inaccurate since there is a penalty for cancellations. Coughlan [23] extends the multi fare-class overbooking problem by introducing the last minute passengers (also called go-shows, are customers who show-up during service time without any prior reservation). His model assumes the demand, the show demand, and the cancellations are all independently normally distributed. However, the assumption that the booking is independently normally distributed is incorrect [19], and in the literature it commonly is assumed to follow a Poisson distribution. Furthermore, his proposed direct search algorithm for solving the complicated closed form overbooking model doesn't guarantee optimality. [24] developed a mathematically tractable static and dynamic overbooking model that provides an upper and lower bound for the overbooking level based on the expected revenue approach. They proposed two different static overbooking models based on the demand information available for the user. Moreover, no-shows and cancellation probabilities are considered class based in order to make the model more realistic. However, their model like all the models in this class does not consider the interaction that exists between classes.

The methods to make overbooking decisions are broadly categorized in to two, namely, static and dynamic overbooking. A review of the two approaches along with their advantages and disadvantages is presented below.

2.2.1 Static overbooking

Because of its simplicity and applicability for real world data, the static overbooking model is the widely used approach in the airline industry. The static overbooking model did not consider the dynamic nature of customer cancellations overtime. What this model does is that it determines the excess amount of ticket sells to be made by considering the number of cancellation and no-shows from a probability distribution, where the parameters of it will be updated every time a new data is available for consideration. In static model the distinction between no-shows and cancellation is unnecessary, since cancellations (that happened before the time of departure) could be substituted by other customers. However, cancellation that may occur during the day of departure may simply be considered as no-shows in the formulation of the static overbooking problem. Hence, the important factor in getting the overbooking level in a static overbooking model is to determine the show-up (show demand as it is commonly referred) rates of that particular flight. One of the widely used static overbooking models is based on the binomial distribution for the show-ups [25], since the cancellation which occurs during the service time can lump with the no-shows. Other models use the normal distribution and the beta distribution for the show-ups in modeling the overbooking [26]. In its simplified form, the static overbooking model is similar to that of the newsboy problem. Though the static model is simple, flexible, and mathematically tractable for real world data, it failed to capture the dynamic nature of customer cancellations that occur in the course of the reservation period. The approach that considers the dynamic nature of cancellations so that treating the overbooking problem relatively in its realistic state is known as dynamic overbooking model. A comprehensive discussion the available literature on dynamic overbooking is presented in the following section.

2.2.2 Dynamic overbooking

The need to include cancellations that occur during the course of the booking process (thereby eliminate the drawback of the static overbooking model) made researchers to model the overbooking problem as a dynamic model. Though, there

are so many dynamic overbooking models available for the single leg overbooking problem, due to their mathematically intractability for real world data, airlines use the static overbooking models. The models in this class are generally formulated based on the Markov Decision Process (MDP) [27],[28],[29]. Rothstein [27] was the first to formulate the single leg single fare-class overbooking as a dynamic programming problem that determines an overbooking policy. In addition, Rothstein assumed that the probability of cancellation is independent of the number of already made reservations, which could affect the accuracy of the model. His general model, however, were mathematically intractable due the curse of dimensionality as the state space of the system is the number of reservation, which is substantially large for a dynamic programming approach. In order to overcome the computational difficulties [15] proposed two methods: (1) reducing the size of the state by aggregating them or (2) "develop a theory of structure of optimal solution" in order to reduce the time for computation. Alstrup et al [28] followed the first approach proposed by Chatwin in their dynamic programming model formulation for a two fare-class problem. Their model not only extends Rosthetein by considering a two fare class problem, but also considers the cost of down grading customers (that is, the cost of allocating seats of high fare contenders in a lower fare-class seat). In order to reduce the size of the state space of the system, they grouped the reservation and cancellations in group of five, reducing the size by a factor of 25. Their model is solved by two dimensional stochastic dynamic programming. The second approach was used by Chatwin himself in his thesis, in which he proposes two multi-period overbooking models for a single-leg single fare-class service.

2.3 Overbooking in practice

The overbooking models in use today are mainly the static models based on simplifying assumptions regarding the distribution of no-shows, demand, and fare-class. However, the literature is full of dynamic overbooking models, which has found relatively no use in practice since those models are mainly mathematically intractable and require a lot of time to solve them. Moreover, the booking personnel in the airline industries are not optimization experts to understand and fully utilize the

advantages of the dynamic overbooking model. To that effect, the booking experts would prefer to use a simple model (one that has fewer input data) to estimate the level of overbooking. They also prefer the static overbooking over the dynamic overbooking for the static overbooking model requires a single run while the dynamic overbooking requires running the model now and then as far as new booking and cancellations are made. The widely known commercial revenue management software (PROS) has an overbooking module embedded in it.

2.4 Literature gap and contribution

Generally, the literature on overbooking could be categorized as static and dynamic models. The static overbooking model could be further categorized based on the number of classes the overbooking model deals with (usually, a single fare class is considered). Those models which are constructed for a single fare class fail to capture the value of the different seats in the classes by making all classes as having equal value. Furthermore, models constructed for two class case did not consider the interaction between classes that exist in the real world system. That is, since upgrading a low fare class seat customer to a high fare class seat is possible, the interaction between classes should not be ignored in the overbooking model.

In this thesis a static overbooking model is developed using two different probability distributions (Binomial and the generalized extreme value distribution) for the show-up or no-show in modeling the cost function. An attempt was made to solve the model using both closed form expression and a Mote-Carlo simulation using the derivative free optimization approaches. Furthermore, the model was made to be flexible so that it could be transformed with a user defined constraint into a constrained optimization problem. This particular feature of this overbooking model is important for decision makers who are sensitive to both customer reaction upon denied boarding and profit loss. The model developed in this thesis could be used for any classes the airline wish to make and for any kind of distribution that the particular airline's data may have. An attempt to include the loss of good will cost, which was not included in overbooking models in past papers, in order to make the cost function realistic, was made using the Taguchi Quality Loss function [30]. Furthermore, the fact that the paper models the cost function based on realistic probability

www.manaraa.com

distributions based on the historical data is a relaxation of the assumptions made in prior studies since in the past the cost function was mainly modeled based on the binomial distribution.

2.5 Ethiopian Airlines

2.5.1 Company Overview

Ethiopian airlines, also called Ethiopian, were founded in 1945 as the flag carrier of Ethiopia, operating out of Bole International Airport, Addis Ababa, Ethiopia. Currently, Ethiopian is one of the youngest and largest air carriers in the region, known for its service excellence and multiple routes in the region. Ethiopian is also known for its operational excellence and one of the most profitable carrier even during the recent financial crisis. Ethiopian serves 63 international destinations of which 40 destinations are in Africa, 8 destination in Europe and the Americas, 15 destinations Middle East and Asia, and 17 destinations domestic with a total of 46 aircrafts. Ethiopian has received awards and recognitions for its service and operational excellence from different organizations. The following are some of the awards the airline has archived during the course of its service.

- Bombardier's "Airline Reliability Performance Award",2011
- "Deal of the Year 2010 Award " , 2011
- "The African Cargo Airline of the Year",2011
- "The NEPAD Transport Infrastructure Excellence Awards", 2009
- "Airline of the Year Award",2009
- "2008 Best Airline in Africa Award"
- "The 2008 Brussels Airport Company Award"
- "The 2008 Corporate Achievement Award"

The following table summarizes the number and type of aircrafts that Ethiopian uses for its fleets to serve the 63 international destinations[1].

---

[1] http://www.ethiopianairlines.com/en/corporate/fleet.aspx

21

Aircraft type	Number	Seat capacity		Total
		B	E	
Boeing737-700	5	16	102	118
Boeing737-800	5	16	138	154
Boeing757-200	3		144	160
	1		154	170
	2	16	155	171
	1		159	175
Boeing767-300ER	1		208	232
	1		210	234
	2	24	211	235
	2		213	237
	3		221	245
	1	30	190	220
	1		195	225
Boeing777-200LR	5	34	287	321
Boeing787-DreamLiner (on order)	10	24	246	270
A350-900(on order from Air Bus)	12	30	318	348
Bombardier Dash 8 Q400	8	0	78	78

Table 1: Number and type of aircrafts operated by Ethiopian

2.5.2 Overbooking practice at Ethiopian airlines

Like many airlines, Ethiopian also claims to use overbooking and seat inventory control models to boost its revenue and compete in the market. Currently, Ethiopian uses the commercial revenue management system (i.e. PROS). PROS uses forecasting models (techniques) to determine the expected number of no-shows and cancellation, and make recommendation using the built in algorithms on the number of overbooking. As all other revenue management systems PROS also try to balance the risk of flying with empty seats and the risk of denying boarding. PROS has the passenger name record (PNR) and non PNR forecasting techniques. Considering Ethiopian data recording management system one can clearly see how Ethiopian

can practice overbooking using PROS. From the collected data one can see that there is no instance that show a practice of overbooking though the personnel claims they did. However, if they do overbook it can be said that Ethiopian can only use the non PNR no-show forecasting technique to determine the number of overbooking. Though Ethiopian claims to use overbooking in its practice, the number of empty flight seats due to either no-shows or cancellations or both in 2008 for example was estimated at 146,153. In 2009, 209,330 of cancellation and no-shows was recorded, which could be translated as 1774 full Boeing 737-800. In other words, the average load factor of all ET flights for 2008 and 2009 are approximately 73% and 69% respectively. In other words, almost 27% and 31% of the seats in 2008 and 2009 were spoiled due to no-shows and cancellations, resulting in lost revenue or cost of lost opportunity. This figure evidently shows that the overbooking model or procedure that Ethiopian is currently using could not capture the lost opportunity. The problem could be originated either from the wrong application (use) of the revenue management system that they are using, or the inherent drawback of the overbooking model/module embedded in PROS. However, I could not check where the problem exactly is, due to the company's policy that prohibits examination of internal working procedures including primary data collection. Nevertheless, according to one of the heads of the booking section, Ethiopian overbooks based on the no-show rate without using the optimization module. That is, the overbooking level is calculated or set equal to the forecasted number of no-shows. The total average rate of no-shows and cancellations for Ethiopian considering all destination flights is approximately 27%, which is a significant rate that has to be addressed. That is, approximately 27 % of the booked passengers did not show-up at the gate for flight due to either cancellation or no-shows or both. The following charts summarize the number of no-shows that are reported during a six month time in 2008 for flights to Middle East and African destinations respectively.
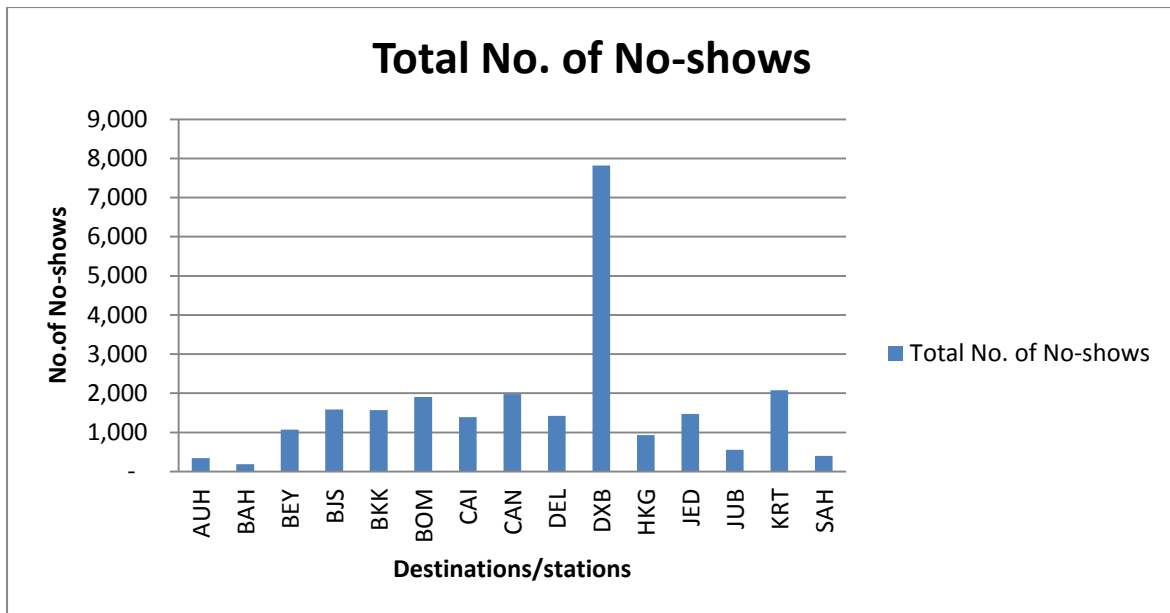
Figure 1: No-show data for Asia and Middle East destination during a six month period
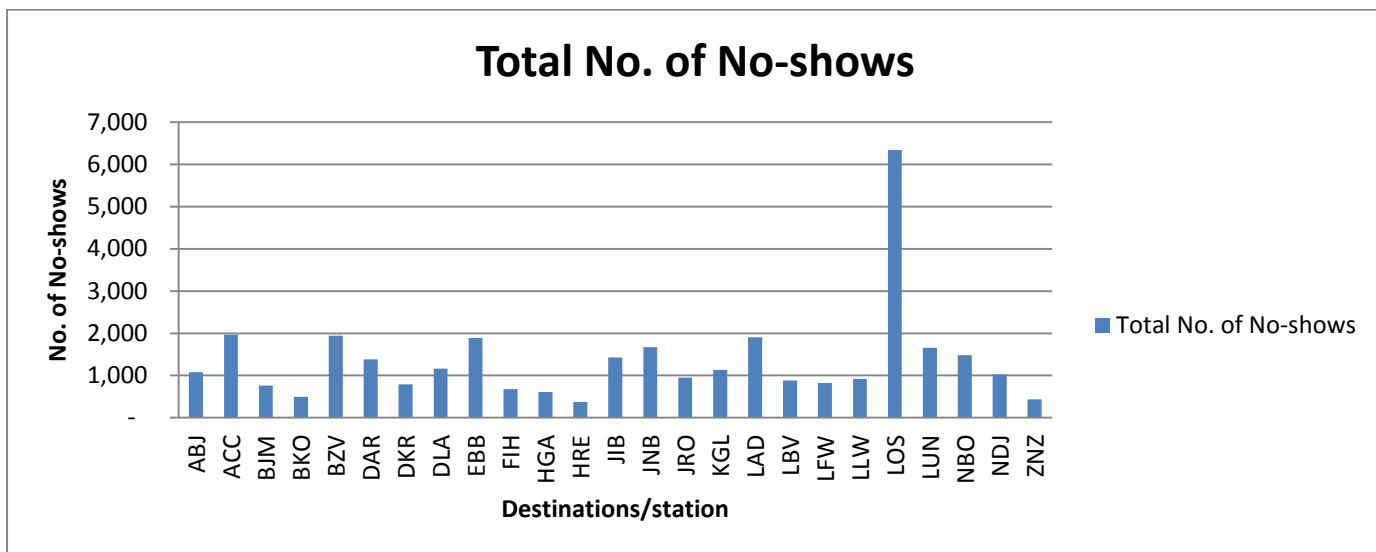
Figure 2: No-show data for African destination during a six month period

