STATISTICAL METHODS FOR FORECASTING AND ESTIMATING PASSENGER WILLINGNESS-TO-PAY IN AIRLINE REVENUE MANAGEMENT

by

Christopher A. Boyer

B.S., Operations Research, United States Air Force Academy, 2008

Submitted to the Sloan School of Management
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2010

Signature of Author: ______________________

Operations Research Center
May 12, 2010

Certified by: ______________________

Dr. Peter P. Belobaba
Principal Research Scientist
Department of Aeronautics and Astronautics
Thesis Supervisor

Certified by: ______________________

Dr. Hamsa Balakrishnan
Assistant Professor
Department of Aeronautics and Astronautics
Thesis Reader, Operations Research Center

Accepted by: ______________________

Dr. Dimitris J. Bertsimas
Boeing Professor of Operations Research
Co-Director, Operations Research Center

STATISTICAL METHODS FOR FORECASTING AND ESTIMATING PASSENGER WILLINGNESS-TO-PAY IN AIRLINE REVENUE MANAGEMENT

by

Christopher A. Boyer

Submitted to the Sloan School of Management
on May 12, 2010, in Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Operations Research

ABSTRACT

The emergence of less restricted fare structures in the airline industry reduced the capability of airlines to segment demand through restrictions such as Saturday night minimum stay, advance purchase, non-refundability, and cancellation fees. As a result, new forecasting techniques such as Hybrid Forecasting and optimization methods such as Fare Adjustment were developed to account for passenger willingness-to-pay.

This thesis explores statistical methods for estimating sell-up, or the likelihood of a passenger to purchase a higher fare class than they originally intended, based solely on historical booking data available in revenue management databases. Due to the inherent sparseness of sell-up data over the booking period, sell-up estimation is often difficult to perform on a per-market basis. On the other hand, estimating sell-up over an entire airline network creates estimates that are too broad and over-generalized. We apply the K-Means clustering algorithm to cluster markets with similar sell-up estimates in an attempt to address this problem, creating a middle ground between system-wide and per-market sell-up estimation.

This thesis also formally introduces a new regression-based forecasting method known as Rational Choice. Rational Choice Forecasting creates passenger type categories based on potential willingness-to-pay levels and the lowest open fare class. Using this information, sell-up is accounted for within the passenger type categories, making Rational Choice Forecasting less complex than Hybrid Forecasting.

This thesis uses the Passenger Origin-Destination Simulator to analyze the impact of these forecasting and sell-up methods in a controlled, competitive airline environment. The simulation results indicate that determining an appropriate level of market sell-up aggregation through clustering both increases revenue

and generates sell-up estimates with a sufficient number of observations. In addition, the findings show that Hybrid Forecasting creates aggressive forecasts that result in more low fare class closures, leaving room for not only sell-up, but for recapture and spill-in passengers in higher fare classes. On the contrary, Rational Choice Forecasting, while simpler than Hybrid Forecasting with sell-up estimation, consistently generates lower revenues than Hybrid Forecasting (but still better than standard pick-up forecasting).

To gain a better understanding of why different markets are grouped into different clusters, this thesis uses regression analyses to determine the relationship between a market's characteristics and its estimated sell-up rate. These results indicate that several market factors, in addition to the actual historical bookings, may predict to some degree passenger willingness-to-pay within a market. Consequently, this research illustrates the importance of passenger willingness-to-pay estimation and its relationship to forecasting in airline revenue management.

Thesis Supervisor: Peter P. Belobaba
Title: Principal Research Scientist, Department of Aeronautics and Astronautics

ACKNOWLEDGEMENTS

Writing a thesis at MIT is no easy task. Showing up with no background in airline revenue management does not really help, either. Therefore, I owe much gratitude to my thesis and research advisor, Dr. Peter Belobaba. He was kind enough to provide the funding for me to attend MIT in the first place, and his guidance for the past two years has been unmatched.

In regards to PODS, I would like to thank Craig Hopperstad. His expertise in airline revenue management methodology and programming made this thesis possible. I appreciate his patience with my questions and the work he puts into creating the program that essentially fueled this thesis.

Next, I would like to thank Claire Cizaire. Upon arrival to the ICAT lab, she introduced me to PODS and always answered my questions with a smile, no matter how unintelligent they were. Her help and wisdom has been extraordinary and greatly appreciated. In addition to Claire, I'd like to thank the rest of the PODS team—Olivier, Sarvee, and Himanshu. We have had some amazing trips the past two years. Hiking around Patagonia with Olivier, even with his extreme Frenchness, was a blast and a trip of a lifetime. However, I have no regrets that my French is still limited to "oui, oui." In addition, to the rest of the ICAT lab, thank you for making work every day exciting. To Nikolas, Yannis, Diana, Mehdi, and Thomas, you will be missed.

To my roommates Tommy and Jake, thank you for the great times that we shared here in Boston. You made this place turn into a home away from home. To Parker and Zeus, I will always miss the days of watching you two relentlessly ravage each other. I know it's hard, but try to stay clean and don't eat too many chicken bones.

To my fellow Air Force Academy, Naval Academy, and West Point friends here, thank you for the great weekends we spent together. Also, let it be said that we are legends in MIT Intramural Flag Football.

I owe a lot of appreciation to Jessica. Thank you for your support and patience, as I know that I do not always make it easy. It is difficult to be tied to the military when you didn't sign up for it, so thank you for always understanding, and I cannot wait for what the future holds for us.

Last, I would like to dedicate this thesis to my parents and my older brother. Thank you for getting me to where I am today and for crafting me into the young man I have become. Your encouragement is always something that I appreciate and that I will always lean on.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 7

CHAPTER 8

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 6

CHAPTER 7

CHAPTER 1

INTRODUCTION

As the airline industry evolves to adapt to an ever-changing competitive climate, airlines are searching for every opportunity to increase their revenue. While there are numerous economic factors that are largely uncontrollable by the airlines, the area of airline revenue management (RM) remains a constant area of focus. Airline revenue management is the practice of controlling a fixed and perishable resource—an inventory of seats on an aircraft for a future flight—by allocating a certain number of seats to specific predetermined fares, known as fare classes. Setting the proper booking limit, or the maximum number of seats made available to a specific fare class, is crucial to ensuring that there are not too many seats available to lower fare class passengers. If this occurs, the future flight will sell out too quickly and generate a smaller amount of revenue than if it had "protected" more seats for later-booking, higher-paying customers. Likewise, if there are too many seats left for higher-paying passengers, planes will depart with empty seats, which again reduces revenue.

The goal of airline revenue management is to maximize revenue by getting every passenger to pay his or her maximum willingness-to-pay (WTP). For instance, if a passenger's maximum WTP is $300, and he or she only pays $250, there is $50 of "lost" revenue to the airline (also known as "consumer surplus"). In order to achieve this goal, creating the proper booking limits for each fare class on each flight is essential. There are several pieces of revenue management that play a vital role in maximizing revenue.

A fundamental part of the puzzle in RM systems is the demand forecasting model, which relies on both historical bookings for the same flight in the past (same time and day of the week, season, etc) and the actual bookings-to-date for the specific flight in question. The model uses these data to create fare class forecasts that are fed into a seat allocation optimizer, which determines the appropriate number of seats to make available to a specific class. Inside the forecasting model, in addition to the historical and current bookings, there lies an estimate of what is called "sell-up."

The sell-up estimate accounts for the potential for passengers to buy the next higher fare class than they originally intended to purchase (given that it was unavailable), moving closer to their maximum WTP. There are several different methods to estimate sell-up, whether it is determined separately from the forecasting algorithm or within the algorithm using the historical data at hand. In addition, one may determine these sell-up estimates and aggregate them on different levels. This may be by market (city-to-city pairing), by groups of markets that share similar characteristics, or on a system-wide basis.

The goal of this thesis is to investigate statistical methods for forecasting and estimating sell-up. More specifically, it will determine the appropriate levels of market aggregation for these estimates based on the characteristics of each market, airline networks, and various fare structures in order to emulate the various aspects of the airline industry.

1.1 The Airline Industry and the Need for Revenue Management

Since deregulation of the U.S. airline industry in 1978, information technology and operations research (OR) have contributed greatly to the industry's performance. Several advances and new capabilities in both computing power and in OR methods and theory enable the airlines to maximize their profit in regards to aircraft planning, crew schedule planning, and airline revenue management. Barnhart et al. discuss the gains from OR in each of these areas, emphasizing that even though there is no one single optimization model to perfectly model an airline's profit maximization problem, it may successfully be broken into more manageable pieces that work together (Barnhart, Belobaba, & Odoni, 2003).

For example, an airline must first determine what origin-destination (O-D) markets to serve based on demand and competition, followed by the creation of a feasible schedule. Next, the airline must assign its fleet of aircraft to the routes in the schedule, based on the demand matching the capacity of aircraft assigned to the route, costs, aircraft maintenance needs, and the feasibility of the schedule, among other things. Finally, the airline must assign crew to those flights and aircraft, subject to factors such as work hour constraints and pilot qualifications. While these decisions are crucial to the success of the airline, revenue management plays a vital role in actually filling the aircraft with the optimal, revenue-maximizing passengers.

The goal of revenue management is to establish booking policies that maximize an airline's profits, or because short term costs are fixed, to maximize an airline's revenues. Both McGill and van Ryzin, as well as Barnhart et al. describe revenue management as developing these policies to determine the appropriate number of seats made available for each booking class, largely in an effort to save, or "protect," seats for the business passengers who tend to book closer to departure and pay more (McGill &

Van Ryzin, 1999), (Barnhart, Belobaba, & Odoni, 2003). At the start of the progression of airline revenue management in the early 1980s, computer reservations systems stored data from the booking process. This advanced further in the mid-1980s to a system that could track flight reservations and compare them to an ideal booking curve for the flight over time. Today's revenue management systems, now considered to be the third generation of systems, are comprised of multiple components, as described in Figure 1.

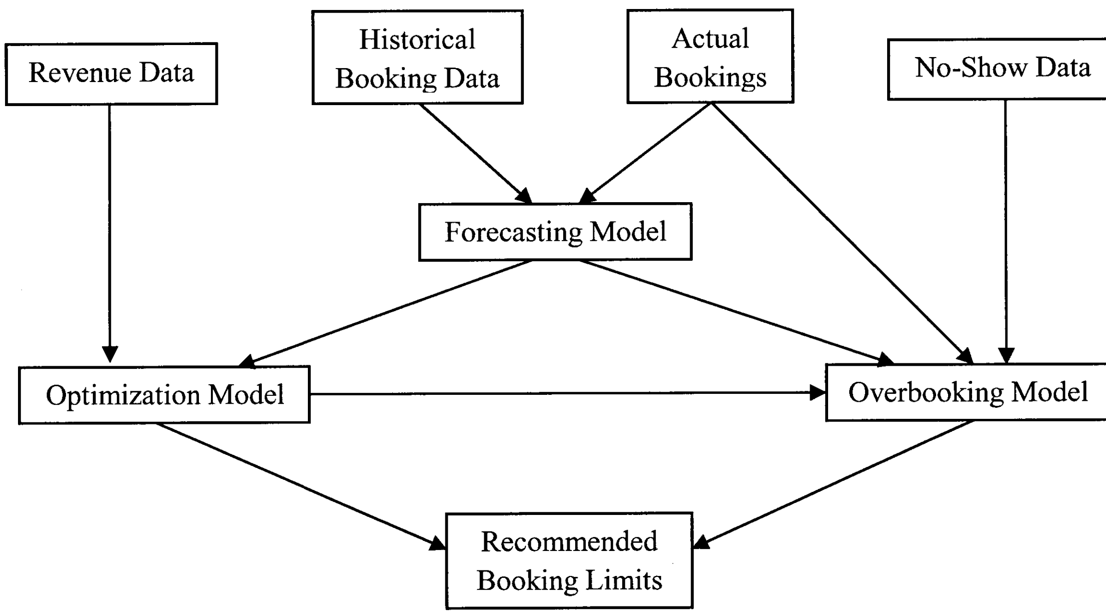


Figure 1: Third-Generation Airline RM System (Belobaba, Fundamentals of Pricing and Revenue Management, 2009)

First, historical booking data from the same flight in the past (keeping in mind seasonal, day-of-the-week, and other elements) and the actual booking progress data combine to create a forecast for the upcoming flight on a per-class basis. This forecasting model, in conjunction with the revenue data, or pricing information and value for each fare class, are inputs into the optimization model that work to create the best booking strategy to maximize revenue on the flight leg. At the same time, the optimization model, forecasting model, actual bookings-to-date, and the no-show data from similar historical flights combine to create an overbooking model, which determines an appropriate number of seats of each class to make available with the consideration that people will not show up for various reasons. Finally, the optimization model and the overbooking model provide inputs to create the recommended booking limits per class for the flight leg. In addition, revenue management analysts often review these booking limits to ensure sensibility and also to edit booking limits that should change due to sudden unforeseen increases or decreases in demand.

These RM models provide tools for airlines to optimize flight revenues based on the airline's operating network schedule and fleet. The use of airline revenue management and the advent of seat inventory optimization and overbooking models give airlines a sizeable advantage over those who do not use them, with simulations and actual experience citing a 4-6 percent increase in revenue (Barnhart, Belobaba, & Odoni, 2003).

1.2 The Rise of Low Cost Carriers and Changing Fare Structures

In his PhD dissertation, Emmanuel Carrier discusses the growth of low cost carriers (LCCs) in the market as a competitive response to the industry's extensive use of hub-and-spoke networks and price discrimination strategies. A hub-and-spoke network enables legacy carriers to provide numerous travel options, all connecting through a hub. This diverse set of origin-destination markets produces more destinations for all travelers. In addition, the network legacy carriers' use of price discrimination effectively separated the demand between leisure and business passengers through the use of fare restrictions. Business travelers, who tend to be more price inelastic, were forced to pay more to avoid such restrictions, especially that of the "Saturday Night Minimum Stay" requirement. Because business travel is done primarily during the week, rarely does a business traveler wish to stay over a weekend. On the contrary, price sensitive leisure passengers were offered lower fares as long as they could meet advance purchase, non-changeable, and non-refundability restrictions in addition to the minimum stay requirement. This segmented the demand between price-oriented (leisure) and product-oriented (business) travelers (Carrier, 2008).

With network carriers focused on providing the most travel options and gaining the most revenue from high yield-business travelers, the door opened for LCCs to focus on providing non-stop point-to-point service in major markets. In response to the legacy carriers' strongholds on various hub cities, several of the major LCCs operate with service into and out of high demand "focus cities." In effect, these are small hubs for the LCCs. With the network carriers' pricing policies gradually creating increasingly more expensive tickets, some business passengers began to avoid the higher priced tickets by opting to stay over Saturday nights, among other things.

Carrier asserts that LCCs established their fare structures using two main strategies with the hopes of still having a segmented market—creating a lower fare dispersion (fare difference between the highest and lowest classes), but still achieving demand segmentation solely through advance purchase requirements and accurate demand forecasting for their revenue management systems. The LCCs effectively removed the entire Saturday Night Stay requirement by offering fares on a one-way basis. With the removal of most travel restrictions, a "semi-restricted" fare structure is now available to the markets served by LCCs. Potential issues for legacy carriers rose, the most important being whether or not to match the lowest fares and fare structures in the competitive markets, as well as the question of whether or not their current revenue management systems could adapt to an unrestricted or semi-restricted fare structure. The former did happen, with the several legacy airlines creating their own low-cost subsidiary airlines featuring overall fare "simplification" to match the LCCs, and the latter is discussed later in this thesis.

1.3 Thesis Objective

Central to an airline's revenue management system is its ability to create forecasts that are not only accurate, but favorable enough to protect the high-fare seats so that the spiral-down effect, or the systemic under-forecasting of high-fare seats in response to a lack of high-fare class demand and purchases, does not occur. (Spiral down will be discussed more in Chapter 2.) In order to combat the erosion of protection levels of these seats, and thus the overall revenues, an effective method of estimating sell-up is a necessity to keep the forecasts at their revenue-maximizing levels. Sell-up is further defined as the occurrence of a price-oriented passenger (one who will always purchase the lowest available class) paying more for the next available higher fare class (as long as it is less than or equal to his or her maximum willingness-to-pay) when the revenue management system closes the originally sought-after lower fare class.

Past work on the estimation of sell-up includes development and analysis in the Passenger Origin-Destination Simulator (PODS) on two separate data-based estimation methods, each with the option of estimating sell-up on a per-market Origin-Destination basis, or over the entire airline system. An alternative to this approach is to estimate the sell-up for similar markets as a group, as defined by several characteristics and parameters of each market. This thesis will discuss and apply a clustering algorithm to sell-up estimation in conjunction with Hybrid Forecasting, a new forecasting approach that incorporates passenger willingness-to-pay, and analyze the results. The data-based sell-up estimation methods will be compared and contrasted with "input" sell-up estimation methods, in which the airline chooses an arbitrary estimate of sell-up. The sell-up estimates for the input methods are determined independently of previous booking data. Following the clustering results, the thesis includes a chapter that focuses on the use of regression analysis to determine the relationship between a market's sell-up estimate and the market's characteristics.

Furthermore, this thesis formally introduces the method of Rational Choice Forecasting, which develops forecasts and sell-up estimation internally, as compared to Hybrid Forecasting. Rational Choice

Forecasting is a regression-based forecasting method based on booking observations and the partitioning of demand and passenger types.

1.4 Thesis Organization

This thesis contains eight chapters. The introduction lays out the current situation in the airline industry today, with the development of the potential problems facing airline revenue management. The literature review contains all pertinent information on techniques and forecasting methods already established to adapt to today's airline fare structure environment. The next chapter on sell-up estimation discusses the current efforts to combat spiral down including methods used to estimate passenger willingness-to-pay. The following chapter describes Rational Choice Forecasting, a new linear regression-based forecasting method that creates forecasts that incorporate passenger sell-up probabilities in a single step, to include its use with fare adjustment. The next chapter introduces the clustering algorithm, a new method for market aggregation and application of sell-up estimation. Chapter 6 lays the groundwork for all testing in the thesis, using various airline networks and fare structures. It includes the application and results for all new sell-up estimation methods presented in the thesis compared to past sell-up estimation methodology, with their application to Hybrid Forecasting and Rational Choice Forecasting. The Passenger Origin-Destination Simulator (PODS) serves as the primary method for creating a controlled simulated environment for testing and analyzing these various revenue management methods. Chapter 7 highlights the results of a regression analysis describing the relationship between sell-up estimates and various market characteristics, independent of the actual booking data. Last, the conclusion describes the impacts of the thesis in addition to laying the groundwork for future work and research directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Airline Revenue Management

The first notion of the existence of airline revenue management began in 1972 at the British Overseas Airways Corporation (now British Airways), where they presented a two-class fare structure—one class, or the "earlybird" class, would be available for 21-day advance purchase and offer a discounted fare in order to fill seats that would otherwise be empty. The problem arose when, in an effort to avoid filling the airplane with all earlybird passengers, the airline needed to determine how many seats to protect for the later, higher paying customers. Littlewood claimed that in order to maximize the flight's revenue in this two-class situation, the airline should accept a discount fare booking as long as its revenue exceeded the expected revenue of a future full fare booking (Littlewood, 1972). This later became known as Littlewood's Rule, and thus gave rise to the field of yield management, or revenue management in today's terms. The literature review below describes the important and relevant seat allocation optimization methods, in addition to the evolution of the airline industry and the adaptation of revenue management systems.

2.1.1. Important Seat Allocation Optimizers (EMSRb, DAVN, Bid Price Control)

In order to maximize revenue on a particular flight, there are two necessary steps. First, one must establish the fare structure. Despite current practices of lessening or removing the booking restrictions applied to each fare class, the practice of differential pricing, or the assignment of different fares to different fare class products is crucial to separate demand. Differential pricing creates the need for the second step, seat inventory control, or managing the amount of seats available for purchase for each fare class product. There are multiple methods to perform seat inventory control. Belobaba discusses the basic options available to airlines today through revenue management software or in-house development

(Belobaba, Airline Network Revenue Management: Recent Developments and State of the Practice, 2002), (Belobaba, Fundamentals of Pricing and Revenue Management, 2009). Here, we will examine three of them.

Developed in his doctoral thesis, Belobaba established the most commonly used basic seat allocation model, the Expected Marginal Seat Revenue (EMSR) heuristic (Belobaba, Air Travel Demand and Airline Seat Inventory Management, 1987). He later refined it to EMSR Version B, or EMSRb (Belobaba, The Revenue Enhancement Potential of Airline Revenue Management Systems, 1992). In summary, the airline of interest has serially nested booking classes, which means that for a given aircraft capacity, all of the seats in the cabin are available to the highest, or most expensive booking class. For example, if an aircraft has a 120 seat capacity, and all 120 passengers will pay the highest fare, 120 seats should be available at the highest fare. Realistically, however, the goal of the heuristic is to determine how many seats to protect for these higher paying customers, and then establish booking limits for the lower classes. After the highest booking class protection level is determined, the heuristic then calculates how many seats to protect for the second highest booking class from the rest of the lower booking class passengers. This process repeats down to the second lowest booking class, creating a nested effect:

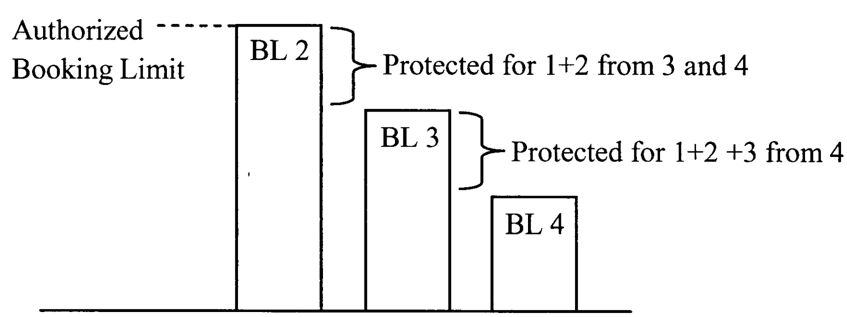


Figure 2: Nested Seat Protection with Booking Limits (Belobaba, Fundamentals of Pricing and Revenue Management, 2009)

In order to establish limits for each booking class, EMSRb uses the airline's estimates of mean demand and standard deviation for each booking class (given a certain probability distribution), assuming that the demand is independent for each booking class on each flight leg. Other assumptions include that the demand for each class is stochastic, and that fare classes book in order from the lowest class to the highest class. The expected marginal seat revenue for each incremental seat is then determined as the average fare for the booking class multiplied by the probability that the demand for the seat will actually materialize. Therefore, if the expected marginal seat revenue for one additional seat is higher than if the seat was given to the next lowest class, it will remain protected for the higher class. Starting from the highest booking class, the heuristic determines the booking limits by subtracting the number of seats to be protected for the given class from the remaining capacity on the flight leg. Note that this process repeats throughout the booking period as long as the remaining demand and standard deviations, as well as the remaining capacities, are updated. For more thorough detail, reference Belobaba's thesis and paper mentioned above.

While this method is employed by numerous airlines, claiming increases of 2-4 percent compared to situations where airlines do not use a seat inventory control algorithm, the EMSRb method has shortcomings (Belobaba, Application of a Probabilistic Decision Model to Airline Seat Inventory Control, 1989). The EMSRb method maximizes revenue for single leg flights and assumes that the demand for each leg is independent. However, with the current widespread use of hub networks, many passengers have connections, which creates network issues for the EMSRb heuristic—the assumption of independence for each leg's demand does not hold. Therefore, maximizing the revenue on one flight leg does not guarantee that the entire network's revenues will be maximized. For more explanation and detailed examples of why this causes problems, refer to (Belobaba, Airline Network Revenue Management: Recent Developments and State of the Practice, 2002). In order to account for flights across a network, two methods are widely used—displacement adjusted virtual nesting (DAVN) and bid price control, which belong to what many consider the 4th generation, or the "path-based" RM system.

In order to determine the availability of a booking class on a flight leg that may be just one piece of a larger Origin-Destination (O-D) itinerary, DAVN uses the following mechanism. The total network value of an itinerary is not always its O-D fare, for this would place too much value on a connecting itinerary and leave few seats to be protected for just the "local," or non-connecting single leg passengers. Rather, the DAVN mechanism sets the network value of the itinerary to the total O-D fare minus the potential revenue loss of displacing a passenger from their local flight leg, whether it is on down-line legs or up-line legs in the original itinerary. For example, suppose a passenger wishes to fly from Denver to Pittsburgh through Chicago for $500. Given that the passenger will displace a local Denver to Chicago passenger whose value would be $150 and then another Chicago to Pittsburgh passenger whose value would be $125, the total network value of the original Denver to Pittsburgh passenger's fare would be $500 - ($150 + $125) = $225. Note that the local passengers were given values of $150 and $125, which are not necessarily fares, for there are several different methods for estimating displacement costs. Using this method, DAVN places each of these values into "virtual buckets," or bins with a given network value range hidden in the airline's computers and revenue management system. Then, like the EMSRb method, the airline then determines the availability and protection levels for each of these buckets.

Another network seat allocation method, similar to DAVN in that it uses displacement adjusted network values for a flight leg, is bid price control. This method uses a simple rule that determines if a "bid price" for an itinerary should be accepted or rejected. The bid price method determines the value of the itinerary as the O-D Fare minus the Network Displacement Cost. The method says to accept the request if this value is greater than the value of the last, or lowest-valued seat on a leg in the connecting itinerary. This rule for accepting may then be expressed as:

Accept if: O-D Fare > Value of Last Seat on Leg + Network Displacement Cost

O-D Fare > Minimum Bid Price

Therefore, if the request is greater than the minimum value of the bid price determined above, the seat is available for purchase. Note that this is less complicated and requires less data storage than the DAVN method, for the airline only needs to store bid prices for each future flight leg based on the current bookings, rather than entire virtual bucket booking limits for every class on every flight leg. In simulations, these network O-D Controls (DAVN and Bid Price Control) in the 4^{th} generation of RM systems are responsible for 1-2 percent increases in revenue over existing leg-based methods. This 1-2 percent increase is in addition to the 4-6 percent increase in revenue just due to fare class mix, or seat allocation optimization and overbooking methods. For greater detail on DAVN and Bid Price Control methods, see (Belobaba, Airline Network Revenue Management: Recent Developments and State of the Practice, 2002) or (Belobaba, Fundamentals of Pricing and Revenue Management, 2009).

2.1.2. Traditional Forecasting Models

Forecasting is an integral part in determining the optimal booking limits in revenue management. While airlines have databases of historical purchases for every class on every flight, these data do not indicate the true uncensored demand for that fare class. Rather, the data is censored by the booking limits on the previous flights that prevented a potential passenger from purchasing the fare, which also denies the airline from realizing that passenger's demand. It is therefore the goal of the airline to find the uncensored demand in order to create accurate forecasts for the disaggregate fare class level for each flight. According to McGill and Van Ryzin, several attempts were made by American Airlines and Sa to estimate unconstrained demand on an aggregate level through regression and time series data analysis (McGill & Van Ryzin, 1999), (Sa, 1987). However, at the disaggregate fare class level, the best tool for forecasting demand is the use of demand data from the same flight in recent weeks.

Forecasts must contain an element of the historical bookings in addition to having the capability of adjusting for current bookings throughout the booking process. What sounds like a classic time series analysis problem, forecasting in airline revenue management requires computationally fast and simple, yet accurate estimates of demand. Talluri and van Ryzin devote an entire chapter of their book to various estimation and forecasting methods, to include stationary and non-stationary time series forecasting, ad hoc forecasting, Bayesian forecasting and neural network forecasting, which is a type of machine learning (Talluri & Van Ryzin, 2004). While all of these are outstanding methods to forecast demand, several require individual attention and analysis before one can determine the proper application of the forecasting method. For example, in the time series case, one must first make sure the data is stationary, determine the appropriate lagging measure, and apply a certain auto-regressive and/or moving average process to create a forecast. If the data was non-stationary, it must then be converted back to its original form. Another forecasting method presented is neural networks, which in short, unlike the other methods, determines its own "best" functions for forecasting demand based on a given input. For example, given ten inputs of historical demand, the neural network uses machine learning (setting aside a portion of the data as a training set) to create three outputs of demand. The potential problem is that the functions lay within a "black box," making it difficult to see and understand what is going on with the forecast, an essential piece of explaining how each forecast is developed to revenue managers.

While "ad hoc" forecasting (exponential smoothing with trend and seasonality) is prevalent in the revenue management world, in regards to this thesis, the most important basic forecasting method presented is pick-up forecasting (or standard forecasting). Pick-up forecasting essentially uses historical bookings for a given flight and class to determine the future bookings-to-come forecast.