2.6 Seat inventory control

The seat inventory control is one of the most studied revenue management in the airline industry. The seat inventory control problem is concerned with allocating the seats to different prices so as to maximize revenue. In short, it tries to find the right mix of low fare seats and high fare seats in order to capture the demand over time

thereby maximize revenue. Thought it is possible to capture the maximum demand by selling more discounted (low fare) seats, it could also result in losing potential customers who are willing to pay the full fare. On the other hand, denying requests of booking of low fare customers anticipating future arrival of high fare customers may result in flying with empty seats, which is lost revenue. Hence, managing the trade of is the primary purpose of the seat inventory control. However, it is difficult to determine the demand of each class earlier in time, since demand is erratic, that makes the seat inventory control problem difficult. That is, when and how to make the trade of is the essential question that the seat inventory control module could answer. Generally, the seat inventory control module will provide a protection level of the seats for each fare classes so that revenue will be maximized. This could be done either ahead of time (static seat inventory control) or during the booking process (dynamic seat inventory control). The seat inventory control could be applied either for a single leg flight or for a network of flights. For example, consider the two leg flight from Addis Ababa-Khartoum-Cairo. If the single leg approach is used, a low fare passenger who wants to travel from ADD-KTR-CAI might be denied a seat in preference of a high fare passenger who wants to travel from ADD-KTR. The single leg seat inventory control model could result in loss of revenue that could have been generated from a consideration of the combination of the full flight network. However, since airlines have a huge number of flight networks the seat inventory problem will be more difficult to analyze as compared to the example presented here in this section.

The static seat inventory control model determines the right mix of low fare class seats and high fare class seats before the booking process starts using a demand forecast as an input [5]. However, since demand arrival could be different from the forecast (that is used as an input in the static seat inventory control), the model could fail to account this fact. Hence, in order to consider the realistic situation and come up with a more accurate seat protection level, researchers have developed a dynamic seat inventory control model [5].

The interested read could find a detailed literature review of seat inventory control in [4]. However, here are some of the prominent papers that deal with the problem in discussion. Brumelle and McGill [31], Littlewood [32], and Belobaba [33] present a static single leg seat inventory control model. Belobaba's Expected Marginal Seat

Revenue (EMSR) laid the foundation and framework for dealing multiple fare class seat inventory problems. Williamson [34] developed a model that accounts the network interactions that eliminates the draw backs of the static seat inventory model.

26

3. Model Development

3.1 Proposed Mathematical Model

Problem statement

Overbooking is an airline revenue management (ARM) technique which seeks to account for the no-shows and cancellations by making more reservations than the available capacity in order to maximize revenue. The approaches for the overbooking problem can be broadly categorized as static and dynamic models. In the static model, the dynamic nature of reservation (cancellations over a period of time) is ignored, and the concern is to find the optimal number of overbooking at the opening period of the reservation that minimizes cost. The dynamic model considers the dynamic nature of reservation, and seeks to find a policy by which the booking operator decides whether to accept or reject a request made by a customer for a reservation of a certain class at time T. Although dynamic overbooking models treat the overbooking problem in its realistic state, generally the models are mathematically intractable for a real world problem. As such, many of the commercial RM systems used by the airlines are static models (Amaruchkul et al., 2011). Therefore, this paper seeks to extend the static overbooking model by incorporating a realistic cost function of overbooking and relaxing some of the assumption made in prior studies.

Mathematical formulation of the Problem

Consider a single leg flight having a maximum capacity of C, with multiple (m) fare classes. The booking operator accepts customers request for booking or cancellations for an already made reservation until the day of departure. A passenger who made a reservation may not show-up on the departure day or cancels his reservation at any time before and on the day of departure. Cancellations have a refund which is proportional to and a fraction of the fare ticket that the customer already bought. In order to accommodate for the no-shows and cancellations, the airline should make overbooking. However, if the number of customers that show up exceeds the maximum capacity of the airplane, customers will be bumped either voluntarily or involuntarily, which in both cases the airline has

to make a compensation for its bumped passengers. If the number of show-ups during the time of departure is less than the capacity, the aircraft will fly with empty seats resulting in lost revenue. Although the overbooking problem was extensively studied, the proposed models are mainly based on some simplifying assumptions. Hence the objective is, to develop a mathematical model that determine the optimal number of overbooking which minimizes the compensation, loss of revenue and loss of customer goodwill cost in order to maximize the expected revenue of the airline while relaxing some of the assumption made in prior studies. Two static overbooking models will be developed. The first will determine the total overbooking limit without considering the different fare classes. The second model will consider the class dependent cancellations and no-shows, and the associated costs to model the overbooking problem. However, since these models were developed before having the data from Ethiopian airlines, some of the input parameters used in developing these two approaches were found to be inapplicable with the current data the airline has. Therefore, it was necessary to develop a model that can be used with the current data structure without compromising the qualities of those already developed models. In effect, a stochastic overbooking model using Mote Carlo simulation approach was used to determine the optimal number of overbooking.

Notations

C= Capacity, $C = \sum_{i=1}^{m} c_i$, where i is an index of the booking class i=1,2,3,…,m

t_i= Demand in fare class i.

f_i= ticket price for fare-class i

y_i= number of overbooking for fare class i , $Y = \sum_{i=1}^{m} y_i$

s_i = penalty cost of an overbooking corresponding to fare-class i

e_i= the amount of refund for fare class i

P_i= probability that a booked seat is in fare class i $\sum_{i=1}^{m} P_i = 1$

β_i = show-up probability of fare-class i

δ_i = cancellation probability of fare class i

α_i = probability of involuntarily bumped passengers in fare class i

Theoretical Framework

The following chart shows the conceptual model of the overbooking problem, up on which the mathematical model will be built on.

```
                    ┌─────────────────────┐
                    │ Total number of show-ups │
                    │         (w)          │
                    └─────────────────────┘
                              │
                              ▼
                    ╱─────────────────╲            ┌──────────────────────┐
                   ╱   W=capacity(      ╲   Yes    │ -No compensation cost │
                   ╲        c)          ╱ ───────> │ -No lost revenue      │
                    ╲─────────────────╱            │ -Flight is full       │
                              │                    └──────────────────────┘
                              │ No
                              ▼
           ┌──────────────────┴──────────────────┐
           ▼                                      ▼
     ┌───────────┐                          ┌───────────┐
     │    w>c    │                          │    w<c    │
     └───────────┘                          └───────────┘
```

Involuntarily Bumped Passengers (BP)
 ➢ Cost of loss of customer goodwill
 ➢ Compensation cost

Voluntarily Bumped Passengers (BP)
 ➢ Compensation cost

Fly with empty seat
 ➢ Cost of lost opportunity or lost revenue
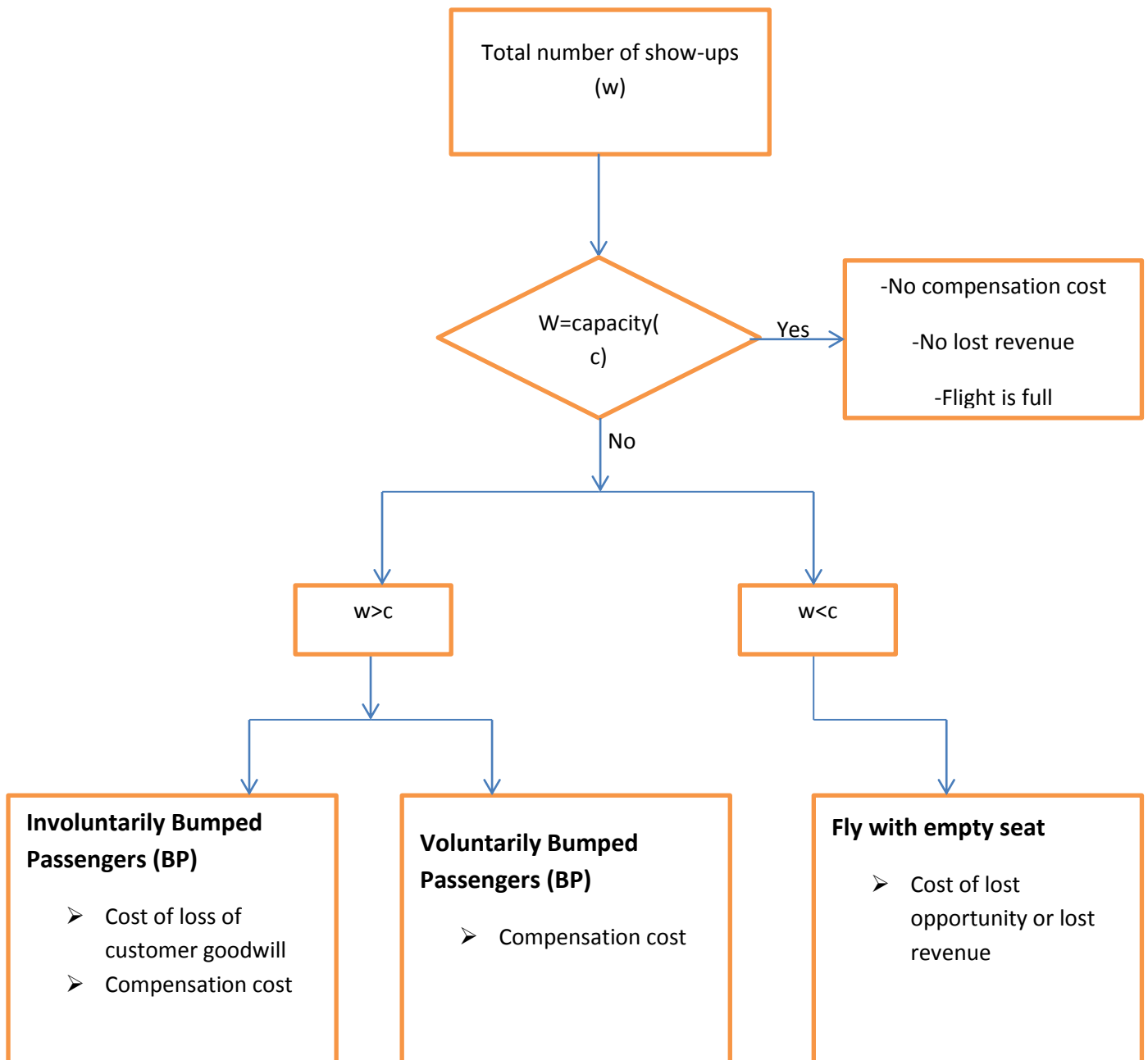
Figure 3: a theoretical framework showing the four possible outcomes under overbooking

www.manaraa.com

3.2 Closed Form approach

3.2.1 Case I- show-up follows Binomial Distribution

Revenue

The number of bookings (n_i) for a certain class is the minimum of the total random demand or the capacity add up with the optimal overbooking of that specific class.

i.e, $n_i = \min\{C_i + y_i, t_i\}$ for   i=1, 2, 3,…,m

Hence the total random booking for a certain flight will be:

$N = \sum_{i=1}^{m} \min\{C_i + y_i, t_i\}$

The demand for each class or the total random demand is assumed to follow binomial distribution. Hence, the expected revenue from all the different classes can be modeled as follows.

The expected number of bookings for a certain class i is:

$$E(n_i) = p_i \times N$$

Hence the corresponding expected revenue R will be:

$$E(R) = \sum_{i=1}^{m} f_i \times p_i \times N = \sum_{i=1}^{m} f_i \times E(n_i) \dots\dots\dots\dots\dots\dots\dots (1)$$

However, the above expected revenue is just calculated without considering the number of cancellations, which are entitled for a fraction of the fare they paid. The lost revenue due to cancellations will be included under the spoilage cost and will be subtracted from the above expected revenue in order to find the actual revenue.