-3 Days	-2 Days	-1 Day	0 Days	Flight Date	Bookings-in-Hand	Bookings-to-Come
8	13	3	13	9-Jan	37	0
11	5	4	2	10-Jan	22	0
6	2	6	8	Today	22	0
6	3	2	7.67	12-Jan	11	7.67
1	2	3.75	7.67	13-Jan	3	11.42
4	5	3.75	7.67	14-Jan	4	16.42

Table 1: Incremental Bookings: an Example of Pick-up Forecasting (Talluri & Van Ryzin, 2004)

For example, given a certain daily flight and a given class denoted above in Table 1, one may wish to determine the forecasted bookings-to-come for the next three days. Pick-up forecasting determines the number of bookings for 12 January simply as the mean of the bookings occurring on the day of the flight for the previous days, or (13+2+8)/3. The same holds true for the -1 Day situation, as well as the -2 Days situation. The bookings-to-come forecast is simply the sum of the forecasts for the days of booking remaining for a particular flight in the future.

Another method for forecasting demand is a simple regression-based forecaster, which, like pick-up forecasting, is used in the Passenger Origin-Destination Simulator (PODS), which will be discussed later. According to Hopperstad, this method simply regresses the predicted bookings-at-departure (a cumulative number) on the given historical bookings up to the current time frame for a given flight and a given class. Once the coefficients are determined, one can solve for the predicted bookings-at-departure. The predicted bookings-to-come is simply the predicted bookings-at-departure minus the cumulative historical bookings-in-hand (Hopperstad, 2005).

2.2 Low Fare Airlines and Simplified Fare Structures

Swelbar discusses the current situation in the airline industry, citing many challenges (Swelbar, 2009). What Swelbar refers to as "midscale airlines," comprised mainly of major LCCs, have a sizeable advantage over the network legacy airlines in terms of costs. In 2008, the midscales had a 5.55 cents per available seat mile advantage over the network airlines, with 1.01 cents of that being labor and related costs. This benefits the midscales greatly when it comes to revenue management and pricing, for they are able to offer lower fares while still maintaining profitability. However, the overall trend for the oil prices remains volatile. Between 1995 and 2004, the fuel cost per available seat mile remained between one and two cents. In 2008, this jumped to around five cents, causing profitability to become more difficult. Additionally, between 2000 and 2008, the US airlines lost over $34 billion (Swelbar, 2009). This information only reinforces the importance for the airlines to squeeze every cent of revenue out of its business through various methods. One of those methods, airline revenue management, is a vital tool for every airline. This thesis focuses on just one aspect of airline revenue management, but it is still important to note that improving an airline's revenue by just half a percent could be crucial to the airlines staying profitable.

An inherent assumption in most revenue management systems is the independence of demand between fare class types. This assumption worked well in the past, before the emergence of LCCs. Previously, legacy carriers embraced a fully differentiated fare structure that automatically separated demand and split it among the fare classes through the use of restrictions, such as the Saturday night minimum stay, advance purchase, cancellation fees, and refundability restrictions. For example, if a business traveler did not want to spend a weekend at his destination, or needed to change his plans at the last minute, the low-cost fare would not be available to him. Likewise, a leisure traveler who was able to make solid plans further in advance for a vacation and who was staying over a weekend was able to purchase a lower fare. These fare structures therefore segmented demand, enabling the assumption of the independence of demand for each different fare product. However, the emergence of the LCCs caused these thoughts to change. With much lower operating costs, LCCs were able to able to offer much cheaper fares, gaining a significant amount of market share. They also began to offer a simpler, less-restrictive fare, leaving legacy carriers no choice but to match restrictions and prices.

2.2.1. Impacts of New Fare Structures on RM Systems and the Spiral Down Effect

Revenue management systems suffered significantly, as these new fare structures created demand patterns that violated their crucial assumption. Belobaba claims that the fallout of the differential pricing mechanism raises many questions about the ability of RM methods to account for new "unrestricted" fare structures (Belobaba, Fundamentals of Pricing and Revenue Management, 2009). If a business passenger needs to travel, despite having a very high willingness-to-pay, a low-priced unrestricted fare is obviously preferred over the former high-priced, unrestricted fare product. Now, nothing will prevent former high-paying customers from purchasing in the lowest fare class and receiving the same freedoms and lack of restrictions. RM systems will subsequently forecast more demand to the lower classes, and thus under-forecast the demand for higher-priced seats. Furthermore, even more low-class seats (and fewer higher-class seats) would in-turn be purchased, reinforcing the over-forecasting of demand for low-class seats, creating a vicious cycle known as the spiral-down effect. This spiral-down phenomenon leaves virtually no seats left at the higher fare, resulting in huge revenue losses to the airline, and is shown in Figure 3.

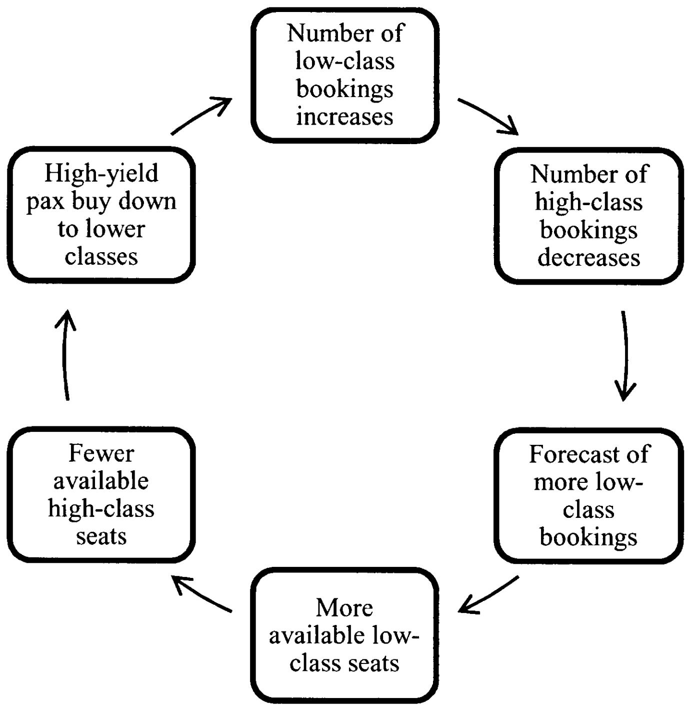


Figure 3: The Spiral-Down Effect

For a mathematical representation of the spiral down model, reference Cooper et al. (Cooper, Homem-de-Mello, & Kleywegt, 2006).

According to Richard Zeni, once the LCCs became more established, their costs also increased with more experienced workers and rising maintenance hours. In addition, the increase in fuel costs put more pressure on the current unrestricted fare structure dilemma that they helped create (Zeni, 2007). This created the need for RM systems to adapt to the new fare environment, to which much of today's research is devoted. It is the goal of the RM system to continue to close down lower fare classes, despite there still being demand for them, so that higher-priced seats are still available for passengers with a higher willingness-to-pay. Therefore, incorporating this concept of estimating passengers' willingness to "sell-up" into the higher fares is crucial to combat spiral down and the potential revenue losses of an unrestricted fare structure.

In addition, there are other prospective methods to create sell-up adjusted forecasts that bypass the use of historical bookings. Zeni suggests that just by looking at the historical booking environment, the demand is already constrained by the fact that not every fare class is open and available for purchase. The goal of forecasting demand is to use the unconstrained demand, demand that is uncensored by any conditions, such as a fare class being closed. The use of online airline search engines, he believes, is the answer. He writes that by monitoring the number of searches and purchases for a city-to-city market pairing represents a viable way to account for demand. If the booking activity for a market is high, perhaps the price is too low and the lower fare class should be closed. Likewise, if there is little activity, opening a lower fare class would be a viable option. Therefore, estimated demand would no longer depend on the past; rather, these estimates are purely current and accurate estimates for a future departure date (Zeni, 2007). While this seems like an acceptable idea, it may not be practical at this moment. This thesis focuses more on the adaptation and analysis of the current use of historical bookings to create forecasts.

2.3 Revenue Management Tools for Today's Environment

Gorin and Belobaba examined the effects of low-cost carrier entry on the incumbent airlines. Their experiments in the Passenger Origin-Destination Simulator (PODS) showed that the entry of a low-cost carrier dramatically reduces the revenues of their competition, as expected. Revenue losses for an LCC's head-to-head competitor range between 5 and 11 percent, depending on various parameters, such as price matching and the LCC average capacities. However, the incumbent airlines' best tool to combat revenue loss is choosing a proper revenue management strategy. The experiments show that matching the pricing strategy reduces revenue loss by limiting the number of local passengers stolen by the LCC. In addition, choosing a Network RM system is the most robust solution, for it consistently offers incremental benefits between 1.2 and 1.4 percent over a leg-based RM system in a variety of scenarios with LCC competition (Gorin & Belobaba, 2004). In addition to their experiments, new forecasting and optimization methods provide relief against revenue loss in unrestricted fare environments.

2.3.1. Price-Oriented versus Product-Oriented Demand

Boyd and Kellesen discuss the two types of demand relevant to airline revenue management and the booking process in changing and less restricted fare structures (Boyd & Kallesen, 2004). The key assumption in most revenue management models is that the demand for each fare class is exogenous and independent. However, recent fare structures that remove restrictions designed to segment demand on a fare class basis remove the applicability of this assumption. For example, in a two class fare structure, if the cheaper fare is available and has the same restrictions (or lack thereof) as the more expensive business fare, the business passengers will buy down to the lower fare. Therefore, with the removal of advanced purchase, minimum stay, and other restrictions, the only difference separating a leisure and business traveler is the maximum willingness-to-pay. This is what is known as *priceable demand*. In this thesis, this type of passenger will also be referred to as "price-oriented." On the contrary, for a flight with a more restricted fare structure designed to segment passengers, there is *yieldable demand*. This thesis also considers this type of demand as "product-oriented." Boyd and Kallesen ask the simple question, "Are airline fare classes different products (yieldable demand) or different prices for the same product (priceable demand)? Airlines cannot simply choose one or the other, for different results will come from both extremes. However, it is important to research the types of markets each flight serves. For example, for a less restricted market, often served by low cost carriers who introduced the concept of simplified fares, airlines often choose to match the policy by implementing an unrestricted fare structure assuming more price-oriented, or leisure demand.

If airlines create forecast models that incorrectly forecast the demand, either too much low-fare demand will be forecasted (spiral down), resulting in more lower class seats sold and less revenue, or too much high-fare demand will be forecasted, resulting in more higher class seats, but no one to sit in them. The resulting problem remains that if the demand is price-oriented in a multi-fare class environment, how does one determine the customer's maximum willingness-to-pay? This information is therefore considered "censored," because a passenger's true willingness-to-pay may be masked by their purchase of lower fares with the same restrictions as higher priced fares. The passenger's maximum willingness-to-pay must then be estimated in an effort to get the passenger to sell-up into the higher fare if the lower fare class is closed.

		Forecasting Model for:	
		Yieldable	*Priceable*
Actual	*Yieldable*	Good	Overestimates high fare demand at the expense of low fare demand
	Priceable	Overestimates low fare demand at the expense of high fare demand (spiral down)	Good

Table 2: Effects of Using the Wrong Forecasting Model (Boyd & Kallesen, 2004)

Table 2 shows the scenarios of the effects of using forecasting models designed for either yieldable or priceable demand, when the demand is actually yieldable or priceable. The lower left scenario shows the classic case of spiral down, where a forecasting model built for yieldable demand will eventually overestimate low fare demand because all of the demand is truly priceable. The opposite case in the upper right, showing a priceable demand forecasting model being applied to yieldable demand, results in spoilage of seats because too many higher-class seats were protected.

Boyd and Kallesen suggest that airlines develop the price versus product-oriented demands separately, and then combine these into a single hybrid forecast, which is presented in the following section.

2.3.2. Q and Hybrid Forecasting

Given that there are two types of passengers, product-oriented (business) passengers and price-oriented (leisure) passengers, the Q-forecasting method was developed by Belobaba and Hopperstad in application to the price-oriented demand (Belobaba & Hopperstad, Algorithms for Revenue Management in Unrestricted Fare Markets, 2004). Q-forecasting counters the spiral down effect that resulted from a lack of product differentiation in unrestricted fare structures. In an unrestricted fare environment, it is assumed that all passengers are price-oriented, for they will all want to purchase the lowest open class, given it does not have any different fare restrictions. The purpose of Q-forecasting is to generate forecasts by passenger willingness-to-pay for price-oriented demand. In short, the method works as follows:

Sell-up probabilities are first established between the lowest fare class (Q) and the rest of the *f* fare classes. Note that these will change over the time frames, for later booking product-oriented travelers will be more price-inelastic and have a higher sell-up rate. The number of observed bookings for each time frame and class are converted into the equivalent number of Q-class bookings by dividing by the probability of sell-up from Q to fare class *f*. Next, pick-up forecasting and detruncating are applied to the total Q-class equivalent bookings to estimate the total unconstrained Q-class bookings for the time frame. This Q-class equivalent forecast is then partitioned into the other fare classes via the sell-up probabilities. This process is repeated for all other time frames, followed by the summation over each fare class to develop the total bookings-to-come forecast per fare class.

Another method of forecasting, known as Hybrid Forecasting, exists for use in semi- and fully restricted fare structures where both price and product-oriented demand exists. Hybrid Forecasting uses a combination of Q-forecasting to estimate price-oriented demand and standard pick-up forecasting to estimate product-oriented demand. For Hybrid Forecasting, a product-oriented person is defined as a passenger who purchases a fare class higher than the lowest open fare class. The bookings-to-come forecast developed from pick-up forecasting is added to the bookings-to-come forecast created from the Q-forecasting method, which results in the overall per class forecast for the flight. For more information on Q- and Hybrid Forecasting methodology and revenue impacts, see (Reyes, 2006).

2.3.3. Fare Adjustment

Fiig et al. discuss the development of a new method of revenue optimization that is unaffected by changes in fare structure, whether it is the less restricted fares influenced by the growth of LCCs in certain markets, or the traditional restricted fare structure that partitions demand. In addition, this method allows the continued use of traditional revenue management systems, even though they operate under the precarious assumption that demand is independent for each fare class. This method, referred to as "fare adjustment," alters the fares based on an estimated price elasticity cost, lending itself as a useful application for any RM optimization method (Fiig, Isler, Hopperstad, & Belobaba, 2010).

Consider a case in an unrestricted fare structure, meaning that there are no restrictions partitioning demand and separating it into different fare classes. Given the lowest open fare class k, the total quantity sold in k, Q_k, the demand for class j, d_j, the total revenue for class k, TR_k, and the fare for class k, f_k, the following equations are established:

$$Q_k = \sum_{j=1}^{k} d_j \text{ and } TR_k = f_k Q_k$$

The total quantity sold is the sum of the total demand in all classes down to class k (1 denoting the highest class), because everyone with a higher willingness-to-pay will still pay less for the same level of restrictions. The total revenue for class k is then the fare for class k times the total quantity sold.

The next step is to find the incremental revenue, or the revenue loss due to buy down by opening up an additional lower fare. This is not simply the demand times the fare, $d_k \times f_k$, but a smaller amount because the demand at the higher class k-1 will now buy down to the lower class k. Therefore, the revenue loss due to buy down for class k-1 is $Q_{k-1}(f_{k-1} - f_k)$. This adjustment, or reduction amount, when applied to the original thought of incremental revenue yields the equation $d_k f_k - Q_{k-1}(f_{k-1} - f_k)$. Fiig et al. describe this "adjusted fare" as the marginal revenue for class k, MR_k, which may be rewritten as:

$$MR_k = \frac{TR_k - TR_{k-1}}{Q_k - Q_{k-1}} = f_k'$$

Thus, according to Fiig et al., the optimization rule when applying fare adjustment to any fare structure is to "order the fares in decreasing marginal revenue and open fares until capacity is reached or the marginal revenue becomes negative" (Fiig, Isler, Hopperstad, & Belobaba, 2010).

In regards to sell-up, in the case of the fully undifferentiated fare structure, this formula may be again rewritten with sell-up probabilities. The demand for k, the lowest open fare class, is $d_k = Q_k$, for the demand for all other fare classes is zero. Denote the sell-up probability from the lowest class n to the class k as $psup_k$, and the base demand as Q_n. Therefore, the demand $Q_k = Q_n psup_k$, and the marginal revenue fare adjustment equation may be rewritten as:

$$f_k' = \frac{f_k psup_k - f_{k-1} psup_{k-1}}{psup_k - psup_{k-1}}$$

Because the application of this method transforms the original fare structure into a set of marginal revenues for each fare class, the fare products for each class may now be considered independent, enabling the correct use of RM optimizers. However, an important part of determining which fare products to have available remains a crucial step. Based on demand with an associated total revenue, a convex hull of feasible fare product choice sets is developed with the maximum capacity set as the constraint. The set of fare products with the maximum total revenue on the "efficient frontier," subject to the capacity constraint, is chosen (Fiig, Isler, Hopperstad, & Belobaba, 2010).

2.4 Chapter Summary

The evolution of airline revenue management is remarkable in that new methods and heuristics are constantly in development to adapt to the ever-changing airline environment. From Littlewood's Rule in 1972, to Belobaba's adaptation to a nested fare structure with the EMSRb heuristic, to network optimization techniques with DAVN and Bid Price Control, methods for seat allocation will never be perfect, but will enable airlines to further extract every dollar possible from their flight networks.

The rise of the LCC brought about a new problem of adapting to different fare structures. The use of Q and Hybrid Forecasting enables airlines to estimate passenger sell-up probabilities and adapt to both price-oriented and product-oriented demand. This, coupled with fare adjustment's ability to adapt to any fare structure, serve as crucial tools for today's airlines.

CHAPTER 3

SELL-UP ESTIMATION

3.1 The Importance of Forecast Accuracy

Weatherford and Belobaba suggest that forecast accuracy is vital to the performance of a revenue management system (Weatherford & Belobaba, 2002). Forecasts for future demand in current revenue management systems are not based on consumer choice models that would feature a price elasticity component; rather they are mostly based on exponential smoothing and moving average functions of previous demand for the same flight on the same day in earlier weeks and similar seasons. With the demand forecast having both a mean and a standard deviation input for the basic EMSR models used widely throughout the airline industry, the consequences of inaccurate forecasting are very interesting.

Weatherford and Belobaba propose that overestimating or underestimating the forecasts in a simulated single flight leg environment using EMSRb does not produce symmetrical or proportional results. For a predominately business flight, revenue decreases as the overestimation error increases, which is largely due to closing lower fare classes too early, leaving too many seats open for too few business passengers. In addition, underestimating the demand forecast for a business flight causes too few seats to be available for business passengers, resulting in a larger loss of revenue than in the equivalent overestimation case.

The interesting results reside in the case of the leisure flight. As expected, underestimating the demand results in too few seats left for higher-paying passengers, with results similar, but less extreme to the business flight case above. However, overestimating the demand for the leisure flight creates cases where there was an increase in revenue. While one would expect overestimation to allow too many seats for non-existent higher-paying passengers, moderate (12.5%) overestimation forced many of the leisure passengers without seats to purchase the higher priced seats. This is a look into the idea of sell-up, the occurrence of passengers buying the next higher fare class than they originally intended because it was closed down (Weatherford & Belobaba, 2002).

3.2 Recent Research in Sell-up Estimation

Belobaba and Weatherford adapted Belobaba's EMSRb heuristic by incorporating sell-up into the model, with the goal of adjusting the EMSRb fare ratios to increase protection levels for higher classes (Belobaba & Weatherford, Comparing Decision Rules that Incorporate Customer Diversion in Perishable Asset Revenue Management Situations, 1996). This would increase revenues by accounting for sell-up potential into these higher classes. More specifically, the model is:

$$P(\pi_n) = \frac{R_{n+1} - R_{1,n} \times psup_{n+1,n}}{R_{1,n}(1 - psup_{n+1,n})}$$

where π_n is the protection level for class 1 to n, $P(\pi_n)$ is the probability of selling the π_nth seat in class n or higher, R_{n+1} is the revenue from the class below class n, $R_{1,n}$ is the weighted average for revenue for classes 1 to n, and $psup_{n+1,n}$ is the probability of sell-up from class n + 1 to n (Belobaba & Weatherford, Comparing Decision Rules that Incorporate Customer Diversion in Perishable Asset Revenue Management Situations, 1996).

In a simulation study, with a four airline competitive environment, the airline that switched to the EMSRb sell-up model from a base EMSRb model realized revenue gains of up to 1.8 percent. However, overestimates in sell-up created large revenue losses for the airline. Because of this, coupled with the fact that airlines would have to input the expected sell-up rates, the method was not used by many airlines.

Andersson discusses a research project at Scandinavian Airlines (SAS) that focuses on estimating passenger preference when their desired class is unavailable for a particular flight. Their model captures the three options for a passenger: choose a competitor, choose a different flight, or sell-up into a higher class on the same flight (Andersson, 1998). More specifically, SAS chose to apply a logit choice theory model to determine the probability of a passenger choosing flight i, class j from the choice set S, where flight k, and class l are additional classes available after flight i, class j are unavailable:

$$P(i,j|i,j \in S) = \frac{\exp(\beta x_{ij})}{\sum_{k,l \in S} \exp(\beta x_{kl})}$$

The β utility function parameters in the model above were determined from passenger behavior data and interviews conducted by SAS.

Using this choice model, it is possible to define two more probabilities:

$P(k,l|k,l \in S\_)$: This is the probability that flight *k*, class *l*, is chosen after flight *i*, class *j*, is closed.

$P(k,l|k,l \in S)$: This is the probability that flight *k*, class *l*, is chosen regardless if flight *i*, class *j*, is closed (all flights and classes open).

Using these probabilities, it is possible to determine a sell-up and recapture rate, or α_{ijkl} :

$$\alpha_{ijkl} = \frac{P(k,l|k,l \in S\_) - P(k,l|k,l \in S)}{P(i,j|i,j \in S)}$$

This α_{ijkl} denotes the probability of a passenger getting rejected from flight *i*, class *j*, and choosing flight *k*, and class *l*, combining both recapture and sell-up.

In addition, Andersson discusses the potential markets where the model will make the largest impact for either sell-up or recapture. When other carriers offer fewer flights for a certain market, SAS will have an advantage for sell-up, regardless of how many flights they offer. However, when SAS offers fewer flights than other carriers for a certain market, they will have less recapture.

The model then uses the probability α that a passenger will sell-up to determine booking class closure rules: The marginal expectation claims that the airline should be indifferent between class 1 and class 2 purchases if the class 2 revenue equals the expected revenue from class 1. For a two class system, with the net revenue denoted as r_1 and r_2, for classes 1 and 2, respectively, the "optional" net revenue for class 2 is:

$$r_2 = r_1 \left[\alpha + (1-\propto) \int_{L_{1(\propto)}}^{\infty} p_C(x;T)dx \right]$$

The expectation of net revenue for retaining a seat in class 2 is the revenue for class 1 multiplied by the probability it actually happens. This is the sell-up probability plus the probability of no sell-up times the probability of at least a certain number (L_1) of class 1 passengers requesting a class 1 booking over the time, T. Therefore, if $\alpha > f_2/f_1$, then all class 2 booking requests should be rejected, for it is more likely that either the passengers will sell-up or there will be enough class 1 passengers to create more revenue.

Unlike Andersson, Talluri and Van Ryzin created a dynamic programming based optimization model to capture sell-up and buy-down behavior (Talluri & Van Ryzin, 2004). They claim that there is no complete revenue management methodology that contains a complete and correct passenger choice decision model, or more specifically, models passenger choice as a function of all of the available fare classes. This is an extremely difficult concept because it is impossible to observe no-purchase decisions from viewing the booking history for a particular flight. It is important to note that their model applies to a single leg, and is very complex, using a multinomial logit model. Applying this methodology to a network level would be computationally expensive, so approximation methods must be developed. However, their application of consumer choice in revenue management is an important alternative approach to addressing the sell-up and buy-down issues.

Like Zeni, Ratliff and Vinod suggest that in order for RM systems to combat restriction-free pricing, they must consider consumer choice behavior and essentially develop a demand curve in order to have optimal booking class closure times. Ratliff says that there is an abundance of data from online booking sources in the form of requests, which whether or not they actually result in a booking, still represent a source of unconstrained demand (Zeni, 2007), (Ratliff & Vinod, 2005). If gathered correctly, this would be more valuable than the current method of using historical bookings, which fails to represent a true notion of uncensored demand due to zero no-purchase data, such as bookings lost to competition or from fare class closure. In addition to a simple demand model, this online data may be used in a consumer choice model, analyzing factors such as price and cross elasticities, sell-up, buy-down, and recapture possibilities.

3.3 The FRAT5

In order to evaluate a passenger's willingness-to-pay, some revenue management models make use of an estimate of a passenger's likelihood to sell-up based on a given fare ratio, which is known as the FRAT5, or the fare ratio between the lowest base "Q" fare and the fare in question at which 50% of passengers are willing to sell-up. Visually, this is much easier to describe in Figure 4.

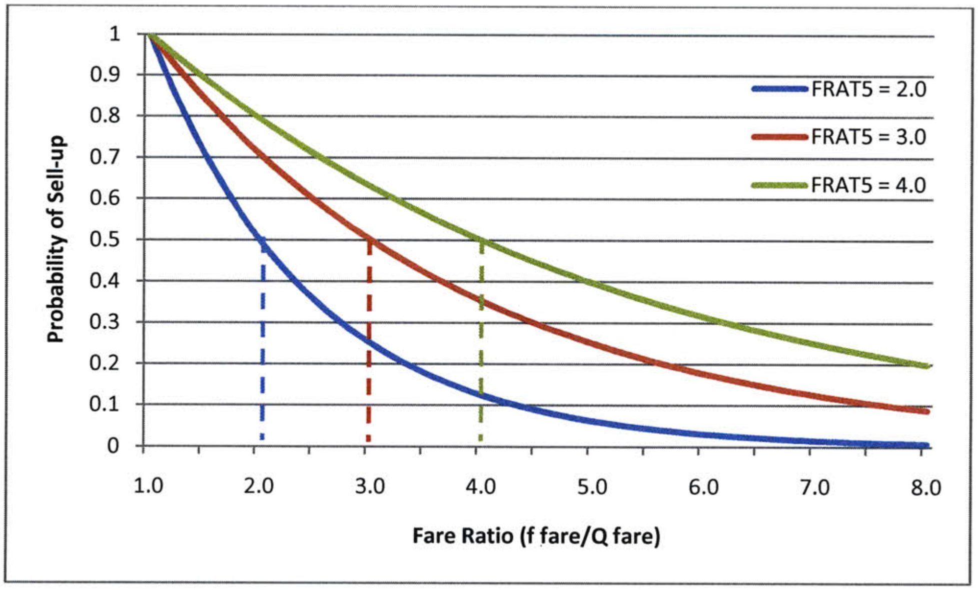


Figure 4: Development of FRAT5 Values

Figure 4 shows how FRAT5 values are obtained for one time frame in the booking process. The probability of sell-up ($psup$) from the base fare Q to a higher class *f*, is simply an inverse exponential function of the fare ratio between Q and *f*, as well as a sell-up constant ($supcon$), which includes the user input of the desired FRAT5 value, or fare ratio at which 50% of the passengers will sell-up into class *f*:

$$psup_{Q \to f}(fare_f) = e^{\left(-supcon \cdot \left(\frac{fare_f}{fare_Q} - 1\right)\right)}$$

Where:

$$supcon = -\frac{\ln(0.5)}{FRAT5 - 1}$$

$fare_f$: fare of the higher class, *f*

$fare_Q$: fare of the lowest class, *Q*

FRAT5: user input for the fare ratio at which 50% of the passengers will sell-up into class f from class Q

As one moves from left to right across Figure 4, increasing the FRAT5 values, passengers are predicted to be less price sensitive, or more willing to sell-up. Also, recall that Figure 4 represents FRAT5 values for just one time frame. As one moves later through the booking process, FRAT5 values will increase, for more passengers with a higher willingness-to-pay (i.e. business passengers) purchase their tickets closer to the departure date, and sell-up is more prevalent overall. To account for this, different series of FRAT5 values were created to model passenger behavior based on the time frame. Three of these FRAT5 series, known as "input FRAT5s" throughout this thesis, are shown in Figure 5.