Compensation cost

The compensation cost is incurred when the number of show-ups during the departure time exceeds the available capacity of the aircraft. In such cases, the extra passengers, who are either voluntarily or involuntarily bumped from boarding, should be compensated by providing them accommodation until they get a seat on the next

flight on the same airline or on a different air carrier. Compensation may also include monetary values in addition to accommodation.

Assuming that the show-up follows binomial distribution, the probability that there are exactly ω_i show-ups out of the $(C_i + y_i)$ bookings is:

$$P(W_i = \omega_i) = \binom{C_i + y_i}{\omega_i} \beta_i^{\omega_i} (1 - \beta_i)^{(C_i + y_i - \omega_i)} \qquad \text{for } i = 1,2,3,\dots,m$$

The expected number of show-ups will be:

$$E(\omega_i) = \sum_{\omega_i=0}^{c_i+y_i} \omega_i \binom{C_i + y_i}{\omega_i} \beta_i^{\omega_i} (1 - \beta_i)^{(C_i + y_i - \omega_i)}$$

$$E(\omega_i) = \beta_i \times p_i \times N$$

$$E(\omega) = \sum_{i=1}^{m} \sum_{\omega_i=0}^{c_i+y_i} \omega_i \binom{C_i + y_i}{\omega_i} \beta_i^{\omega_i} (1 - \beta_i)^{(C_i + y_i - \omega_i)}$$

$$E(\omega) = \sum_{i=1}^{m} E(\omega_i)$$

Now let us say the function $F(\omega_i, c_i)$ is the compensation cost, then:

$$F(\omega_i, c_i) = \begin{cases} 0 & \omega_i \leq C_i \\ \sum_{i=1}^{m} s_{i[E(\omega_i) - C_i - \sum_{i=i+1}^{m}(C_i - E(\omega_i))]} & \text{for} \quad \omega_i > C_i \text{ and } \omega_{i+1} < C_{i+1} \\ \sum_{i=1}^{m} s_{i[E(\omega_i) - C_i]} & \omega_i > C_i \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots(2) \end{cases}$$

The second element of the cost function explains the fact that if there are extra numbers of show-ups in the ith class, they can be made to board to the (i+n) classes if there are empty seats in the (i+n) classes. Allocating high fare class contenders to an empty lower fare class seat is considered as downgrading, and is not allowed to be practiced as an option (at least in theory). In the first element of the above equation the value of the cost function is zero, though there is a loss in revenue (when $\omega_i < c_i$) since there will be an empty seat in the flight, and this cost is termed as a spoilage cost, which should be considered differently.

Cost due to loss of customer goodwill

Customer goodwill is an important market share factor in the hospitality industry in general. Involuntarily bumped passengers regardless of the monetary compensations they are entitled with upon denied boarding, they will retain a bad image about the service of the airline especially if they knew they were overbooked than it was because of a minor clerical error. Though airlines did not inform their denied boarding customers about the true reason, customers will find out the reason in one way or another. As such, considering the loss of customer goodwill cost in the overbooking model will improve the performance of the model in managing customer perception of quality, which could have an impact on the revenue of the airline in the long run if not considered. The overbooking models in the literature did not consider this important factor, though many explained its qualitative impact on the airline's market share in the long run. Therefore, modeling the loss of customer goodwill cost is important in minimizing the risk of losing potential customers. Being the case, this thesis proposes the use of quality loss function of Taguchi in modeling [30] the cost of loss of customer goodwill.

Taguchi method could be used to model the cost of loss of goodwill of involuntarily bumped passengers. Of the expected ω_i show-ups let us assume xi passengers are involuntarily bumped passengers. The nominal value/number of involuntarily bumped passengers should always be assumed zero, since involuntarily bumping is undesirable. Now, using the quality loss function, the loss of goodwill cost can be modeled as follows.

$$L(x_i) = \sum_{i=1}^{m} kx_i^2$$

Furthermore, using the binomial model, the probability that there are exactly x_i involuntarily bumped passengers out of the expected ω_i show-ups will be:

$$P\,(X_i = x_i) \; = \; \binom{\omega_i - c_i}{x_i} \alpha_i^{x_i}\,(1 - \alpha_i)^{(\omega_i - c_i - x_i)} \qquad \text{for } i = 1,2,3,\ldots,m$$

Hence, the expected number of involuntarily bumped passengers is:

$$E(x_i) = \sum_{x_i=0}^{\omega_i - c_i} x_i \binom{\omega_i - c_i}{x_i} \alpha_i^{x_i} (1 - \alpha_i)^{(\omega_i - c_i - x_i)}$$

However, in its shortest form the expected number of the involuntarily bumped passengers is:

$$E(x_i) = \alpha_i \times (\beta_i \times p_i \times N - C_i)$$

Hence, the expected loss of customer good will cost due to involuntarily bumping is:

$$E(L) = \sum_{i=1}^{m} k\{\alpha_i \times (\beta_i \times p_i \times N - C_i)\}^2 \dots \dots \dots \dots \dots \dots \dots \dots (3)$$

Spoilage cost

Spoilage cost is incurred when the number of reserved show-ups is less than the capacity available. This could happen when customers cancel their reservation or did not show-up without cancelling. In the former case customers are entitled to a refund, which is a fraction of the ticket fare, up on cancellation. Customers who did not show-up at the departure time without cancelling their reservations will not be refunded whatsoever.

No show probability $= 1 - \beta_i$

Hence the probability of a refund will be:

$$= (\text{No show probability}) \times (\text{cancellation probability})$$

$$= (1 - \beta_i) \times \delta_i$$

Let r_i be the expected number of no-shows with cancellations, who are entitled for a refund, out of the $(C_i + y_i)$ bookings. Hence, the probability that there are exactly r_i no-shows with cancellation out of the $(C_i + y_i)$ bookings is:

$$P(R_i = r_i) = \binom{C_i + y_i}{r_i} ((1 - \beta_i) \times \delta_i)^{r_i} (1 - (1 - \beta_i) \times \delta_i)^{(C_i + y_i - r_i)}$$

The expected number of no-shows with cancellations would be:

33

$$E(r_i) = \sum_{r_i=1}^{c_i+y_i} r_i \binom{C_i + y_i}{r_i} \left((1 - \beta_i) \times \delta_i\right)^{r_i} \left(1 - (1 - \beta_i) \times \delta_i\right)^{(C_i+y_i-r_i)}$$

Or in short,

$$E(r_i) = (1 - \beta_i) \times \delta_i \times p_i \times N$$

The expected lost revenue due to cancellations, L(r)

$$L(r) = \sum_{i=1}^{m} \sum_{r_i=1}^{c_i+y_i} e_i r_i \binom{C_i + y_i}{r_i} \left((1 - \beta_i) \times \delta_i\right)^{r_i} \left(1 - (1 - \beta_i) \times \delta_i\right)^{(C_i+y_i-r_i)}$$

$$L(r) = \sum_{i=1}^{m} \delta_i \times (1 - \beta_i) \times p_i \times N \times e_i \qquad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (4)$$

No shows without cancellations could be obtained as follows:

$$E(r_i) = (1 - \beta_i) \times (1 - \delta_i) \times p_i \times N$$

Hence the expected lost revenue due to no-shows without cancellations could be obtained by multiplying the fare of a class by the number of no-shows as given below:

$$L(r_i, no - shows\ without\ cancellations) = \sum_{i=1}^{m} \delta_i \times (1 - \beta_i) \times p_i \times N \times f_i$$

Finally, the Net Expected Revenue (NER) at departure time will be:

$$NER = E(R) - F(\omega_i, c_i) - E(L) - L(r) \dots\dots\dots\dots\dots\dots (5)$$

Hence the objective is to maximize the Net Expected Revenue:

Maximize

$$NER\,(N) = \sum_{i=1}^{m} f_i \times p_i \times N - F(\omega_i, c_i) - \sum_{i=1}^{m} k\{\alpha_i \times (\beta_i \times p_i \times N - C_i)\}^2$$
$$- \sum_{i=1}^{m} \delta_i \times (1 - \beta_i) \times p_i \times N \times e_i$$

The above model could be used to find the optimal number of the overall overbooking pad for the specific flight. However, sometimes it might be desirable to find the optimal number of overbooking for each class rather than the optimal number of overbooking for the flight. In such cases, a slight modification of the above model is essential, and the following section presents a model for class dependent overbooking.

Class Dependent Overbooking Model

Keeping all the assumption regarding the show-up distribution as binomial, the optimal overbooking limit for each class could be modeled as follows.

Revenue

The revenue generated by a certain class is the product of the ticket fare by the number of bookings of that class. Hence the total revenue generated by all the bookings could be:

$$E(R) = \sum_{i=1}^{m} f_i \times E(n_i) \dots \dots \dots \dots \dots \dots \dots \dots \dots . (1)$$

Compensation cost

Following the same reasoning as presented in the optimal overall overbooking model presented in section I, the compensation cost in case of a class dependent overbooking will be:

$$F\left(\omega_i, c_i\right) = \begin{cases} 0 & \omega_i \leq C_i \\ \sum_{i=1}^{m} s_{i[E(\omega_i) - C_i - \sum_{i=i+1}^{m}(C_i - E(\omega_i))]} & \text{for} \quad \omega_i > C_i \text{ and } \omega_{i+1} < C_{i+1} \\ \sum_{i=1}^{m} s_{i[E(\omega_i) - C_i]} & \omega_i > C_i \quad \dots \dots \dots \dots \dots \dots \dots . (2) \end{cases}$$

Where,

$$E(\omega_i) = \sum_{\omega_i = c_i + 1}^{c_i + y_i} \omega_i \binom{C_i + y_i}{\omega_i} \beta_i^{\omega_i} (1 - \beta_i)^{(C_i + y_i - \omega_i)}$$

And

$$E(\omega) = \sum_{i=1}^{m} \sum_{\omega_i = c_i + 1}^{c_i + y_i} \omega_i \binom{C_i + y_i}{\omega_i} \beta_i^{\omega_i} (1 - \beta_i)^{(C_i + y_i - \omega_i)}$$

$$E(\omega) = \sum_{i=1}^{m} E(\omega_i)$$

Cost due to loss of customer Goodwill

The model developed in section-1 could easily be modified by replacing the term $p_i \times N$ by n_i. Hence, the model will be:

$$E(L) = \sum_{i=1}^{m} k \{\alpha_i (\beta_i n_i - c_i)\}^2 \dots \dots \dots \dots \dots \dots (3)$$

Lost revenue or spoilage cost

Following similar reason as in the cost due to customer good will, the expected loss of revenue due to empty seat flight could be modeled by replacing the term $p_i \times N$ by n_i. Hence the model will be:

$$L(R) = \sum_{i=1}^{m} \delta_i (1 - \beta_i) \, e_i n_i \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (4)$$

Therefore, the net expected revenue as a function of n_i will be:

Maximize $NER(n_i)$:

$$NER(n_i) = \sum_{i=1}^{m} f_i \times E(n_i) - F(\omega_i, c_i) - \sum_{i=1}^{m} k \{\alpha_i (\beta_i n_i - c_i)\}^2 - \sum_{i=1}^{m} \delta_i (1 - \beta_i) \, e_i n_i$$

3.2.2 Case II: No-show rate follow Generalized Extreme Distribution

In section I, the show-up distribution was assumed to follow binomial. Furthermore, the model developed in that section was prior to collecting the data from the Airline, and as such the input parameters used in developing the model did not reflect the actual data that the airline has. Accordingly, in this section and in the next section a variant of the previous model that captures the nature of the airlines data with minimal input parameters will be developed. In addition, the model developed in this section considers an alternative no-show distribution, the Generalized Extreme value Distribution (GED), which has been found appropriate for describing the no-show distribution of the historical data for Ethiopian.

_____

Notations and Terms

P_i= ticket price for fare-class i

y_i= number of overbooking for fare class i , $Y = \sum_{i=1}^{m} y_i$

x_i = The number of no-shows and cancellations in class i, with p.d.f f(x). (x_i is a r.v.)

S_i = penalty cost of an overbooking corresponding to fare-class i

e_i= the opportunity cost of flying with an empty seat for fare class i

_____

1. Revenue

The revenue generated from the booking (y) passengers in each class could be obtained by multiplying the price of each ticket the overbooking level made in that class.

$$R = \sum_{i=1}^{m} p_i \times y_i \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots . (1)$$

2. Compensation cost

$$C(x,y) = \sum_{i=1}^{m} S_i\ [y_i - x_i] \qquad\qquad \forall y_i > \ x_i$$

Otherwise,

$$C(x,y) = \sum_{i=1}^{m} e_i max\left[y_i - x_i - \sum_{i=i+1}^{m} max(x_i - y_i, 0), 0\right] \ldots \ldots \ldots \ldots \ldots \ldots \ldots . (2)$$

The second term of the above equation implies the fact that, extra arrivals for a seat in one class may be assigned a seat if there is empty seat in another class.

3. Spoilage cost (cost of lost opportunity)

$$L(x,y) = \sum_{i=1}^{m} e_i\ [x_i - y_i] \qquad\qquad \forall y_i < \ x_i$$

Otherwise,

$$L(x,y) = \sum_{i=m}^{1} e_i max\left[x_i - y_i - \sum_{i=m-1}^{1} max(y_i - x_i, 0), 0\right] \qquad \ldots \ldots \ldots (3)$$

Therefore, the net revenue would be modeled as:

38

$$Z(y) = \sum_{i=1}^{m} p_i \times y_i - \begin{cases} \begin{cases} \sum_{i=1}^{m} \max(S_i \left[y_i - x_i - \sum_{i=i+1}^{m} \max(x_i - y_i, 0) \right], 0) \\ \sum_{i=1}^{m} S_i[y_i - x_i] \qquad \text{for} \quad \forall y_i > x_i \\ \sum_{i=m}^{1} \max(e_i \left[x_i - y_i - \sum_{i=m-1}^{1} \max(y_i - x_i, 0) \right], 0) \\ \sum_{i=1}^{m} e_i[x_i - y_i] \qquad for \ \forall x_i > y_i \end{cases} \end{cases}$$

Since the number of no-shows is a continuous variable, the expected net revenue could be rewritten as:

$$E(Z) = \sum_{i=1}^{m} p_i y_i - \sum_{i=1}^{m} max \left(\int_0^{y_i} s_i(y_i - x_i - \left(\sum_{i=i+1}^{m} max \int_{y_i}^{\infty} (x_i - y_i), 0 \right), 0 \right) f(x) dx$$
$$- \sum_{i=m}^{1} max \left(\int_{y_i}^{\infty} e_i(x_i - y_i - \left(\sum_{i=m-1}^{1} max \int_0^{y_i} (y_i - x_i), 0 \right) 0 \right) f(x) dx$$

This model could be easily solved using Mat Lab's numerical integration function instead of using derivatives and iterative solution approaches such as the Quasi-newton method and other derivative based unconstrained optimization algorithms. However, since the closed form equation of the expected revenue, in this case has been found difficult solving using derivatives, a Monte Carlo simulation with a derivative free unconstrained optimization of the Nelder Mead algorithm [14] was adopted. However, assuming that the cost of lost opportunity is the mean value of all the classes, this model could be solved for any number of classes using derivatives. As such, a procedure of solving the model using derivatives will be presented as follows.

In order to simplify the task let us consider minimizing the expected cost instead of maximizing the expected revenue. In this case the model would be reduced in to the following form.

Minimize Z(y):

$$E(Z) = \sum_{i=1}^{m} max \left(\int_{0}^{y_i} s_i(y_i - x_i - \left(\sum_{i=i+1}^{m} max \int_{y_i}^{\infty} (x_i - y_i), 0 \right), 0 \right) f(x) dx$$

$$+ \sum_{i=m}^{1} max \left(\int_{y_i}^{\infty} e_i(x_i - y_i - \left(\sum_{i=m-1}^{1} max \int_{0}^{y_i} (y_i - x_i), 0 \right) 0 \right) f(x) dx$$

Remember that the above model is based on the assumption that higher fare class contenders could not be bumped into lower fare class seats, or in other words downgrading is not allowed. However, this is a hypothetical case which does not have any application in reality (at least at Ethiopian). In fact, it is strictly forbidden to bump cloud nine passengers into the economy class in any flight. Therefore, there is the freedom to bump economy class of those extra high fare passenger show-ups into a low fare seat in the economy or a business class seat if there is any empty seat there. Considering the practical situation, that is downgrading within the economy class and upgrading economy class passengers into cloud nine seats, the model will further be restructured and simplified as follows.

$$E(Z) = \sum_{i=1}^{m} \int_{0}^{y_i} s_i(y_i - x_i - max \sum_{i=i+1}^{m} \int_{y_i}^{\infty} (x_i - y_i), 0) f(x) dx$$

$$+ \sum_{i=1}^{m} \int_{y_i}^{\infty} e_i(x_i - y_i - max \sum_{i=i+1}^{m} \int_{0}^{y_i} (y_i - x_i), 0) f(x) dx$$

Now, let us consider the two fare class scenario in the economy and the no-overbooking rule in the cloud nine case. In order to accommodate the number of no-shows in the cloud nine, it should be overbooked in the economy without violating the rule of no-overbooking in cloud nine. To do so, consider the whole sit as if it is an economy class seat and then make the overbooking, finally set aside the number of business class seats not overbooked.

Let y_1, y_2 be the overbooking levels in fare class-1 and fare class-2 respectively

x_1, x_2 Represent the number of no-shows in fare class-1 and fare class-2

e_1, e_2 Represent the cost of lost opportunity.

S_1, S_2 Represent the compensation cost of each class. In practice a linear compensation plan is used, and hence the compensation cost for all denied passengers will be the same regardless of which class they belong.

Now let us assume that the cost of lost opportunity for each class be the average of the individual fare classes. The weighted mean of the cost of lost opportunity will be considered in order to minimize the error that could be introduced as a result of the assumption made. Hence, the weighted mean for the cost of lost opportunity could be obtained as:

$$e = \frac{e_1 r_2 + e_2 r_2}{r_1 + r_2} \qquad where\ r\ is\ the\ no-show\ rate$$

This approximation of the cost of lost opportunity greatly simplifies our objective function into a form that finally would give a closed form solution approach. Using the weighted cost of lost opportunity makes all the seats as having the same value, and consequently reducing the objective function into a single variable minimization problem. Since the objective function has been solved without making such a simplifying assumption it would be good to compare the solutions and other measuring parameters of the two approaches. With our assumption, the no-shows in each class could sum up without any multiplying factor (since all the seats are having the same value) as in the overbooking case; the sum of each variable could be reduced into a single variable of overbooking and no-show.

Hence, $y = y_1 + y_2$ will be the total overbooking

Likewise, $x = x_1 + x_2$ will be the overall no-show

Therefore, the reduced form of the objective function would be:

$$E(Z) = S \int_0^y (y - x)f(x)d(x) + e \int_y^\infty (x - y)f(x)d(x)$$

41

Solving this equation would be straightforward using Leibniz Rule, which provides a means of differentiation under the integral. The Leibniz Rule says that if we have an integral of the form []:

$$\int_{y_0}^{y_1} f(x,y)d(y)$$

Then for $x \in (x_0, x_1)$ the derivative of this integral is thus expressible

$$\frac{d}{dx}\int_{y_0}^{y_1} f(x,y)d(y) = \int_{y_0}^{y_1} \frac{\partial}{\partial x}f(x,y)d(y)$$

Provided that both f and $\frac{\partial f}{\partial x}$ are continuous over the region $[x_0, x_1] \times [y_0, y_1]$

Making use of Leibniz Integral Rule, the objective function could be minimized at a relative ease.

Taking the first order derivative on both sides of the objective function:

$$\frac{dE(Z)}{dy} = S\int_0^y \frac{\partial}{\partial y}(y-x)f(x)d(x) + e\int_y^\infty \frac{\partial}{\partial y}(x-y)f(x)d(x)$$

$$= S\int_0^y f(x)d(x) - e\int_y^\infty f(x)d(x)$$

$$= SF(y) - (e - eF(y))$$

$$where\ F(y)\ is\ the\ cumulative\ distribution function\ of\ the\ no-show$$

Now set the first derivative to zero to find the closed form expression for the overbooking level.

$$SF(y) - (e - eF(y)) = 0$$

$$(S + e)F(y) = e$$

Solving this for F(y) would give us:

$$F(y) = \frac{e}{e+s}$$

42

Now let us check the convexity of the objective function by taking the second derivative of the objective function.

$$\frac{dE(Z)}{dy} = SF(y) - e\big(1 - F(y)\big)$$

$$\frac{d^2E(Z)}{dy^2} = (S + e)f(y) > 0$$

Since the second derivative of the function is nonnegative, our objective function is convex. The value of $F(y)$ is the probability that the number of no-shows will not exceed y for the given values of the compensation and cost of lost opportunity.

The expression for $F(y)$ under the assumption of the weighted mean for the cost of lost opportunity is the same as that of the News Boy problem. This expression will be used to find the overbooking level and will be compared with the results of the derivative free approach to investigate its applicability for the case considered.

3.2.2 Solution approach using Monte Carlo Simulation

The above stochastic model is solved by using a derivative free optimization algorithms (both the Nelder mead and Genetic algorithms were used) in order to eliminate some of the drawbacks of using the derivative based solution approach. The Monte Carlo simulation does not only, eliminate the assumption of making the whole seat as if they have identical values, but also has the flexibility to run it for a variety of probability loss values as required by the decision maker. The simulation approach can be used for any number of fare-classes that the airline may have. This would in effect make the Mote Carlo simulation approach an advantage over the closed form equation.

The objective is to find the optimal overbooking level (y_i) that maximizes revenue.

Revenue

$$R(y) = \sum_{i=1}^{m} P_i \times y_i \ldots \ldots \ldots \ldots \ldots \ldots \ldots (1)$$

Compensation cost

$$C(x,y) = \sum_{i=1}^{m} S_i \ [y_i - x_i] \qquad\qquad \forall y_i > \ x_i$$

Otherwise,

$$C(x,y) = \sum_{i=1}^{m} e_i max \left[y_i - x_i - \sum_{i=i+1}^{m} max(x_i - y_i, 0), 0) \right]$$

Lost Revenue (cost of lost opportunity)

$$L(x,y) = \sum_{i=1}^{m} e_i \ [x_i - y_i] \qquad\qquad \forall x_i > \ y_i$$

Otherwise,

$$L(x,y) = \sum_{i=1}^{m} e_i max \left[x_i - y_i - \sum_{i=m-1}^{1} max(y_i - x_i, 0), 0) \right]$$

Objective:

Maximize:

Net Revenue $= R(y) - C(x, y) - L(x, y)$

This model could be modified to accept a user defined constraint function. As an example, the probability of loss could be considered as a constraint and the above unconstrained maximization problem would be transformed into a constrained maximization problem as shown below:

Maximize:

Net Revenue $= R(y) - C(x, y) - L(x, y)$

Subject to:

$$probability(mean\ net\ revenue < 0) <= v$$

Where, v is the user defined value for the probability of loss. This is also one of the advantages of this model over previous models which lack this flexibility.

3.3   Model Characterization and identification

The model, which developed based on the extreme value distribution for the no-show data, is to be identified as unconstrained non-linear programming (UNLP). For a large number of fare ticket classes and variable demand rates solving the model using derivatives could be extremely difficult and take a considerable amount of time. However, the model could be simplified with some reasonable approximations as explained in the previous section. Furthermore, with the current data management system and relatively low variation between fare classes in use at Ethiopian, the model could be solved in relatively minimal time. Again, for large values of the authorization level, and fare classes the simulation approach proposed for solving this model requires a much less amount of time as compared to the derivative based solution approaches. The Monte Carlo simulation approach uses direct search algorithms as the solution approach. More specifically, the Nelder Mead direct

45

search algorithm of Mat Lab was used for solving the models, though this algorithm does not guarantee optimality.

3.4   Sample Data

For the purpose of this study in verifying and measuring the performance of the proposed model, a historical data of booking, no-shows, and cancellation was collected. An 18 months (six months from each of the year of 2008/09/10) data was collected for the purpose of fitting the data in to a probability density function (PDF). An out bound station with a daily flight (ADD-DXB) and another station with a lower load factor as compared to other stations (due to no-shows, ADD-CAI) were chosen for the analysis of the data. Then, the six months data of no-shows and rate of no-shows from each year were fit separately in to a PDF.

Since the number of bookings for each day differs, first the rate of no-shows was fitted to see the probability density function (PDF) of the smoothed variable. Then, the no-show data was fitted without considering the variation in the number of bookings, to see if there could be a significant difference in the PDF of the two variable fits. For the flight destinations in our case example it was found that both the rate of no-show and the no-show data's PDF follow the same distribution. A closer look at the number of bookings, no-shows as well as the load factor of Ethiopian airlines shows that Ethiopian has insignificant number of denied boarding (one per twenty thousand). However, this could be the case not only because of the low overbooking level but sometimes demand goes below the capacity. When it is the case that demands are expected to be lower than the available seat capacity, a competitive air fare structure should be used in order to attract potential customers. Ethiopian has affixed fare structure from which a customer could choose, and this fare structure is calculated mainly based on the minimum number of load factors forecasted so that the airline operates with an anticipated profit even if it is flying with a lot of empty seats.