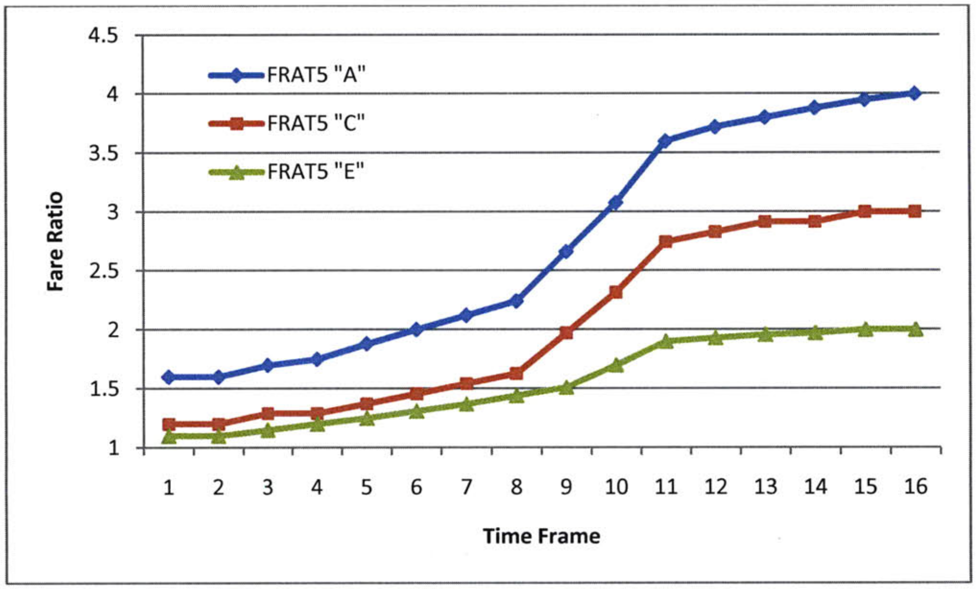


Figure 5: Input FRAT5 Curves

In Figure 5, note that a higher FRAT5 curves implies more aggressive passenger sell-up behavior. For example, at time frame 16, FRAT5 "A" has the highest value at 4.0, meaning that passengers booking in that time frame will sell up 50% of the time when the fare ratio equals 4.0. However, FRAT5 "E" only assumes that 50% of passengers will sell up when the fare ratio is 2.0, which is much less aggressive than FRAT5 "A." See Michael Reyes' thesis for more information on FRAT5s and traditional sell-up methodology (Reyes, 2006).

3.4 Methods to Estimate Sell-up

Sell-up focuses primarily on price-oriented behavior. Therefore, in a completely unrestricted fare environment, every passenger is considered price-oriented, for there are no restrictions partitioning the demand. All passengers will want to buy the lowest open fare class. In a semi-restricted fare structure, those passengers whose decisions do not depend on restrictions are considered price-oriented. All price-oriented passengers will try to buy the lowest fare class, or class Q, when possible. If the Q-class is closed, the price-oriented passenger may pay more for the next available fare class above Q (there may be more than just Q closed), as long as it does not exceed their predetermined maximum willingness-to-pay. If the passenger buys a fare higher than Q, this is considered sell-up.

As mentioned previously, sell-up may be estimated by an input FRAT5 curve. However, even though these curves are based on various airline data, they are not determined from the historical data at hand from the booking process. Therefore, estimating sell-up instead of using an input FRAT5 should provide a more robust solution applicable to any situation where enough sell-up exists. Using the historical bookings, there currently exist two methods to estimate sell-up, as described in detail in Charles Guo's thesis and presentation (Guo, Estimation of Sell-up Potential in Airline Revenue Management Systems, 2008), (Guo, Review: Methods for Estimating Sell-up Potential, 2008). A summary of the two methods to estimate sell-up are provided in the following sections.

3.4.1. Direct Observation

Suppose you are given s samples of data in a given time frame, where each sample refers to a recorded number of bookings in the lowest open fare class. For example, if there are only bookings in M-class for a given sample, then M was the lowest open class.

Sample Bookings

Y			3			
B		8				
M				10		14
Q	20				28	

Number of Samples	6

Table 3: Sample Data for Sell-up Estimation Examples

Using this data, one can develop sell-up estimates with the following methods.

The Direct Observation (DO) estimation method is the simpler of the two methods, focusing solely on the average number of bookings per fare class. Using the example established in Table 3, the following shows the application of the DO method.

	Average	Sell-Up Probability
Y	3	0.1250
B	8	0.3333
M	12	0.5000
Q	24	1.0000

Table 4: Direct Observation Estimation Example

Given s samples, for classes 1, 2,…, Q, the DO method uses the average ($b_{f,s}$) bookings to determine the sell-up probabilities, $p_{f,s}$, for each class. The sell-up probabilities are simply the ratio of the average number of bookings per class to the number of Q class bookings:

$$p_{f,s} = \frac{b_{f,s}}{b_{Q,s}}$$

3.4.2. Forecast Prediction

The Forecast Prediction (FP) method is slightly more complex than the DO estimation method.

	Average	Total	Initial Sell-Up Probability	Total Q Bookings	Revised Sell-Up Probability
Y	3	3	0.20	15	0.1463
B	8	8	0.40	20	0.3902
M	12	24	0.60	40	0.5854
Q	24	48	1.00	48	1.0000

Average Q	20.5

Table 5: Forecast Prediction Estimation Example

In summary, given *s* samples, for classes *f* = 1,2,..., Q, the first step for FP estimation is to find the average ($b_{f,s}$) and total bookings for each fare class *f* and number of samples *s*. Then, select an arbitrary initial sell-up estimate from Q to *f* for each fare class for only the first iteration of the estimation process. Given *s* samples, FP uses the sell-up probability from the previous number of samples, $p_{f,s-1}$, to scale the current number of bookings in each class *f*, or $t_{f,s}$ to the equivalent number of bookings for Q. Then, the sum of the scaled number of Q bookings is divided by *s* to produce the average forecasted booking values for Q:

$$b_{Q,s} = \frac{\sum_{i=1}^{Q}(t_{f,s})(p_{i,k-1})}{s}$$

Last, the probability of selling up from Q to another fare class *f* after a given number of samples is:

$$p_{f,s} = \begin{cases} \frac{b_{f,s}}{b_{Q,s}}, & f \neq Q \\ 1, & f = Q \end{cases}$$

3.5 Fitting Time Frame Sell-up Estimates

The average sell-up estimate per time frame (a single FRAT5 value), created from the sell-up probabilities obtained in one of the above methods, may serve as inputs for an entire FRAT5 curve over a given number of time frames. However, this method does not guarantee that the FRAT5 curve will be smooth, monotonically increasing, lie in a particular range of values, or have a sell-up estimate in every time frame. In order to convert the data-based sell-up estimates per time frame into a FRAT5 curve, a data fitter, or smoothing technique, may be applied to the time frame FRAT5 estimates. (In the following example, 16 time frames are used). In this thesis, both a logistic fitter and regression-based cross-time frame fitter are applied to the data.

Logistic Fitter

For the logistic fitter, the goal is to minimize the sum of the squared differences between the developed FRAT5 curve and the actual sell-up estimate per time frame. Both a two parameter and three parameter version of the cross-time frame fit will be used.

For the two parameter version, pick $x1$ and $x2$ such that

$$\sum_{tf}\left(f5tf_{tf} - f5le_{tf}\right)^2 \text{ is minimized}$$

Where: $f5tf_{tf}$ = actual sell-up estimate per time frame *tf*

$f5le_{tf}$ = logistic equation estimate, time frame *tf*

$= frat5n + \frac{frat5x - frat5n}{1+e^{-x1(tf-x2)}}$

$frat5n$ = min estimated frat5 (hard-coded to 1.1)

$frat5x$ = max estimated frat5 (input)

Furthermore, for the three parameter version, pick $x1$, $x2$, and $x3$ such that

$$\sum_{tf}\left(f5tf_{tf} - f5le_{tf}\right)^2 \text{ is minimized}$$

Where: $f5tf_{tf}$ = actual sell-up estimate per time frame *tf*

$f5le_{tf}$ = logistic equation estimate, time frame *tf*

$= frat5n + \frac{x1}{1+e^{-x2(tf-x3)}}$

$frat5n$ = min estimated frat5 (hard-coded to 1.1)

The basis for the minimum and maximum values is to produce FRAT5 estimates that are monotonically increasing over the 16 time frames, below a value of 2.0 for time frame 1, and to have a difference of at least 1.0 between time frame 1 and time frame 16. In addition, some constraints are necessary to obtain a reasonable fit. When looking at a given market, for a single time frame, across the historical observations (26 departure days in this case), a FRAT5 value may be determined only if there were at least two occurrences of sell-up to a higher fare class. In addition, for a given market, there must be FRAT5 values in at least four of the 16 time frames for a logistic cross-time frame fit to occur. If a market does not receive a logistic fit, then there were simply not enough occurrences of sell-up, and an alternate method must be applied, such as giving the market an input FRAT5 curve.

Consider the following example for a given market with FP estimated FRAT5 values over the 16 time frames. Applying a two parameter logistic cross-time frame fitter to this data produces the following results, with values for the two parameters: $x1 = 0.3131, x2 = 14.7337$

Time Frame	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
FRAT5 value	1.20	1.23	1.24	1.28	1.33	1.35	1.40	1.45	1.57	1.80	2.00	2.32	2.51	2.88	3.19	3.35

Table 6: Sample FP Estimated FRAT5 Values

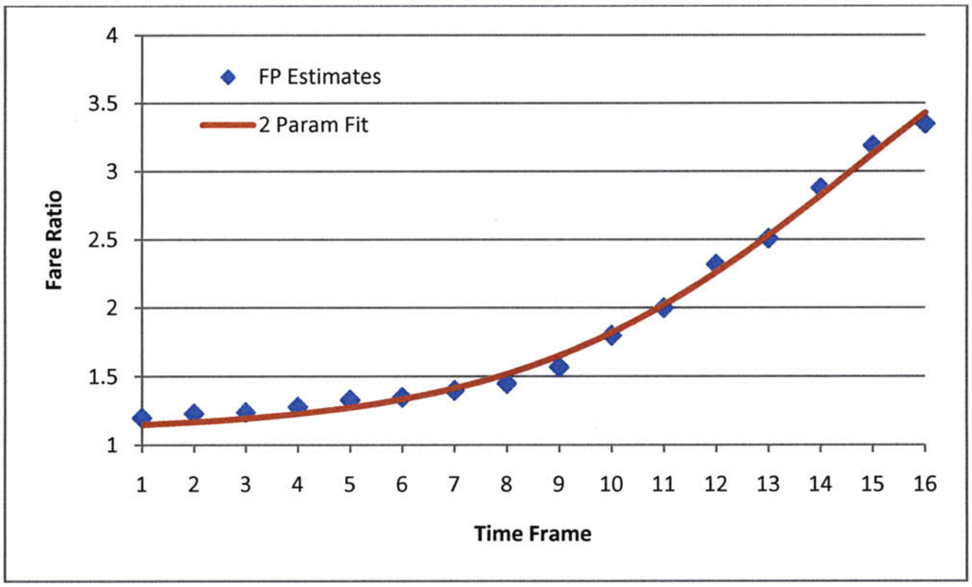


Figure 6: 2-Parameter Logistic Fit Applied to FRAT5 Data

Note that the shape of the fitted FRAT5 curve depends solely on the values of the two or three parameters from the logistic smoother. To distinguish the differences between markets, or to determine their similarity in estimated sell-up, these parameters are perfect to provide a basis for comparison, instead of comparing 16 different sell-up estimates from the 16 time frames. This will be of great importance in Chapter 5.

Regression Fitter

For the regression-based cross-time frame fitter, a slightly different method is used. Instead of directly smoothing the FRAT5 estimates, the methodology begins with the average elasticity constant per time frame, $econ_{tf}$.

The elasticity constant per time frame is the average of the elasticity constants for each combination of fare classes where sell-up may occur. The elasticity constant for sell-up from class k to class j is:

$$econ_{k\to j} = -\frac{\ln(psup_{k\to j})}{frat_j - frat_k}$$

Where:

$psup_{k\to j}$ is the probability of sell-up from fare class *k* to class *j*, and

$frat_j$ is the fare ratio to the lowest fare class for class *j*.

Next, a linear regression is performed, with the goal of picking *bint* and *bslope*, such that

$$\Sigma_{tf}\left(bint + bslope \cdot tf - econ_{tf}\right)^2 \text{ is minimized.}$$

The estimated average elasticity constant, $bint + bslope \cdot tf$, may then be used to create a FRAT5 estimate for the time frame:

$$frat5_{tf} = \frac{-\ln(0.5)}{bint + bslope \cdot tf} + 1$$

3.6 Chapter Summary

This chapter on sell-up estimation lays the groundwork for the new research presented in this thesis. The differences between price- and product-oriented passengers is evident in all sell-up models and serves as a basis for examining the effects on passengers in various markets with different fare structures. The implementation of sell-up adjusted forecasts walks a very fine line. If the forecast is too heavy towards the lower classes, and does not account for buy-down in unrestricted fare structures, the airline's revenues are subject to spiral down. In addition, if the forecasts favor the high-yield passengers too much and are wildly inaccurate, revenues will drop as too many seats will fly empty. Recent research in the sell-up sector of revenue management has driven the field to pursue a sensible, but simple, approach to estimating sell-up. Several questions remain, such as what model serves as the best approach, and on what level should the sell-up estimates be determined and aggregated.

One method of estimating passenger willingness-to-pay, the FRAT5, is a crucial element that aids in developing many sell-up models. Whether or not the FRAT5 is used directly in the model, in most cases the FRAT5 may be reported in order to give a sense of the aggressiveness of the model's sell-up estimates. Recent development of the two methods to estimate sell-up will propel their use in conjunction with various other revenue management tools in several competitive environments to determine their applicability in the real world.

CHAPTER 4

RATIONAL CHOICE (RC) FORECASTING

4.1 Introduction

Rational Choice (RC) Forecasting was developed by Miller, Zawack, and Schrag from Northwest Airlines as an alternative to standard forecasting, Q forecasting, and Hybrid Forecasting (Kayser, Belobaba, & Hopperstad, 2008). The basis for RC forecasting is that it removes the complexity of the Hybrid and Q Forecasting and creates its estimates from the historical bookings in one step, based on a linear regression. This is much faster and simpler than the four-step process required for the aforementioned Q and Hybrid Forecasting, where sell-up probabilities were needed to convert historical bookings into Q-equivalent bookings, followed by detruncating, forecasting, and using the sell-up probabilities again to repartition the Q-equivalent bookings into demand by fare class. Avoiding these steps removes the difficulties of storing all of the data as well as reducing the chance of making an error with fewer steps. The beauty of Rational Choice is that sell-up is already incorporated in the forecast just by the structure of the linear regression, and may be determined by looking at the segmented demand forecast. (There is no external FRAT5 necessary.) The following sections further describe the RC Forecasting process and methodology.