The statistics toolbox of Mat Lab was also utilized in checking the distribution of the historical data after a general distribution fit comparisons were made on the 'EasyFit' software. The 'EasyFit' software [36] is helpful in generating the best distribution fit appropriate for our data. However, I also used the manual fitting toolbox in Mat Lab to check if the results from the 'EasyFit' software are acceptable. The EasyFit

software is commercial software developed by 'mathwave data analysis and simulation' project, and a trial version of it could be downloaded for a one month use. For the case of Ethiopian (with respect to the data at hand), the assumption that the no-show and cancellation data follow beta, normal or gamma distribution is not applicable even though it might be the case for other airlines as pointed out in the literature review. A detailed comparison of the fit based on three (Kolmogorov-smirnov, Anderson-Darling, and chi-squared) goodness of fit (gof) test shows that, the generalized extreme value distribution is the best fit distribution for our no-show and cancellation data. This was also approved by manually fitting the data in Mata Lab's 'dfittool'. An example showing the fitted data for ADD-BXD and ADD-CAI is give below along with the test statistic of the chosen gof test.
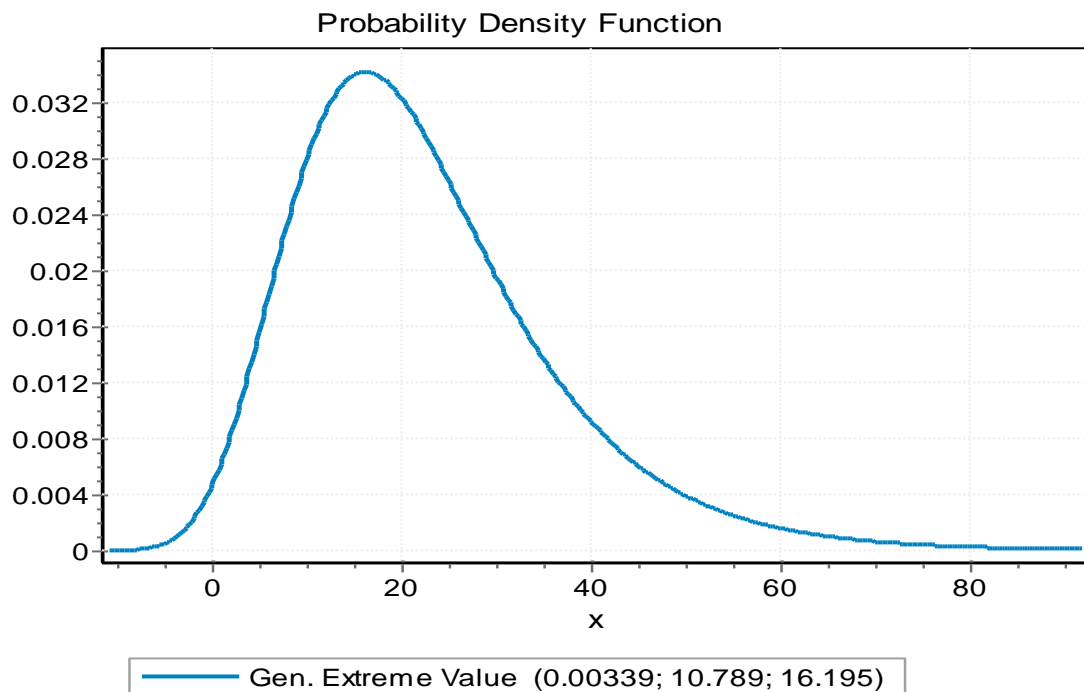

Figure 4: pdf of no-shows for ADD-DBX

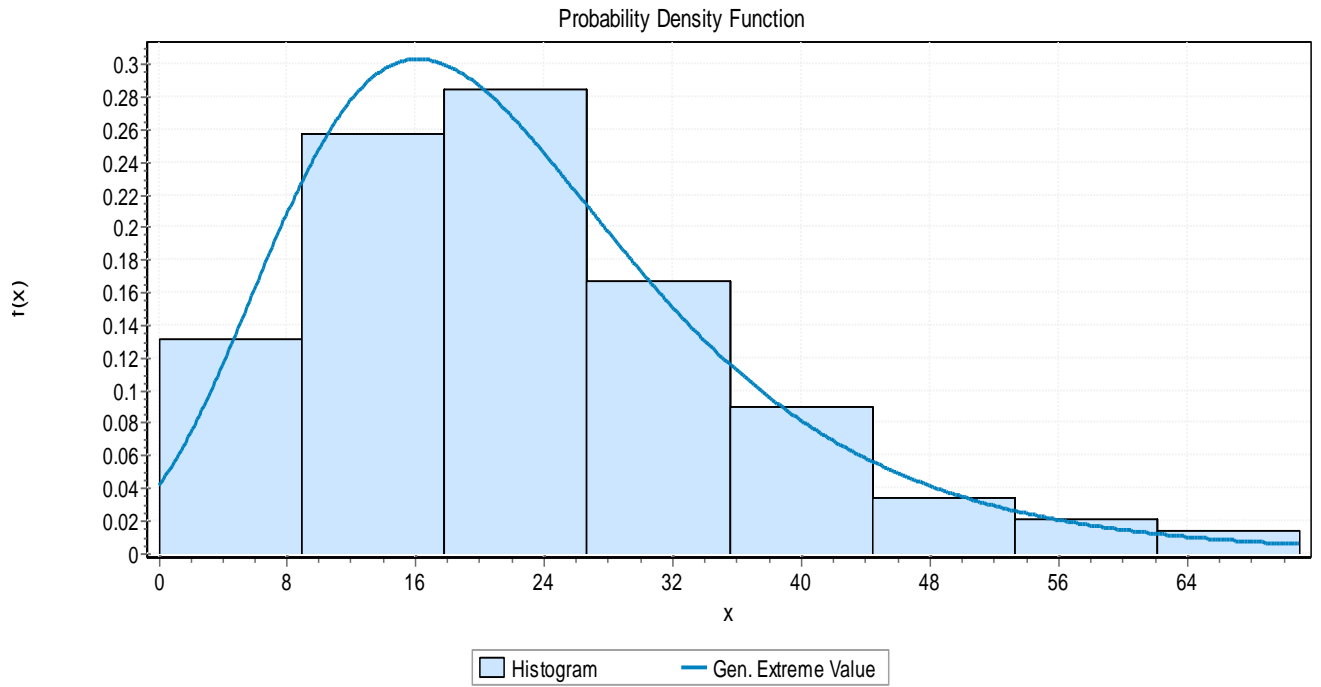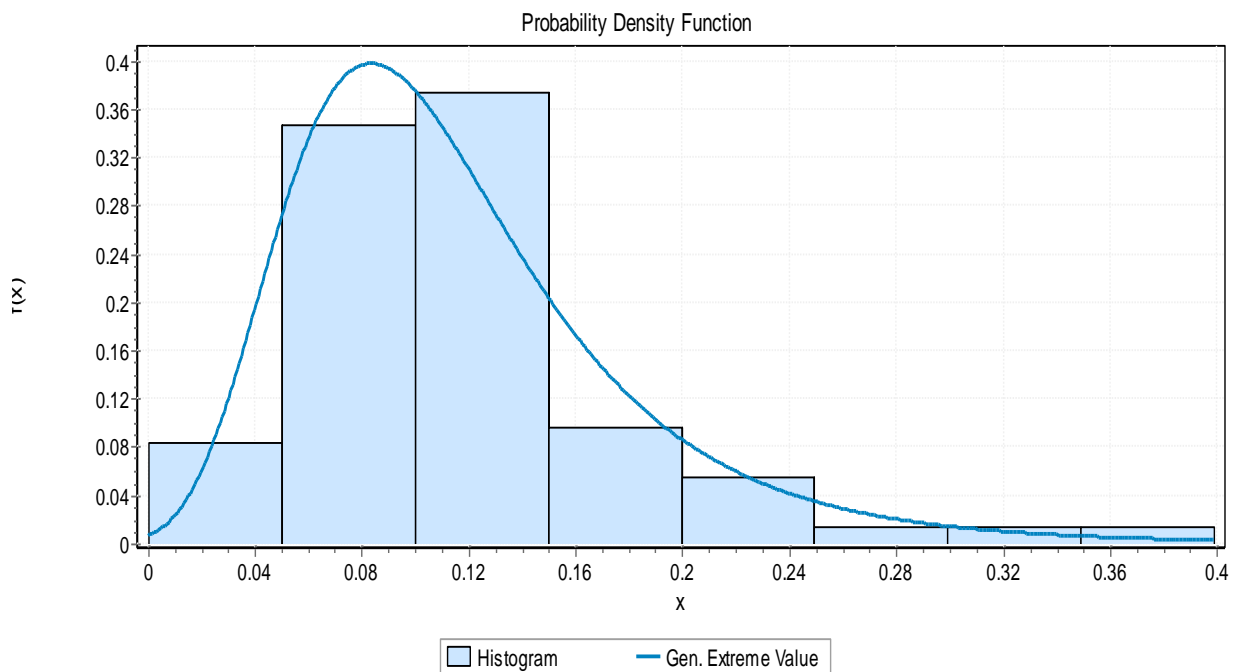**Figure 5: a histogram of no-shows fitted to pdf of GED for ADD-DBX**

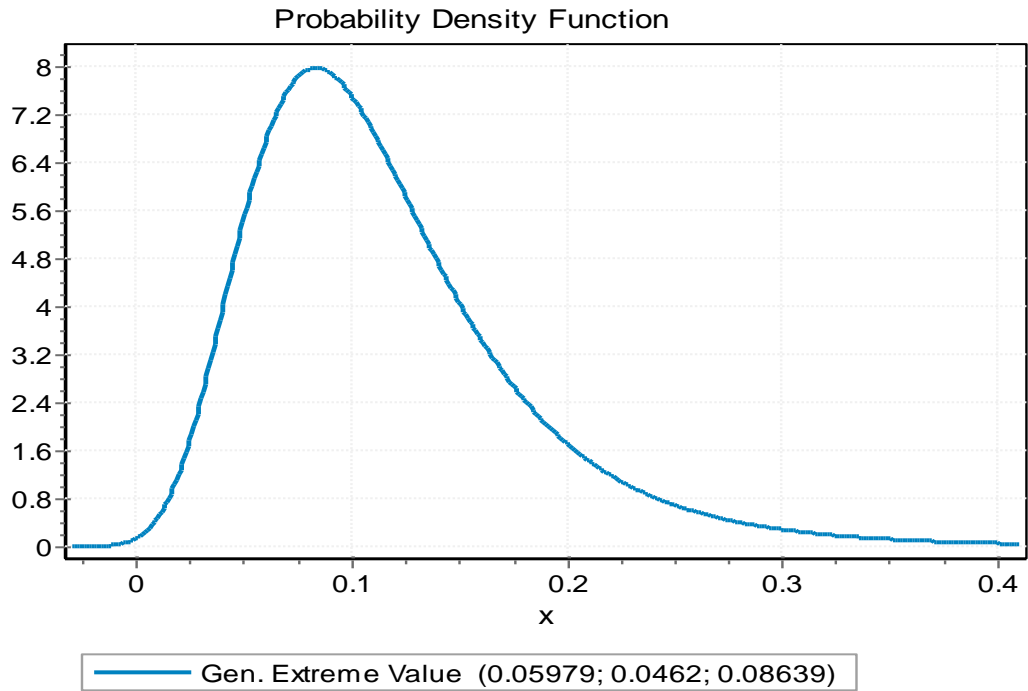**Figure 6: a histogram of no-show rates fitted to pdf of GED for ADD-DBX**

48

Figure 7: pdf of GED for ADD-DBX no-show rate data

The following pdf shows the no-show rate and no-show values fitted for ADD-CAI route respectively.
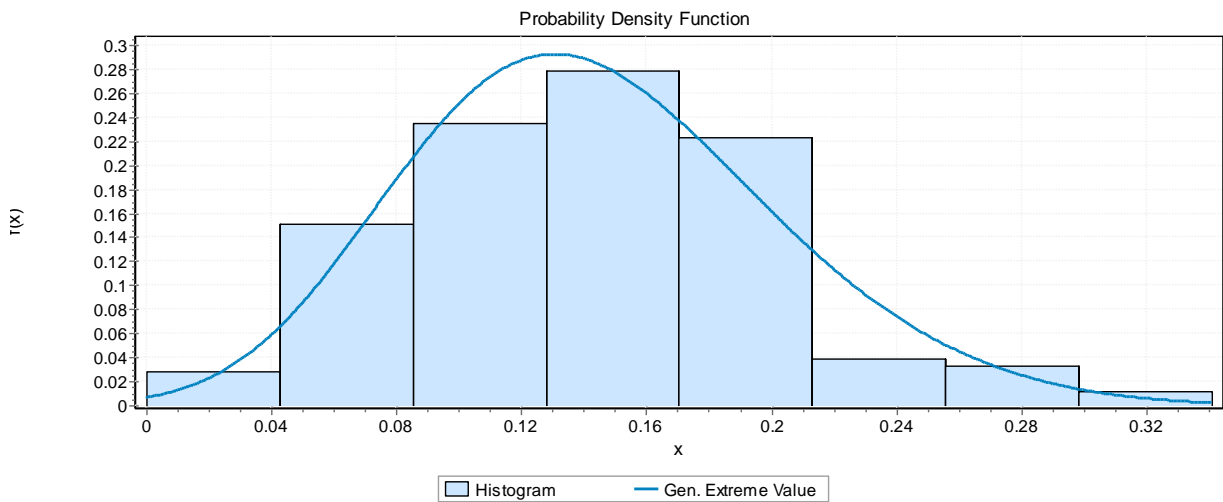

Figure 8: a histogram of no-show rates fitted to pdf of GED for ADD-CAI
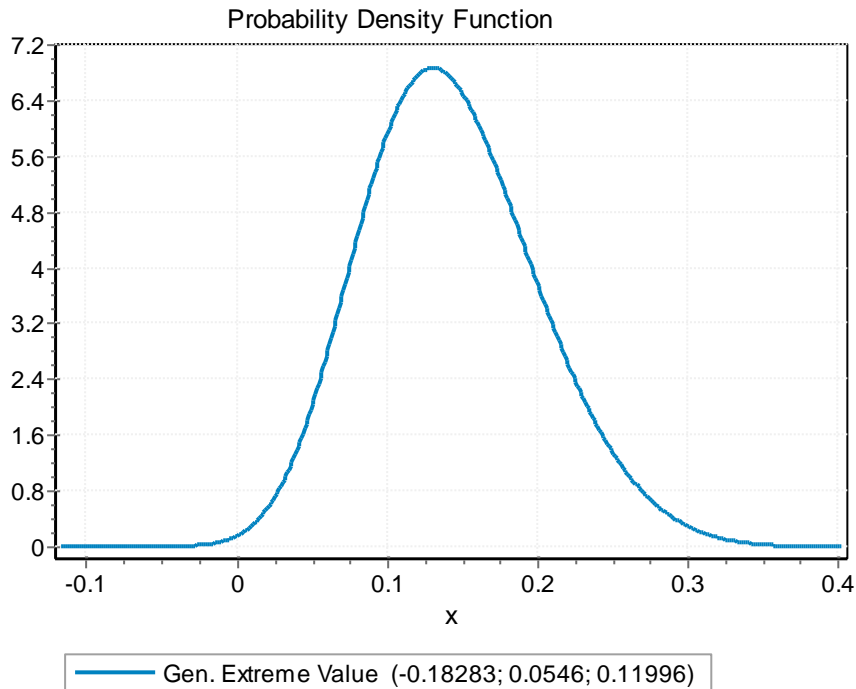
Figure 9:   pdf of GED for ADD-DBX no-show rate

The pdf for the no-show data best fit might not be the general extreme value distribution in some cases. However, since the no-show data does not take the number of booked passengers in to account, the no-show rate best fit describes the nature of the no-show under a smoothed booking values. Therefore, the no-show rate has been found consistently to best fit the Gen. Extreme value distribution for the data tested for our case.

3.5   Solution Approach: using MatLab's optimization toolbox

Mat Lab (matrix laboratory) is a high performance numerical computing programming language developed by MathWorks [37]. Mat Lab integrates visualization, programming, and computation in one environment that made it user friendly and accessible by all professionals in the computing arena. Mat Lab, in addition to its programming environment, also has a variety of toolboxes designed to suite different users. Among the toolboxes available in Mat Lab, the optimization toolbox and the statistics toolbox are extensively used in this thesis. Since Mat Lab incorporates all the known optimization algorithms in its optimization toolbox, this thesis largely depends on the optimization toolbox to solve the unconstrained nonlinear model

50

presented above. Therefore, all the programs required to solve both the models developed (i.e. the expected value model and the simulation model) are programmed in Mat Lab codes and use Mat Lab toolboxes.

The optimization toolbox in MatLab has a built in algorithm for unconstrained nonlinear minimization, which is used to solve our model developed based on the expected value approach. The minimization algorithms built in could solve the model either a derivative free search approach or using derivatives approximated by the solver or derivatives supplied by the user. In this paper, 'fminsearch' which uses a derivative free optimization algorithm is selected as the solution approach in solving our model.

4. Computational Results

The computational results of all the variant models considered will be presented in the following section. A personal computer with capacity of 2.67GHz Intel i5 Core processor and 4GB of RAM was used. The codes for all the models were implemented in MATLAB 7.7.0 running under windows 7.

4.1 No Overbooking

As pointed out above, the loss due to the no-shows for Ethiopian is substantially significant that Ethiopian could not simply see overbooking as an option; rather it should see it as critical for increasing its revenue. In order to see the significance of overbooking let us consider the flight ET452 (this flight is B737-700 with passenger seat capacity of 118). The following data are used in estimating the cost of lost opportunity if overbooking is not to be used. However, since calculating the break-even load factor for each flight (considering the available data) is impossible, the average break-even load factor of the airline has been taken in the following calculation. The break-even load factor is the load factor that the airline should perform generating revenue that equates its expenses. Compared to other airlines in the region and airlines in the west, Ethiopian could be considered as having the least break-even load factor, which could be attributed to the cheap labor cost in Ethiopia than in other countries.

Seat Capacity=118

Break-even Load Factor=52%

Average No-show rate=14.6%

Average Ticket fare=$350

Penalty cost of cancellations and no-shows=$100

Since the break-even Load Factor is 52%, the number of passengers for break-even point (N_b) for flight ET452 will be:

$$N_b = 0.52 * 118 = 62$$

That is, at least 62 passengers are required to cover the capital cost, operating cost of the flight, and costs associated with loading, unloading, and placing a passenger on board. At the break-even point, the airline is operating with no profit.

For this particular flight the average show-up rate is 85.4%, and the maximum number of booked passengers that could be realized is 118, which should be equal to the capacity of the aircraft (since we are assuming no overbooking for this case). Hence, the number of passengers who will arrive at the gate for boarding would be obtained as:

$$N_a = 0.854 * 118 = 101 \ passengers$$

Since N_b covers all the operating costs, the number of passengers beyond the break-even point who board for the flight could be assumed to generate the profit. Hence, assuming $350 profit to be generated from the 63rd passengers onwards, the total profit would be:

$$Total \ profit = (118 - 101) * \$100 + (101 - 62)*\$350 = \underline{\$15,350} \ \text{per flight}$$

Though this profit is sizable, Ethiopian could have generated more profit had it been using overbooking in its practice in order to account the 17 no-shows. The average lost revenue or cost of lost opportunity in this case would be:

Cost of lost opportunity=17*350=$\underline{\$5,950}$ per flight

Therefore, in the case of no-overbooking, Ethiopian could operate the ET452 on only 72.06% of its capacity of generating profit per flight on average. Considering the above analysis and the gross number of no-shows in charts above, it is clear that overbooking is a critical practice if Ethiopian would like to boost its profit and increase customer satisfaction. At this point it is important to explain how overbooking (to some extent) could increase customer satisfaction. Customers denied booking may become frustrated with the airlines capacity of handling the market, and as a result they would show interest in using other air carriers in the future. With overbooking a substantial number of customers (for example, on average 17 customers with the ET 452 flight) could have been accommodated, resulting in an increase in the market share for current and future flights. However,

an uninformed number of overbooking would also result in customer dissatisfaction and consequently with a potential loss of future market share.

4.2 Using proposed Overbooking Models

4.2.1 Case-I: Using a closed form expression for overbooking

From the fitted data we have the following parameters of the distribution, which could be used to generate the cumulative distribution. The parameters and the distribution used in estimating the level of overbooking are given below.
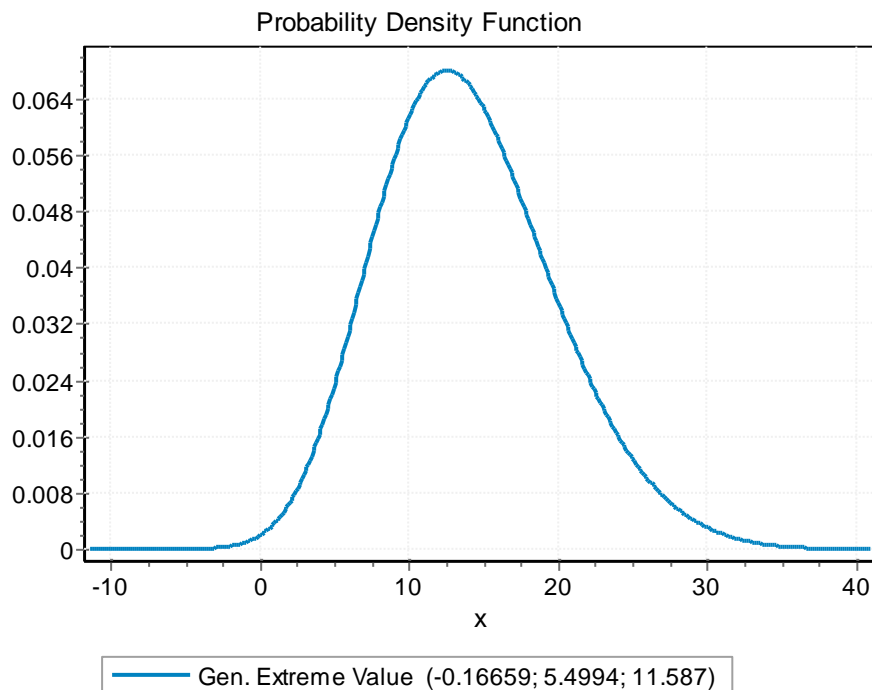
Figure 10:   pdf of GED for ADD-CAI no-show data
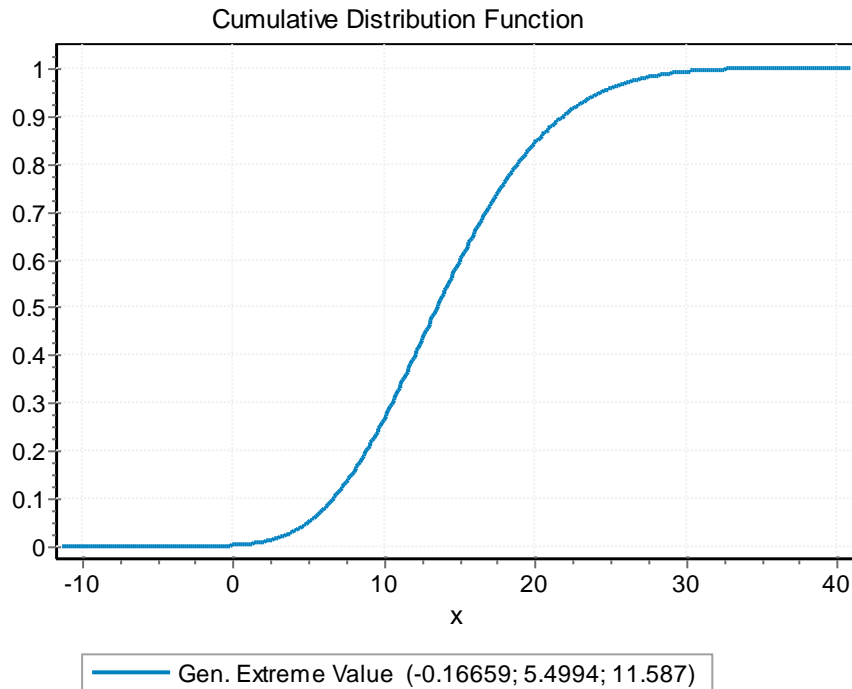
Figure 11: cdf of GED for ADD-DBX no-show data

Now let us determine the overbooking level for different possible linear compensation plans using the cumulative distribution function given above and the value of the critical ratio.

The weighted cost of lost opportunity (e) could be calculated as:

$$e = r_1 e_1 + r_2 e_2$$

$$e = 0.4 * 300 + 0.6 * 400$$

$$e = 360$$

Compensation Cost per passenger	F(y)	Approximate # of Optimal Overbooking
$150	$\frac{360}{150 + 360} = 0.71$	17
$300	0.55	15
$400	0.47	13
$500	0.42	13

55

$600	0.38	12
$700	0.34	11
$800	0.31	11

Table 2: Optimal overbooking numbers for different compensation plans using the closed form solution approach

As we can see from the table, for a compensation plan less than the price of the ticket we have a maximum overbooking level of 17. However, the overbooking level for a compensation plan less than the price should in theory be at least the maximum number of no-shows observed frequently. Strictly speaking, the number of overbooking could be set to infinity for this case as long as the profit is increasing per overbooked passenger. This clearly shows us that, the News Boy Model cannot efficiently optimize the seat overbooking problem. However, as infinite overbooking is impractical, a practical choice of the overbooking would be to set the maximum no-show that has been observed frequently. In addition to failing to capture at minimum compensation plans, the closed form of overbooking model also does not perform well at a large value of the compensation. When the compensation cost increases, in theory we expect the overbooking level to decrease dramatically. Why did the model if constructed correctly could not capture the phenomena that everyone should expect? Two reasons might be proposed for the discrepancy between theoretical results and the result of the closed form solution. First, modeling the problem as minimizing the cost will eliminate the term that generates the profit and consequently a distorted value of the overbooking level will be observed. The second reason is that the closed form solution is set up in such a way that it could determine the probability that the no-show will not exceed a certain value, and hence cannot determine the overbooking value that could maximize the net revenue. However, if the airline wants to be more sensitive to customer reaction upon denied boarding and the consequences of loss of customer good will, the closed form solution approach for determining the number of optimal overbooking would be more appropriate as it determines the maximum number of no-shows that would be expected with a certain level of confidence.

4.2.2. Case-II: Monte-Carlo simulation approach

The following chart shows the data used in determining the level of overbooking in each class using the simulation approach.

Ethiopian Airlines cost data for flight ET-452

	Cloud Nine	Economy
Boeing 737-800 capacity	16	102
Fare	$728	$299,$399
Compensation cost	N/a	$150-$598,$798
Mean no-show rate	3.5%	11%

Ethiopian did not overbook in its 'cloud nine' class though the no-show rate is significant and the associated lost revenue or cost of lost opportunity is huge. Since there is no overbooking in cloud nine, the compensation cost for this particular service is set to be zero or not applicable. Even though it is recommended not to overbook at 'cloud nine' as these Business class customers are highly sensitive if denied boarding, it is reasonable to consider the no-show rate and add this rate in to the economy class so that the overbooking will be made in the economy anticipating no-shows in the cloud nine. In this case, the customers to be overbooked are the economy class customers anticipating no-shows in the cloud nine. With this in mind, the simulation model first generates the combination of all possible overbooking levels in the economy class, and calculates the corresponding revenue to be generated at each overbooking level. This value is simulated 10,000 times, and then the mean of the revenue at each combinations of overbooking levels is obtained. Finally a plot of the revenue and its corresponding overbooking level is drawn to see where the maximum revenue lies. The following table shows the possible combination of overbooking levels for two fare class case and the average revenue generated using the above data as well as the probability of loss. The probability of loss is the probability that the revenue generated would have a negative value.