4.2 Methodology

Rational Choice Forecasting creates partitioned passenger type forecasts irrespective of the assumption that passengers arrive in an inverted willingness-to-pay order, as required by other forecasting methods. The observed historical bookings serve as the basis for partitioning the demand into different passenger categories. An "observation" in the following discussion is the event where the following data was recorded: the lowest open fare class, the fraction of the time frame that the lowest open class was available, and the observed bookings during this time in Y, B, M, or Q classes (from highest to lowest fares). The RC Forecasting methodology is summarized in the following flowchart:

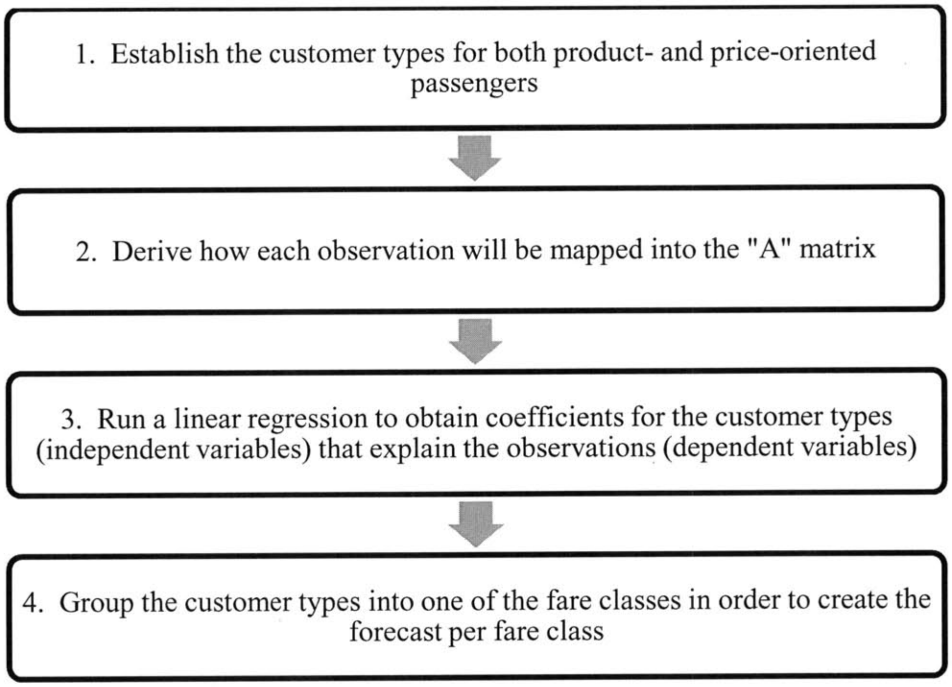


Figure 7: Rational Choice Process

4.2.1. Customer Types

Passenger categories are established based on their potential booking behavior. The first division is between product-oriented (business) and price-oriented (leisure) travelers. The second division establishes customer types within product- and price-oriented passengers. Product-oriented passengers all have a desired fare class, based on a set of restrictions and their maximum willingness-to-pay (WTP). Price-oriented passengers are assumed to all have a maximum willingness-to-pay, set at a particular fare class.

Product-Oriented			Price-Oriented			
wtp Y won't B/M/Q	wtp B won't M/Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M
1						
	1					
		1				
1			1			
	1		1	1		
		1	1	1	1	
			1	1	1	1

Booking In:	Lowest Open Class:
Y	B, M, or Q
B	M or Q
M	Q
Y	Y
B	B
M	M
Q	Q

Y
B
M
Q

Table 7: Customer Types for RC Forecasting

For example, consider the four-class fare structure in Table 7, with Y, B, M, and Q fare classes established in decreasing fare price. In order of increasing willingness-to-pay, there exist four price-oriented passenger categories: "wtp Q, not M"; "wtp M, not B"; "wtp B, not Y"; "wtp Y." Therefore, if a booking occurs in Q class when Q is the lowest open class, it could have been any of the four price-oriented passenger types because all will be willing to pay at least Q. This is denoted by the "1" in Table 7, meaning that given the lowest open class, a purchase in that booking class could have been from any customer type marked with a "1."

A similar process holds true for the product-oriented passengers. For the four class example, there are three categories of product-oriented passenger types—the first of which are those business passengers that will buy Y because they are willing to pay Y, but will not fly in B, M, or Q, because of various restrictions associated with those lower fare classes (denoted as "wtp Y, won't B/M/Q" in Table 7 above). This pattern applies to the two other product-oriented passenger types—those that are willing to pay for class B, but won't fly in M or Q classes, and those that are willing to pay for class M, but won't fly in class Q. A business passenger is assumed to not want the base Q fare. However, they could want the lowest fare available as long as that fare is above the Q fare.

This means that if the lowest open fare is B, M, or Q, and a passenger books in Y, then it must have been a "wtp Y, won't B/M/Q" product-oriented passenger. However, if Y is the lowest open fare class, and there is a booking in Y, it could be the product-oriented "wtp Y, won't B/M/Q" passenger, or the price-oriented "wtp Y" passenger. This possibility of a booking being either price or product-oriented is also true for other cases where B or M is the lowest class open. Therefore, this effectively removes the assumption that if there is a booking in the lowest open fare class it must be a price-oriented passenger, which is inherent to the Q and Hybrid Forecasting models.

In addition to the basic customer types and lowest open class scenarios shown in Table 7, each customer type is color-coded to denote the class to which customer type belongs. The key assumption is that if the customer type can book in more than one class based on their willingness-to-pay, the forecast for that customer type is allocated to the highest class possible for them. For example, if a price-oriented "wtp Y" customer books in M class because M is the lowest open class, the observation will be allocated towards the Y total forecast. Step 4 of the RC process in Section 4.2.4 further explains how the forecasts for each customer type are combined into actual forecasts for the fare classes.

4.2.2. Mapping the Observations

The premise of RC forecasting is to keep all options of different customer types available for who made a specific booking. Therefore, the next step is to connect each booking observation to all possibilities of customer types depending on the lowest open class during the time frame when the booking observation occurred. The process, referred to as "mapping" observations to customer types, expands Table 7 for all lowest open class possibilities and is shown in Table 8 below:

	Product-Oriented			Price-Oriented				
Open Classes	wtp Y won't B/M/Q	wtp B won't M/Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M	Booking In:
Y	1							Y
B		1						B
M			1					M
Q				1	1	1	1	Q

	Product-Oriented			Price-Oriented				
Open Classes	wtp Y won't B/M/Q	wtp B won't M/Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M	Booking In:
Y	1							Y
B		1						B
M			1	1	1	1		M

	Product-Oriented			Price-Oriented				
Open Classes	wtp Y won't B/M/Q	wtp B won't M/Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M	Booking In:
Y	1							Y
B		1		1	1			B

	Product-Oriented			Price-Oriented				
Open Classes	wtp Y won't B/M/Q	wtp B won't M/Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M	Booking In:
Y	1			1				Y

Table 8: Mapping Customer Types

This table may be read as follows: Looking at the scenario where M is the lowest open class, if there is an booking observation in class M, then it could have been any of the three price-oriented customer types who had a willingness-to-pay of at least M, or the product-oriented customer type who was willing to pay for M, but wouldn't fly in Q class. This chart lays the foundation for the "A" matrix, which then serves as the basis for linear regression.

Recall that each overall observation consisted of the set of bookings in Y, B, M, and Q, given that a certain class was the lowest open. These observations serve as the dependent variables for the regression, and the transpose of the observation vector will match up accordingly with the corresponding case in the mapping charts in Table 8. Consider the following example in Table 9 of three observations, showing the lowest open class, the fraction of the time frame that that was the lowest open class, and the observed bookings within each observation.

			Observed Bookings			
Observation	Lowest Open	Fraction TF Open	Y	B	M	Q
1	Q	1	0	1	0	2
6	B	0.5	1	1		
24	M	1	0	1	1	

Table 9: Sample Set of Observations

The next step is to map this sample set of observations into the "A" matrix using Table 8 from above to determine all possible customer types per observation given a particular lowest open class. The customer type possibilities, originally shown by "1"s, are weighted by the amount of the time frame that the particular class was the lowest open.

Observation	Observed Bookings	wtp Y won't B/M/Q	wtp B won't M/Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M
1	0	1	0	0	0	0	0	0
	1	0	1	0	0	0	0	0
	0	0	0	1	0	0	0	0
	2	0	0	0	1	1	1	1
6	1	0.5	0	0	0	0	0	0
	1	0	0.5	0	0.5	0.5	0	0
24	0	1	0	0	0	0	0	0
	1	0	1	0	0	0	0	0
	1	0	0	1	1	1	1	0

Table 10: "A" Matrix Derived from Sample Observations

The "A" Matrix above serves as the basis for the linear regression, with the goal of determining the best coefficients for each customer type to predict the observed bookings for each observation.

4.2.3. Linear Regression

The third step of RC Forecasting uses a linear regression to minimize the sum of squared residuals between each observed number of bookings within an observation and the estimated number bookings in each class determined from the regression coefficients (forecast values per class type) and the values for the independent variables. In math, the generalized formulation is:

Let: neq = number of observations ($i = 1..neq$)
$nfcls$ = number of fare classes ($j = 1..nfcls$)
lof_i = lowest available class, observation i
$flof_i$ = fraction of time frame that class lof_i was open for observation i
$blof_i$ = observed bookings, class lof_i, equation i

For each observation i and fare class j, define the least squared parameters as:

$A_{i,j} = flof_i$ given that $lof_i \leq f$, where f is the fare class in question, otherwise $A_{i,j} = 0$
$B_i = blof_i$

Determine the forecast for each class, fc_j, that minimizes the sum of squared differences between the estimated and observed bookings:

$$\underset{fc_j}{argmin}\left[\sum_i^{neq}\left(\sum_j^{nfcls} fc_j \cdot A_{i,j} - B_i\right)^2\right]$$

$$subject\ to: fc_j \geq 0\ \forall j$$

The explanatory variables in the linear regression are the passenger type possibilities available for the observation, which are also weight-adjusted by the fraction that the lowest open class was available during the time frame. For the four-class example above, the regression equation is:

$$Obs\_Booking = \beta_1(wtp\ Y, won'tB/M/Q) + \beta_2(wtp\ B, won't\ M/Q) + \beta_3(wtp\ M, won't\ Q) + \beta_4(wtp\ Y) + \beta_5(wtp\ B, not\ Y) + \beta_6(wtp\ M, not\ B) + \beta_7(wtp\ Q, not\ M)$$

$$subject\ to:\ \beta_i \geq 0\ \forall i$$

Note that the regression coefficients must be greater than or equal to zero because they serve as the actual forecasts for each customer type.

In the generalized form, additional information about each forecast may be determined, to include the forecasting error (Kayser, Belobaba, & Hopperstad, 2008). This helps to determine if some fare classes receive more accurate forecasts than others.

For the lowest available fare class k, define variance in terms of the residual, *resid*, from the least squares fit as:

Let: $neqk_k$ = number of equations where k is the lowest open class

$$var_k = \frac{\sum_{i|lof_i=k}^{neq} resid_i^{\,2}}{neqk_k}$$

The forecasting error, or variance, for a particular class f is:

$$fce_f = \sum_{k=f}^{nfcls} var_k \frac{fc_f}{\sum_{j=1}^{f} fc_j}$$

4.2.4. Forecast Allocation

The last step of the RC process aggregates the β coefficients by mapping each coefficient to the highest possible class associated with the passenger type as discussed in the first step of the RC methodology. The sum of the regression coefficients for each passenger type belonging to a certain fare class constitutes the total forecast for that class.

Observation	Observed Bookings	wtp Y won't B/M/Q	wtp B won't M/Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M	Estimated Bookings	Observed Bookings	Squared Error
1	0	1	0	0	0	0	0	0	0.6283	0	0.3948
	1	0	1	0	0	0	0	0	0.7179	1	0.0796
	0	0	0	1	0	0	0	0	0.9617	0	0.9249
	2	0	0	0	1	1	1	1	3.1770	2	1.3854
6	1	0.5	0	0	0	0	0	0	0.6283	1	0.1381
	1	0	0.5	0	0.5	0.5	0	0	1.6830	1	0.4665
24	0	1	0	0	0	0	0	0	0.6283	0	0.3948
	1	0	1	0	0	0	0	0	0.7179	1	0.0796
	1	0	0	1	1	1	1	0	1.9268	1	0.8590

Coefficients	0.6283	0.7179	0.9617	0.9651	0	0	2.2120

Total Forecast

Y	1.5934
B	0.7179
M	0.9617
Q	2.2120

Table 11: Rational Choice Forecast Allocation

Table 11 shows how the forecasts are generated from the linear regression in the four-class example. Note that some passenger types received a "zero" forecast, which means it was constrained to be greater than or equal to zero in to the Rational Choice formulation.

4.3 RC Expanded

The previous Rational Choice method assumed seven passenger types, three of which were product-oriented. However, the product-oriented passengers were assumed to want exactly one fare class and nothing else based on the set of restrictions associated with the fare. However, the RC Expanded methodology presents an alternative approach and assumes that a product-oriented passenger may still be sensitive to price and accept more than just one fare class option. For the four-class example, three new product-oriented passenger types are introduced in addition to the existing three: "wtp Y, won't M/Q"; "wtp Y, won't Q"; "wtp B, won't Q." For example, a product-oriented "wtp Y, won't M/Q" passenger with a willingness-to-pay set at Y may now purchase a Y or B class fare, instead of just the Y fare in the previous Rational Choice formulation. The new set of customer types for RC expanded is:

Product-Oriented						Price-Oriented						
wtp Y won't B/M/Q	wtp Y won't M/Q	wtp Y won't Q	wtp B won't M/Q	wtp B won't Q	wtp M won't Q	wtp Y	wtp B not Y	wtp M not B	wtp Q not M	Booking In:	Lowest Open Class:	
1										Y	B, M, or Q	
	1		1							B	M or Q	
		1		1	1					M	Q	
1	1	1				1				Y	Y	Y
	1	1	1	1		1	1			B	B	B
		1		1	1	1	1	1		M	M	M
						1	1	1	1	Q	Q	Q

Table 12: Passenger Types for Rational Choice Expanded Forecasting

In general, the RC Expanded method creates many more product-oriented passenger types as the number of fare classes increases. For every additional fare class, creating a new total of *n* fare classes, this creates *n-1* additional passenger types. While this four-class example only has six product-oriented passenger types, a 26-class fare structure will have 325 product-oriented passenger types, in addition to the four price-oriented passenger types.

The remaining steps of the RC Expanded methodology, to include the customer type mapping, linear regression, and customer type forecast allocation, follow the same process as the basic Rational Choice method. It is important to note again that the forecast allocation still operates under the assumption that the forecast for the customer type is allocated to the highest possible class that the customer type in question may purchase. Referring to Table 12 above, for the "wtp Y, won't Q" product-oriented passenger, their forecast is allocated to Y because their maximum willingness-to-pay is Y, even though they may purchase a B or M fare.