Compensation cost	Optimal # of Overbooking		Current revenue generated	Expected revenue generated	Probability of Loss
	Y_1	Y_2			
150	20	20	9.64E+03	9.93E+03	0
300	20	20	5.60E+03	5.50E+03	0
400	12	20	2.99E+03	2.55E+03	0.1966
500	7	14	1.50E+03	-4.63E+02	0.586
600	6	11	5.69E+02	-3.40E+03	0.7
700	5	9	-1.19E+02	-1.00E+03	0.887

Table 3: Optimal overbooking levels for different compensation plans for a two fare class using the Monte Carlo simulation approach

As one can observe from the table as the compensation cost increases the level of overbooking decreases, and that is in line with our intuitive expectations. For compensation plans less than the price of the ticket and the cost of lost opportunity, the Monte Carlo simulation solution approach gives us a very huge number for overbooking, and that is true theoretically at least from a mathematical (economic) point of view. However, such a huge number for overbooking is not practical and does not reflect reality. Being the case, even if the model recommends huge number of overbooking for small amount of compensation plans, the airline should limit the overbooking level to the maximum no-show observed in the past. Hence, the maximum of the no-show is used as the optimal overbooking for this case.

For the two fare class scenario the following figure presents the mean net revenue values against the overbooking numbers to be made in each class. These plots are generated for a compensation cost plan of five hundred per passenger.
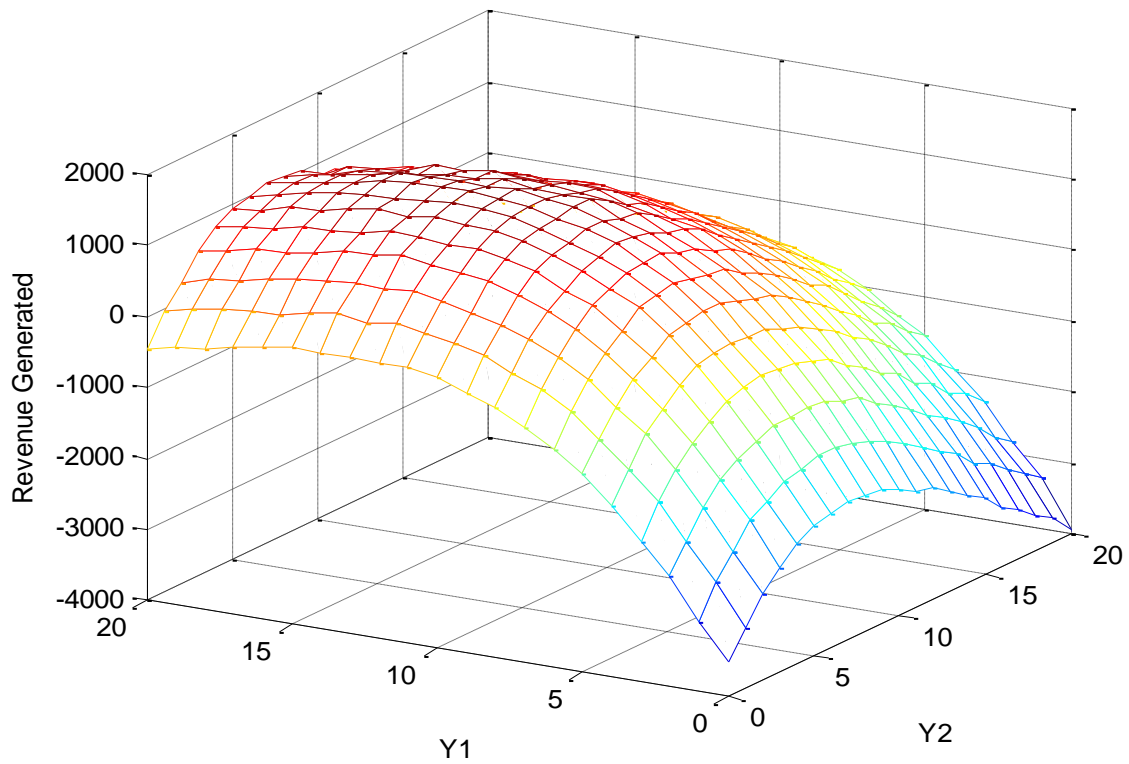
Figure 12: Plot of Revenue generated vs. overbooking for two class case

The following figure shows the probability of loss for a compensation plan of five hundred per denied boarding.
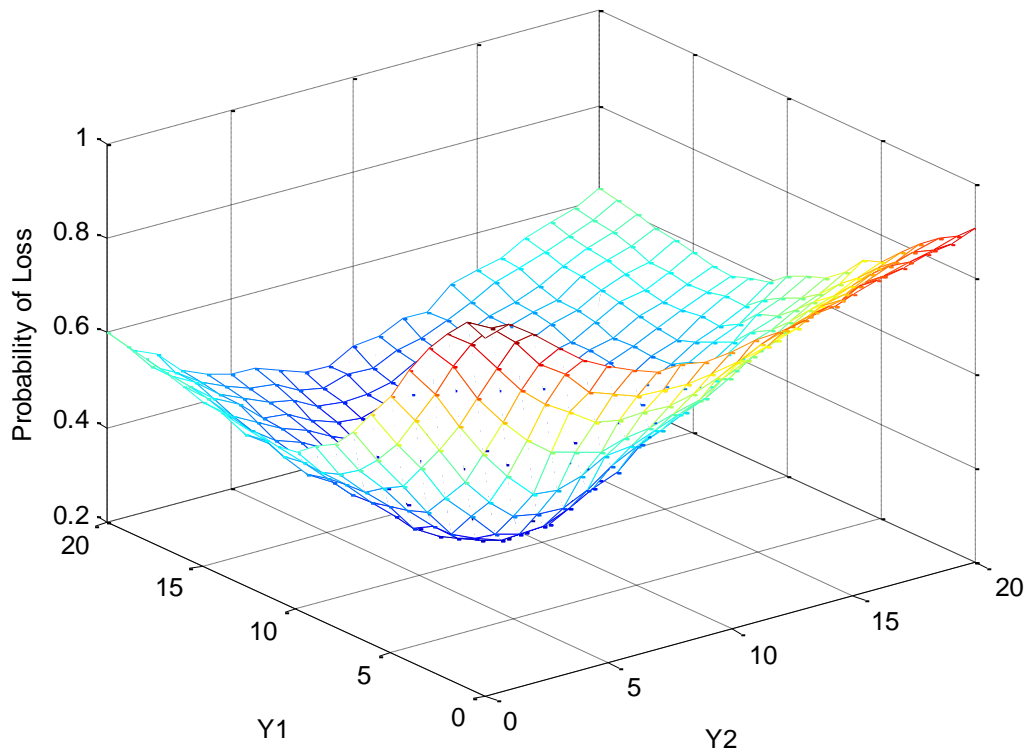
Figure 13: Plot of Probability of loss vs. overbooking for two class case

Looking into the probability of loss plot gives an insight of the critical values of overbooking where potential loss could occur. In the case example, the overbooking level for class-1 is near ten while for class-2 is greater than ten. This once more confirms our simulation result given above as seven for class-1 and fourteen for class-2 with probability of loss around 0.586. One of the advantages of this graph is that it does not only show the optimal values but also gives the decision maker an insight how and what level of overbooking in each class could affect the probability of loss in revenue. Furthermore, the graph gives the decision maker the freedom of relaxing the overbooking value by a certain amount as long the probability of loss is acceptable.

The compensation plan used in the above calculation is based on the denied boarding and overbooking regulation of Ethiopian. However, in practice the compensation plan offered may not be the same as what is stated in the denied

boarding regulation. The compensation plan offered to denied boarding customers may vary from a minimum value of $150, which is less than the price of the seat ticket for our considered flight for any class, to a maximum value of twice the ticket fare (mostly in case of flights to Europe and the Americas). For the currently in use compensation plan the overbooking level would be greater than the overbooking level obtained above. The following simulation result obtained by using a linear compensation plan of $150 shows that the overbooking level required to be made for the same flight considered above.

4.3   Optimal Overbooking Analysis for Multi fare-class problem

The data for cancellation and no-shows was first subdivided into different number of fare classes (in this case a maximum number of seven classes were considered). Then, this new data derived from the aggregate no-show and cancellation data was used to determine the optimal number of overbooking in each fare class. A detailed table of all the values for the cases considered is given in the appendix. However, the comparison of values (in terms of revenue generated, and the probability of generating a negative value) are considered and are given the following table.

| Compensation Cost | optimal number of overbooking | | | | | | |
	Single fare class	Double fare class	3 fare classes	4 fare classes	5 fare classes	6 fare classes	7 fare classes
150	17	40	40	40	40	40	40
300	15	40	36	34	33	31	30
400	13	33	28	27	25	23	23
500	13	21	21	19	19	16	16
600	12	17	16	16	14	13	13
700	11	14	14	13	12	11	9

Table 4: comparison of number of optimal overbooking for different fare classes

The optimal number of overbooking decreases as the number of fare class increase. This could be because of the fact that the cumulative error introduced as a result if increased number of classes. The same effect has been observed for the revenue generated as the number of fare-classes increases as shown in the following table.

Compensation Cost	Revenue Generated for the above Overbooking values					
	Double fare class	**3 fare classes**	**4 fare classes**	**5 fare classes**	**6 fare classes**	**7 fare classes**
150	9.64E+03	7.25E+03	7.88E+03	7.48E+03	6.68E+03	6.58E+03
300	5.60E+03	4.60E+03	3.85E+03	3.60E+03	2.87E+03	2.80E+03
400	2.99E+03	2.60E+03	1.99E+03	1.83E+03	1.34E+03	1.26E+03
500	1.50E+03	1.20E+03	7.53E+02	6.36E+02	4.62E+02	2.94E+02
600	5.69E+02	2.69E+02	3.91E+02	1.98E+02	-4.14E+02	-8.30E+01
700	-1.19E+02	2.56E+02	-6.45E+02	-7.04E+02	-8.83E+02	-9.58E+02

Table 5: comparison of revenue generated for different fare classes

5. Conclusions

In this thesis, it has been attempted to study, model, and optimize the airline overbooking problem with Ethiopian airlines data. Two approaches for solving the model are proposed and their advantages and disadvantages identified. The closed form solution, which is based on the News boy Model, could be effective relative to the Monte Carlo simulation in minimizing the number of denied boarding. However, this approach did fail in attaining the maximum revenue as it is based on minimizing the costs involved in the model. In contrast, the Monte Carlo simulation solution approach of the derivative free optimization using Nelder Mead was observed as maximizing the expected revenue generated from overbooking. Furthermore, this solution approach has the advantage of specifying the overbooking limits for each class with an estimated probability of loss while the News Boy model couldn't. Moreover, the simulation approach can handle problems of any classes of size having different compensation and losses of revenue, which the closed form solution could not handle. However, the derivative free optimization approach adopted and proposed in solving the model in this study does not guarantee optimality.

Genetic Algorithms were used to solve the model and verify the results obtained from the Nelder Mead algorithm. It was found that the solutions of both methods are fairly close that the Nelder Mead could be used independently. Furthermore, the time needed by the Nelder Mead algorithm is slightly shorter than that of the GA algorithm.

The model developed and proposed in this study uses less input data as compared to other models in the literature, which is an advantage for the airline that usually has difficulty in collecting and organizing its data.

Based on the analysis of the booking data for Ethiopian airlines, the distribution of the no-show has been found best to fit the Generalized Extreme Distribution as opposed to the commonly assumed normal distribution.

Generally based on the findings of the study the following conclusion could be drawn

- ➢ Currently Ethiopian flights operate at an average Load factor of 72%, which costs Ethiopian a huge sum of money per flight.
- ➢ When the number of fare-classes increases beyond three, it has been observed an insignificant change in the number of overbookings, while the revenue generated decreases dramatically.
- ➢ The probability of loss (generating negative revenue) increases as the number of fare classes considered increases.

6. Future Work

This study focuses on developing the static overbooking model based on a data obtained from Ethiopian airlines. In doing so, a one way trip was considered in both the development and the analysis of the model. However, a round trip booking is common practice among customers and this should be considered as future work in extending and improving the performance of the static overbooking model.

Furthermore, the static overbooking problem considered here was treated as if it is independent of other activities of the revenue management system such as pricing and seat inventory control. Integrating the overbooking model into the other two major airline revenue management problems would be an interesting future work.

The cost of loss of customer goodwill is an important factor that should be considered in the overbooking model especially when the competition in the market is becoming fierce. The recommendation to test the validity of modeling the cost of loss of customer goodwill using the Taguachi quality loss function could be an extension of this work.

Developing an algorithm that could guarantee optimality of the overbooking problem is another task for the interested researcher.

Considering the number of no-shows and cancellation resulting in empty seat flight at Ethiopian, one can extend to use the static overbooking model in determining the minimum ticket price that could be offered without loss. Hence, developing a flexible or negotiable pricing system (model) to some of the seats could be a future work.

Appendix

Compensation	Optimal # of overbooking			Current revenue generated	Expected revenue generated	Probability of loss
	Y_1	Y_2	Y_3			
150	10	15	15	7.25E+03	5.37E+03	0
300	6	15	15	4.60E+03	3.50E+03	0.00007
400	4	9	15	2.60E+03	1.56E+03	0.235
500	3	7	11	1.20E+03	-351	0.5623
600	2	6	8	2.69E+02	-2.30E+03	0.76
700	2	5	7	2.56E+02	-4.15E+03	0.8584

Table 6: Optimal overbooking levels for different compensation plans for a three fare class

Compensation cost	Optimal # of overbooking				Current revenue generated	Expected revenue generated	Probability of Loss
	Y_1	Y_2	Y_3	Y_4			
150	8	8	12	12	7.88E+3	3.67E+3	0
300	5	5	12	12	3.85E+03	2.40E+03	0.025
400	4	4	7	12	1.99E+03	1.10E+03	0.22
500	3	3	5	8	753	1.14E+02	0.476
600	2	2	5	7	391	-9.24E+02	0.67
700	2	2	4	5	-645	-8.30E+03	0.7911

Table 7: Optimal overbooking levels for different compensation plans for a four fare class

Compensation cost	Optimal # of overbooking					Current revenue generated	Expected revenue generated	Probability of loss
	Y_1	Y_2	Y_3	Y_4	Y_5			
150	4	8	8	8	12	7.48E+03	3.80E+03	0
300	2	5	6	8	12	3.60E+03	2.11E+03	0.021
400	1	3	4	5	12	1.83E+03	1.00E+03	0.24
500	1	3	3	4	8	6.36E+02	-1.10E+02	0.53
600	1	2	2	3	6	1.98E+02	-1.20E+03	0.72
700	1	2	2	2	5	-7.04E+02	-1.30E+04	0.83

Table 8: Optimal overbooking levels for different compensation plans for a five fare class

Compensation cost	Optimal # of overbooking						Current revenue generated	Expected revenue generated	Probability of Loss
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6			
150	4	4	8	8	8	8	6.68E+03	4.00E+03	0
300	2	2	5	6	8	8	2.87E+03	2.20E+03	0.0075
400	1	1	4	4	5	8	1.34E+03	9.93E+02	0.243
500	1	1	3	3	3	5	4.62E+02	-2.70E+02	0.556
600	1	1	2	2	3	4	-4.14E+02	-1.33E+04	0.76
700	1	1	2	2	2	3	-8.83E+02	-2.66E+03	0.85

Table 9: Optimal overbooking levels for different compensation plans for a six fare class

Compensation cost	Optimal # of Overbooking							Current revenue generated	Expected revenue generated	Probability of Loss
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7			
150	4	4	4	4	8	8	8	6.58E+03	3.9E+03	0
300	2	2	2	2	6	8	8	2.80E+03	1.96E+03	0.0051
400	1	2	2	2	4	4	8	1.26E+03	9.38E+02	0.22
500	1	1	1	1	3	4	5	2.94E+02	-2.30E+01	0.511
600	1	1	1	1	2	3	4	-8.30E+01	-1.00E+03	0.71
700	1	1	1	1	1	2	2	-9.58E+02	-4.90E+03	0.99

Table 10: Optimal overbooking levels for different compensation plans for a seven fare class

Sample Mat Lab codes used

Table 20: Matlab codes used to calculate the optimal number of overbooking, expected net revenue, and probability of loss.

```
function [muNetRev,stdNetRev,ProbLoss] =
NetRevObj(y,k,mu,sigma,Prices,compcost,lostrev,NumSimulations)

%This function defines the objective function along with the desired
measuring parameters


[NumClasses,m] = size(y(:));

x = zeros(NumClasses,NumSimulations);   %prelocation for speed
C = zeros(NumClasses,NumSimulations);
L = zeros(NumClasses,NumSimulations);
A = zeros(NumClasses,NumSimulations);
B = zeros(NumClasses,NumSimulations);
R = zeros(NumClasses,NumSimulations);
NetRev = zeros(1,NumSimulations);

% generation of x values


for i=1:NumClasses
    x(i,:) = gevrnd(k(i),mu(i),sigma(i),1,NumSimulations);
    x(x(:,:)<0)=0; %no-shows cannot have negative values.
    I = find(y(i) > x(i,:));
    J = find(y(i) <= x(i,:));

    A(i,I)=y(i)-x(i,I);
    B(i,J)=x(i,J) - y(i);

    aa=sum(A,1);
    bb=sum(B,1);

    a=sum(aa);
    b=sum(bb);

    if a>b
        B(i,I)=0;
        A(A>=a-b)=a-b;
    else
        A(i,J)=0;
        B(B>=b-a)=b-a;

    end
    C(i,I) = compcost(i)*A(i,I); % compensation cost

    L(i,J) =lostrev(i)*B(i,J); %cost of lost opportunity

    R(i,:) = Prices(i)*y(i) - C(i,:) - L(i,:); % net revenue
```

```
    NetRev = NetRev + R(i,:);



end

ProbLoss =  sum(NetRev < 0) / NumSimulations;   %The probability that the
mean net revenue will be less than zero

stdNetRev = std(NetRev);
muNetRev = mean(NetRev);   %the mean net revenue
```

```
clear
close all
 %this code will generate the three dimensional plot of the expected net
revenue vs. the optimal overbooking for the two classes under
consideration.

yLB = [0,0]; %lower limit of overbooking in any class, required when using
the GA algorithm
yUB = [20,20]; %upper limit of overbooking in any class

y1 = 0:20; %possible values of overbooking in class-1
y2 = y1;

[Y1,Y2] = meshgrid(y1,y2);


k = [-0.16629,-0.16629]; %Gen. Extreme distribution shape parameter
mu = [4.6347,5.822];
sigma = [2.1998,2.7355];
Prices = [299,399]; % ticket prices
compcost=[500,500]; %compensation cost
lostrev=[280,380]; %cost of lost opportunity

NumSimulations = 10000; %number of simulations

for i=1:21
    for j=1:21
        y = [Y1(i,j),Y2(i,j)];
        [muNetRev(i,j),stdNetRev(i,j),ProbLoss(i,j)] =
NetRevObj(y,k,mu,sigma,Prices,compcost,lostrev,NumSimulations);
    end
end

figure(1);mesh(Y1,Y2,muNetRev);xlabel('Y2');ylabel('Y1');
figure(2);mesh(Y1,Y2,stdNetRev);xlabel('Y2');ylabel('Y1');
figure(3);mesh(Y1,Y2,ProbLoss);xlabel('Y2');ylabel('Y1');
```

```matlab
clear
close all
 % this sample matlab code will generate the values of the mean net
revenue, the probability of loss and the standard deviation of the revenue
for the two class case.

y = [20,20];

k = [-0.16629,-0.16629];
mu = [4.6347,5.822];
sigma = [2.1998,2.7355];
Prices = [299,399];
compcost=[700,700];
lostrev=[280,380];

NumSimulations = 10000;



[muNetRev,stdNetRev,ProbLoss] =
NetRevObj(y,k,mu,sigma,Prices,compcost,lostrev,NumSimulations)
```

```matlab
clear
close all
 % This code will optimize the objective function using the Nelder Mead
optimization algorithm of the matlab optimization toolbox.

y0 = [10,10];
yLB = [0,0];
yUB = [20,20];

k = [-0.16629,-0.16629];
mu = [4.6347,5.822];
sigma = [2.1998,2.7355];
Prices = [299,399];
compcost=[700,700];
lostrev=[280,380];




NumSimulations = 10000;



% Nelder Mead
 OPTIONS = foptions;
 OPTIONS(14) = 3000;
 [yOpt,muNetRevOpt] =
fminsearch('NetRevObj',y0,OPTIONS,k,mu,sigma,Prices,compcost,lostrev,NumSim
ulations);




% Genetic Algorithms
 % The following code will optimize the objective function using the GA
optimization algorithm developed by Dr. Ashraf.
```

```matlab
[NumVars,m] = size(yLB(:));
PARAMS = [NumVars,0,100];
[muNetRevOpt,yOpt,BestFitness,Generations] =
garu('NetRevObj',yLB,yUB,PARAMS,k,mu,sigma,Prices,compcost,lostrev,NumSimul
ations);



%[F,muNetRev,stdNetRev,ProbLoss] =
NetRevObj(y,k,mu,sigma,Prices,compcost,lostrev,NumSimulations)



% Sequential Quadratic Programming
% this code will optimize the objective function under a user defined
probability of loss constraint.

 OPTIONS = foptions;
 OPTIONS(14) = 4000;
 [yOpt,FOpt] =
fmincon('NetRevObj',y0,[],[],[],[],yLB,yUB,'NetRevConstr',OPTIONS,k,mu,sigm
a,Prices,compcost,lostrev,NumSimulations);
```

71

Reference

[1] E. E. Bailey, R. D. Graham, and K. P. D, Deregulating the Airlines. Cambridge Mass: MIT Press 25, 1985.

[2] V. Ben, "Origin and Destination Yield Management," Handbook of Airline Economics, 1995.

[3] K.Amaruchkul and P. Sae-Lim, "Airline overbooking models with misspecification," Journal of Air Transport Management, Vol.17,pp. 142-146.

[4] S. E. Bodily and P. E. Pfeifer, "Overbooking decision rules," OMEGA, vol. 20, pp. 129–133, 1973.

[5] P. Belobaba, "Survey paper: Airline yield management an overview of seat inventory control," Transportation Science, vol. 21, pp. 63–73, 1987a.

[6] L.R. Weatherford and S. E. Bodily, "A Taxonomy and Research Overview of Perishable-Asset Revenue Management: Yield Management, Overbooking, and Pricing," Operations Research, Vol. 40, pp. 831-844, 1992

[7] I. J. McGill and G. Van Ryzin, "Revenue management: Research overview and prospects," Transportation Science, vol. 33, pp. 233–256, 1999.

[8] W.C Chiang, J.C.H. Chen, X. Xu, "An overview of research on revenue management: current issues and future research," International Journal of Revenue Management, Vol. 1, pp.97-128, 2007

[9] B.L. Monterio, "Airline yield management origin-destination seat inventory control," Ph.D. dissertation, Northwestern University, 1996

[10] "Yield Managers Now Control Tactical Marketing," Lloyd's Aviation Economist, pp.12-13, 1985.

[11] R. Klophaus, S. Pölt, "Airline Overbooking Considering Passengers' Willingness to Pay," Trier University of Applied Sciences,2010.

[12] S.S. Saab and P. P. Zouein, "Forecasting passenger load for a fixed planning horizon," Journal of Air Transportation Management, Vol.7,pp.361-372,2001

[13] B. Smith, J. Leimkuhler, and R. Darrow, "Yield management at american airlines," Interfaces, vol. 22, pp. 8–31, 1992.

[14] J. Nelder and R. Mead, "A simplex method for function minimization," Computer Journal, vol. 7, pp. 308-313 , 1965.

[15] R. E. Chatwin, "Optimal airline overbooking," Ph.D. dissertation, Standford university,1993.

[16] J. Bailey, "Bumped fliers and no plan B," The New York Times, May 30, 2007.

[17] S. Siddappa, "Statistical Modeling Approach to Airline Revenue Management with Overbooking," Ph.D dessertation, University of Texas at Arlington, 2006.

[18] M.J. Beckman and F. Bobkoski, "Airline Demand: An analysis of some Frequency Distributions," Naval Research Quarterly, Vol.5,pp. 43-51

[19] H.R. Thompson, "Statistical problems in airline reservation control," Operations Research Quarterly Vol.12, pp. 167-185,1961.

[20] M. Rothstein and A.W. Stone, "Passenger booking levels," Tech. report, AGIFORS Symposium Pro-ceedings, Noordwijk, The Netherlands, 1967.

[21] S.E. Bodily and P.E. Pfeifer, "Overbooking decision rules," Omega, Vol.20,pp.129-133, 1992.

[22] Z. Chi, "Airline yield management in a dynamic network environment," Ph.D. dissertation, MIT, 1995.

[23] J. Coughlan, "Airline overbooking in the multi-class case," Journal of the Operational Research Society, Vol.50, pp.1098-1103,1999.

[24] N. Aydm, S. Birbil, J. B. G. Frenk and N. Noyan, "Single-Leg Airline Revenue Management with Overbooking," Sabanc University, 2010.

[25] J. Subramanian, S. Stidham, and C. Lautenbacher, "Airline yield management with overbooking, Cancellations, and no-shows," Transportation Science, Vol.33, pp.147-167,1999.

[26] S.Luo, M.C. Akanyildrim, and R.G. Kasilingam, "Two-dimensional cargo overbooking models," European Journal of Operational Research Vol.197, 862-883, 2009.

[27] M. Rothstein, "An airline overbooking model," Transportation Science, vol. 5, pp. 180–192, 1971.

[28] J. Alstrup, S. Boas, O.B.G. Madsen, and R.V.V. Vidal, "Booking policy for flights with two types of passengers," European Journal of Operations Research, Vol.27, pp.274-288, 1986.

[29] J. Subramanian, S. Stidham, and C. J. Lautenbacher, "Airline yield management with overbooking, cancellations and no-shows," Transportation Science, vol. 33, 1999.

[30] T. Taner and J. Anthony, "Applying Taguchi methods to health care," Leadership in Health Services, vol. 19 , pp. 26-35, 2006.

[31] S. L. Brumelle and J. I. McGill, "Airline Seat Allocation with Multiple Nested Fare Classes," Journal of the institute for Operations Research and Management science, Vol.41,pp. 127-137,1993.

[32] K.Littlewood, "Forecasting and Control of Passenger Bookings," AGIFORS Symposium Proc. 12, 1972.

[33] P. Belobaba, "Air Travel Demand and Airline Seat Inventory Management," Ph.D. dissertation, Flight Transportation Laboratory, Cambridge, MIT. 1987.

[34] E. L. Williamson, "Airline network seat inventory control: Methodologies and revenue impacts," Ph.D. dissertation, Massachusetts Institute of Technology, 1992.

[35] W. Kaplan, "Integrals Depending on a Parameter--Leibnitz's Rule," Advanced Calculus, 4th ed. Addison-Wesley, pp. 256-258, 1992.

[36] http://www.mathwave.com/en/home.html (28/09/2011)

[37] http://cimss.ssec.wisc.edu/wxwise/class/aos340/spr00/whatismatlab.htm (07/09/2011)