4.4 RC with Fare Adjustment

In order for a revenue management optimization method to use adjusted fares, estimates of sell-up are necessary, as discussed in section 2.3.3. However, instead of getting sell-up estimates from an input FRAT5 or a FRAT5 estimation method (FP or DO), one may use the Rational Choice price-oriented partitioned forecasts, making Rational Choice Forecasting with fare adjustment a very simple process. Only the price-oriented passenger types are used because sell-up only occurs with price-oriented passengers.

The price-oriented passenger types for both the basic and expanded Rational Choice formulations represent the partitioned forecasts. To create estimates of sell-up between each possible fare class pairing, it is necessary to aggregate the partitioned forecasts into willingness-to-pay categories:

$$\left.\begin{array}{l}\left.\begin{array}{l}\left.\begin{array}{l}\left.\text{wtp Y}\right\}\text{wtp Y}\\ \text{wtp B, not Y}\end{array}\right\}\text{wtp B}\\ \text{wtp M, not B}\end{array}\right\}\text{wtp M}\\ \text{wtp Q, not M}\end{array}\right\}\text{wtp M}$$

Using these non-partitioned forecasts, the sell-up estimate from fare class k to fare class j, $psup_{k\to j}$, is created for all possible sell-up pairs. For n fare classes, there are $\binom{n}{2}$ sell-up combinations. For example, the six sell-up pairs for this four class example are:

Q→Y: (wtp Y) / (wtp Q)
Q→B: (wtp B) / (wtp Q)
Q→M: (wtp M) / (wtp Q)
M→Y: (wtp Y) / (wtp M)
M→B: (wtp B) / (wtp M)
B→Y: (wtp Y) / (wtp B)

Next, each estimate of sell-up is converted into an elasticity constant, or $econ_{k\to j}$:

$$econ_{k\to j} = -\frac{\ln\left(psup_{k\to j}\right)}{frat_j - frat_k}$$

where $frat_k$ is the fare ratio to the lowest fare class (Q in the example) for class k.

The average elasticity constant for the sell-up combinations for the time frame in question, $econ_{tf}$, may be used to create a FRAT5 estimate for the time frame. Even though this estimate is not used in the fare adjustment process, it is still worthy of noting to compare sell-up probability estimates between different forecasting methods:

$$frat5_{tf} = \frac{-\ln(0.5)}{econ_{tf}} + 1$$

The purpose of creating the average elasticity constant is to aggregate the $\binom{n}{2}$ sell-up estimates for the fare class combinations to just n sell-up estimates between the lowest base fare class (Q) to each of the other n-1 fare classes in question, which is what the fare adjustment method uses. Therefore, using the average elasticity constant for the time frame and the fare ratio between class Q and class k, the probability of sell-up between class Q and class k, $psup_{Q\to k}$, is:

$$psup_{Q\to k} = \begin{cases} e^{-\left(econ_{tf}\cdot\left(\frac{fare_k}{fare_Q}-1\right)\right)}, for\ k \neq Q \\ 1, \quad for\ k = Q \end{cases}$$

The last step of the fare adjustment process using Rational Choice forecasts is to create the adjusted fares using the probabilities of sell-up for each class k. The adjusted fare for each class k, f_k', is:

$$f_k' = \frac{f_k psup_k - f_{k-1} psup_{k-1}}{psup_k - psup_{k-1}}$$

Consider a single departure, whose booking process is divided into the 16 time frames that span a given booking period. Because of variable estimates of $psup_{Q\to k}$ between each departure per time frame, two smoothing techniques were applied to the sell-up estimates before they were used to create the adjusted fare per time frame.

The first method places the sell-up estimates into a historical database and uses a moving average of the sell-up estimates per time frame per class over the past 26 samples. The sell-up estimates for this method are averaged on a path basis. The second method, or the "market" method, uses the same moving average smoother, but applies it to all paths associated with a market, thus aggregating the sell-up estimates from the "path" method.

4.5 Chapter Summary

Rational Choice Forecasting serves as a very logical and sensible method to create class forecasts. Avoiding the multi-step processes inherent to Q-and Hybrid Forecasting, as well as the assumption that passengers arrive from lowest to highest willingness-to-pay order, Rational Choice uses booking observations based on conditional reasoning and just one linear regression to create a forecast. While other methods incorporate separate estimates of sell-up into their model, all occurrences of sell-up are already covered by the observations used and passenger types created within the Rational Choice forecasts.

The most difficult aspect of Rational Choice Forecasting is to account for all possible passenger behavior types. Establishing the customer types enables the mapping of each observation, conditioned on the lowest open class, into the "A" matrix used in the linear regression. Using the observed bookings as the dependent variables, with the customer types as independent variables, the linear regression seeks to minimize the squared difference between the two by giving each customer type a regression coefficient. This coefficient, when allocated to a certain class, serves as the forecast for the given class.

Rational Choice Expanded Forecasting does not change the basic methodology of Rational Choice Forecasting, but just creates additional customer types based on the assumption that product-oriented customers may too have a "price-oriented" aspect of being sensitive to price and be willing to fly in multiple fare classes given a maximum willingness-to-pay. However, the total number of product-oriented customer types grows non-linearly with the number of classes used.

Another extension of Rational Choice Forecasting is its use with fare adjustment. Based on the price-oriented customer type structure, Rational Choice Forecasting is able to easily provide sell-up probabilities, an essential piece that the fare adjustment formula uses to create the adjusted fares for an RM optimizer. Additionally, using an elasticity constant created from the sell-up probabilities between combinations of classes, a FRAT5 value may be reported for each time frame. This is important for reporting purposes and will be referenced in later portions of the thesis.

CHAPTER 5

CLUSTERING

5.1 Introduction

In Section 3.4, two methods were discussed that estimate sell-up based on historical bookings for a given time frame. These methods, Direct Observation and Forecast Prediction, create sell-up estimates that may be applied to a forecasting method to create forecasts by willingness-to-pay for a given market. However, a question of the level of aggregation remains. Should sell-up be estimated on a per-market basis, or should it be estimated over the whole system, or network? (Recall that the input FRAT5 curves were applied over the whole system.) Both market-level and system-wide aggregation methods were recently developed and in testing, created revenues comparable and often slightly better than those of an input FRAT5, with system-based estimation performing better than input FRAT5s and market-based estimation.

However, some fundamental issues may exist with both system-wide and market-based sell-up estimation. Despite having more data over the entire system, the system-wide estimates may be too broad, giving some markets sell-up estimates that are uncharacteristic of that market. On the other hand, giving a market its own specific sell-up estimate is much more difficult due to a shortage of sell-up occurrences over the booking process. In order to disaggregate the system-wide estimation, perhaps a more sensible option is to group, or cluster, similar markets that will have analogous sell-up properties and characteristics. Then, individual sell-up estimates for each market may be clustered, and the markets belonging to each cluster will all receive the same sell-up estimate. This provides a middle ground between the system-wide and per-market sell-up estimation used previously, creating a more robust solution.

In order to define "similarity" between sell-up estimates for markets, the parameters from the logistic cross-time frame fit smoother will be used, since they determine the overall shape of the FRAT5 curve. Therefore, instead of needing a sell-up estimate for each of the 16 time frames for a given market, just two or three parameters can define the shape of the FRAT5 curve. For a given number of clusters, those markets belonging to a specific cluster will all use the same parameter values from the cluster mean, and thus have the same FRAT5 curve. In order to determine which markets will be assigned to a given cluster, the K-means clustering algorithm will be used.

5.2 K-Means Algorithm

The purpose of any clustering method is to assign observations to clusters such that the sum of the pairwise dissimilarities between two observations in a given cluster is smaller than if the observations were in different clusters. According to Hastie et al., in every clustering algorithm there exists an encoder, $k = C(i)$, that assigns the i^{th} observation to the k^{th} cluster. The goal of clustering is to use the specific encoder to assign values to each pair of observations through a distance metric $d(x_i, x_i')$. In order to properly assign observations to clusters, one should seek to adjust the cluster assignments until a loss function is minimized. The loss function in this case determines the amount to which the overall clustering goal is not met, and is defined by

$$W(C) = \frac{1}{2}\sum_{k=1}^{K} \sum_{C(i)=k} \sum_{C(i')=k} d(x_i, x_i').$$

The K-means algorithm is a type of iterative descent method. Instead of trying every possible combination of cluster assignments for a given number of observations and clusters, an iterative descent method begins with an initial partition and changes the cluster assignments to improve the loss function in each step. Once the algorithm fails to improve the loss function, it terminates with the current cluster assignments as the solution. However, a potential pitfall occurs if these solutions converge at local minimum, but not the global minimum. Note that this is still more feasible than enumerating the total number of cluster assignments possible for N observations and K clusters:

$$S(N, K) = \frac{1}{K!}\sum_{k=1}^{K}(-1)^{K-k}\binom{K}{k}k^N$$

Having just 20 observations and 5 clusters yields 7.492×10^{11} possible combinations, making the iterative descent method much more efficient.

The K-means algorithm uses the squared Euclidean distance between observations as the dissimilarity metric, where p is the number of features, or dimensions, for each independent observation:

$$d(x_i, x_{i'}) = \sum_{j=1}^{p} (x_{ij} - x_{i'j})^2 = \|x_i - x_{i'}\|^2$$

There are two steps to the K-means algorithm, which uses an alternating optimization method between the two steps. First, in order to find the cluster means, $\{m_1, \ldots, m_K\}$, the algorithm seeks to minimize the total cluster variance for a given cluster assignment C. The total cluster variance is given by

$$\min_{C,\{m_k\}_1^K} \sum_{k=1}^{K} N_k \sum_{C(i)=k} \|x_i - m_k\|^2,$$

producing the means of the currently assigned clusters for a current set of observations S:

$$\bar{x}_S = \underset{m}{\operatorname{argmin}} \sum_{i \in S} \|x_i - m\|^2$$

Second, based the current set of means, one will minimize the total cluster variance by allocating each observation to the nearest current cluster mean, or mathematically,

$$C(i) = \underset{1 \le k \le K}{\operatorname{argmin}} \|x_i - m_k\|^2.$$

These two steps are repeated until the cluster assignments for observations to cluster means no longer change, indicating convergence to a minimum cluster variance (Hastie, Tibshirani, & Friedman, 2001).

In regards to the clustering application presented in this thesis, recall that the logistic-fit parameters, x1 and x2 (or x1, x2, and x3 for the three parameter version), define the shape of the FRAT5 curve for each market. Visually, each market's parameter values can be expressed in a scatter plot, as shown in Figure 8. Applying the K-means clustering algorithm essentially asks the question, "Given that you have to use k=3 clusters, where do you place the cluster centers in order to minimize the total squared Euclidean distance from the center of each cluster to each point assigned to the cluster?"

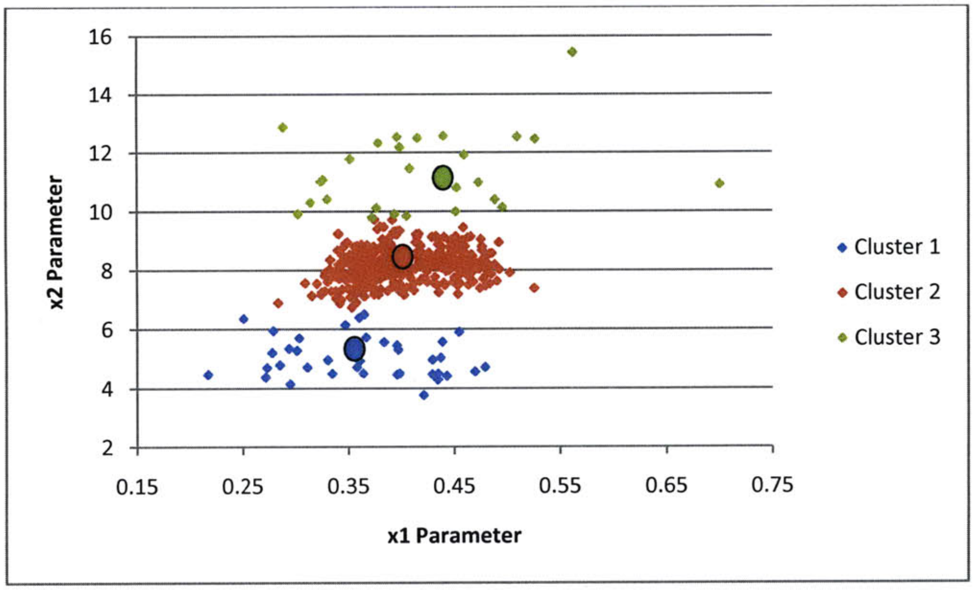


Figure 8: K-Means Clustering Example

In Figure 8, each market is represented by its x1 and x2 parameter. Those markets belonging to Cluster 1, as indicated by the blue markers, have a within sum-of-squares based on the Cluster 1's mean, as shown by the blue circle. The K-means algorithm seeks to minimize the total cluster variance, which is simply the sum of the three clusters' within sum-of-squares.

This algorithm works for different dimensions beyond the two-dimensional case above, and also for higher values of K. However, choosing the correct K value is not entirely intuitive.

5.3 Determining the Number of Clusters: Gap Statistic

With each additional cluster center, the total within sum-of-squared distances, denoted as W_K, will decrease. However, the decrease is non-linear, reaching a saturation point where every additional cluster provides minimal benefit. This is often denoted by an "elbow," or sharp curve in the data, as shown in Figure 9. The gap statistic is one recently developed approach that scientifically quantifies the degree of this "elbow" (Tibshirani, Walther, & Hastie, 2001).

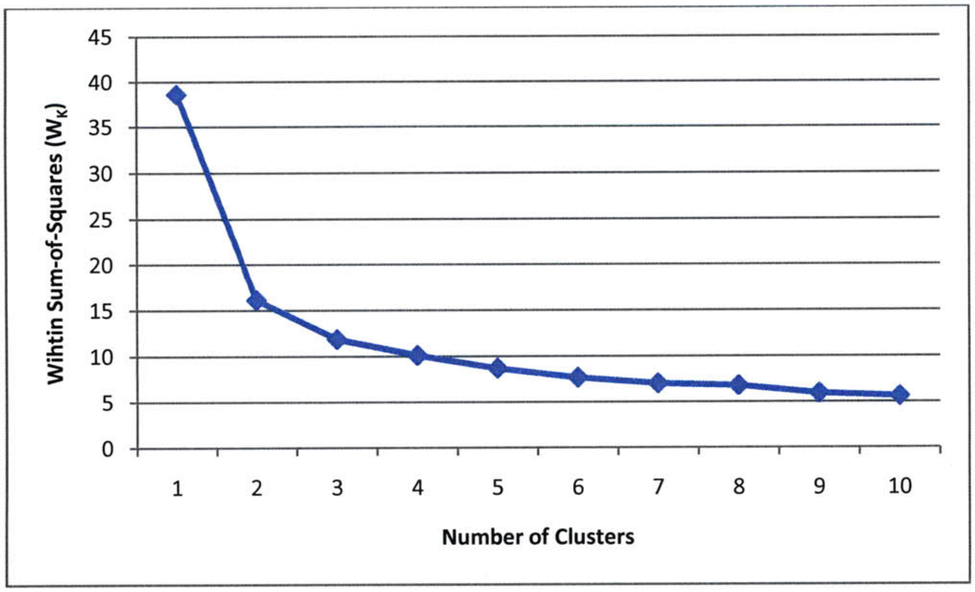


Figure 9: Within Sum-of-Squares Example

Figure 9 shows that there is minimal benefit gained after about three to five clusters, but it is difficult to determine exactly where a distinct elbow occurs beyond that of two clusters. The Gap Statistic Method's approach is to compare the log of the values of W_K to those from a reference distribution. In this case, a uniform distribution of the data is applied over the range of values covered by observations of interest. If this were used with x1 and x2 parameters, the reference uniform distribution would lie in a two-dimensional box. Next, the same clustering method is applied to the uniform distribution, and new W_K values are obtained. This data series is denoted as the "Expectation" for the within sum-of-squares, and is shown in Figure 10.

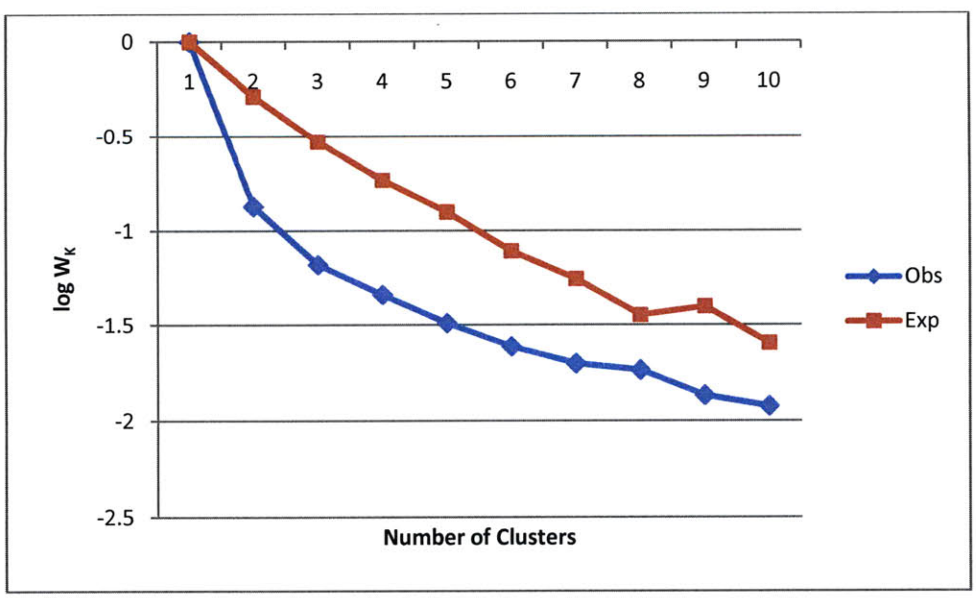


Figure 10: Observed and Expected Log (W_K) Example

The Gap Statistic Method sets the optimal number of K clusters where the observed WK falls farthest from the expected WK. In other words, the gap statistic is the difference between the expected and observed WK, with the optimal k set to its maximum value. As seen in Figure 11, for this example data, the Gap Statistic sets k=3 as the optimal number of clusters.

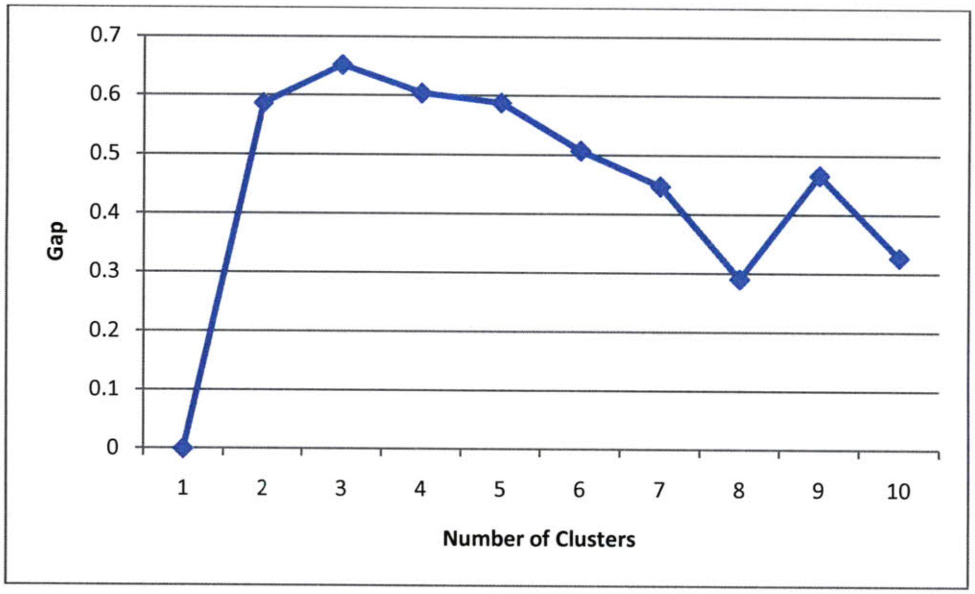


Figure 11: Gap Statistic versus Number of Clusters Example

It is also important to note that the gap statistic is not entirely convex, and is therefore important to look at the entire curve over all values of *k*. If there are other increases, or local maxima in the gap statistic curve, this may mean that there are smaller subclusters within larger, better-defined clusters. For example, if the data is grouped in such a way that there only appears to be one large clump of data, the gap statistic still offers a suggestion about where other divisions may be made.

While the gap statistic offers a sound and statistically-based method for determining the number of clusters, the goal of its use in regards to this thesis is simply that of determining the number of clusters that will maximize revenue. The gap statistic may offer a best guess at an appropriate number of clusters—however, there is no guarantee that the optimal *k* value will actually produce the best revenue. The experiments that test these thoughts are included in the following chapter.

5.4 Chapter Summary

With sell-up estimation playing an increasingly pivotal role in airline revenue management, it is vital to determine proper estimates for a given market. However, the over generalization of a system-wide or network-based estimate does not distinguish sell-up estimates for specific markets. In contrast, estimating sell-up on a market basis results in over-specification and often too sparse of a sell-up data set across all time frames. The K-Means clustering algorithm enables the grouping of similar observations based on a certain number of defining characteristics. In this thesis, those characteristics are the parameters that define shape of the FRAT5 curve for a given market. When clustering on these parameters, markets that have similar sell-up characteristics will be grouped, and will all receive the same parameters from the cluster mean. This greatly increases the distinction of sell-up estimates compared to a system-wide estimation method, and increases the number of sell-up data points for a given cluster, where the market-based estimation method fails. While there is no clear way to predict what exact number of clusters will result in the highest revenue, various methods such as the gap statistic provide a good starting point. To determine the effectiveness of these clustering methods and their variants in different environments, the following chapter uses the Passenger Origin-Destination Simulator (PODS) to evaluate their performance.

CHAPTER 6

SIMULATION OF FORECASTING, SELL-UP ESTIMATION, AND CLUSTERING METHODS

6.1 The Passenger Origin Destination Simulator (PODS) General Background

Developed in the mid-1990s at Boeing by Hopperstad, Berge, and Filipowski, the Passenger Origin-Destination Simulator (PODS) creates a competitive environment to test airline revenue management systems and strategies. With multiple airlines serving numerous markets, PODS is capable of simulating airline competition over a multiple-day booking period with several different networks. There are many components in PODS, and only a brief overview is presented below. For more information, see (Tam, Belobaba, & Hopperstad, 2008), (Hopperstad, 2005), or many other MIT PODS-based theses and dissertations.

6.1.1. Passenger Choice Model and Revenue Management System Components

At the heart of PODS lies the Passenger Choice Model, which simulates passenger preference and bookings that feed into the Revenue Management System. The RM System includes a booking database, forecaster, and RM optimizer.

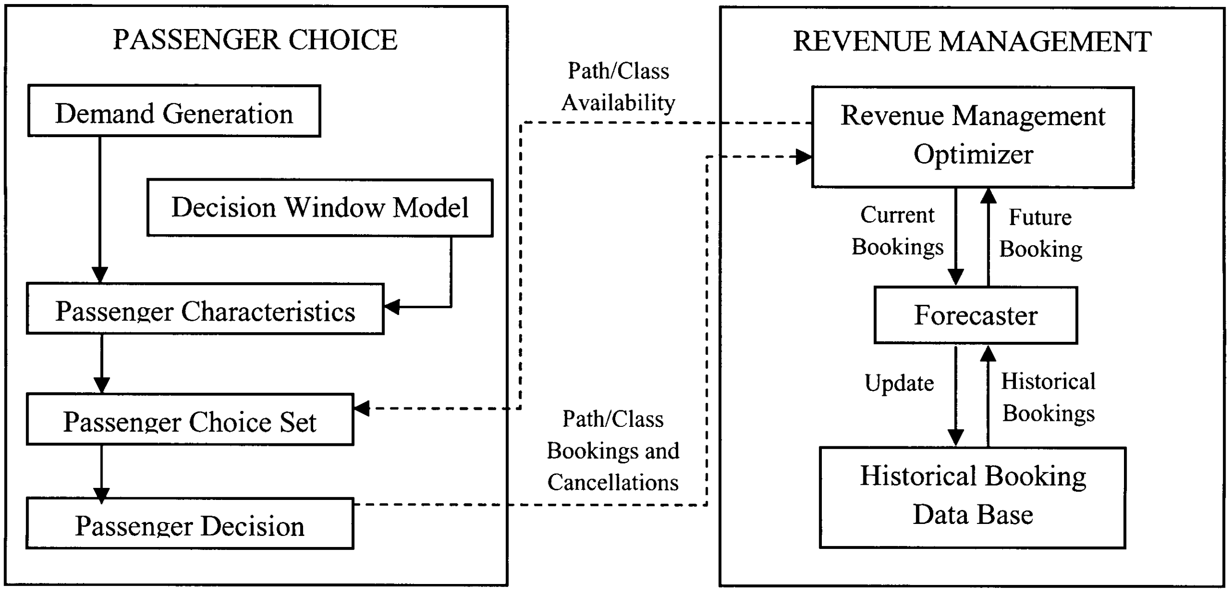


Figure 12: PODS Structure (Tam, Belobaba, & Hopperstad, 2008)

In the Passenger Choice Model, the first step for PODS is demand generation. The demand by market was developed by PODS Consortium members to reflect real world markets. In addition, the demand is divided between leisure and business passengers, roughly based on a 65% leisure to 35% business passenger mix. However, different simulation networks used in PODS may have slightly different passenger type proportions for various markets.

For every passenger generated, he or she is assigned passenger characteristics beyond that of their desired origin and destination, to include a time of travel preference, a maximum willingness-to-pay based on a demand curve for both business and leisure passengers, and a disutility value for various restrictions and limitations for the fare structure set in PODS. For the departure time parameter, the Boeing Decision Window Model is used to determine whether or not a path fits within each passenger's window. If not, the excluded path class option will receive an additional disutility.

Based on all of the passenger requirements, PODS determines the complete available set of fare options to each passenger with all disutilities included. However, if the passenger fails to meet other requirements such as advance purchase or maximum willingness-to-pay, or if the RM system of the airline closed down the desired class, the particular fare is unavailable to the passenger. Last, the passenger decides on the lowest-cost feasible option. This is the fare with the best total value to the passenger—that is the lowest fare price plus fare restriction disutility plus path quality disutility (for non-stop versus connecting paths). This simulated purchase then becomes a historical booking.

The interest of this thesis lies in the forecasting block in the RM system side of PODS. The forecasting section uses current and historical bookings to adjust the forecast for future bookings for the revenue management optimizer. There are numerous different methods of forecasting in airline revenue management, but through their simulation in PODS, it is possible to test their effects while holding all other conditions equal. The item of interest is the method for estimating sell-up (passenger willingness-to-pay), which will adjust the forecast fed into the RM optimizer.

Revenue management research greatly depends on creating controlled environments to test new methods and algorithms. Through the creation of these environments, PODS serves as the main testing ground for new methodology and analysis presented in this thesis. Specifically, multiple networks reflecting different fare environments (restricted and unrestricted) are used to evaluate the potential impact that new methods will have in the real world.

The PODS booking process for a single departure in our tests consists of a 63-day period, which is divided into 16 time frames. As the departure date nears, the length of each time frame shortens, from a duration of 7 days for Time Frame 1 beginning 63 days from the departure date, to a duration of one day for time frame 16, the day before departure. While booking limits are re-optimized before every time frame, passengers may book flights and airlines may close or reopen fare classes within each time frame.

Each 16 time frame booking process serves as one "sample," or departure. One "trial" in PODS consists of 600 samples, the first 200 of which are burned. This removes autocorrelation between samples that may exist due to initial conditions inherent to statistical methods used in the simulation. The overall results for the performance of each airline in the simulation are the average of the last 400 samples. Last, a typical PODS "run," consists of five trials for smaller, simpler networks, or two trials for more complex networks, with the performance for each airline defined as the average of the sample averages from each trial.

6.2 Simulation Environment

Various competitive airline networks can be used as the basic structure for setting up a simulation experiment in PODS. Because different revenue management methods are often designed to operate in variety of environments, it is useful to test them under multiple conditions. These situations are encapsulated by the numerous different controllable attributes for each network in the simulation. The most important factors to consider in the experiments are the airlines' fare structures with various restrictions, as well as the size and complexity of the network and markets served by the airline of interest. This thesis makes use of two airline network structures, known as Network D6 and Network T.

While it is possible to enumerate many possible fare structures in PODS, the thesis will focus on the choice of two fare structures for the airline of interest for both Network D6 and Network T.

6.2.1. Network D6 Semi-Restricted and Unrestricted

Network D6 is the more simple of the two networks used in the simulations. It is a dual airline competitive network, with each airline operating out of its own central hub, serving 40 spoke cities. Traffic flows only from west to east, and must connect through each hub, H1 for Airline 1 (AL1), or H2 for Airline 2 (AL2). There are no hub bypass, or point-to-point, services offered. Airline 1 legs are denoted by blue in Figure 13, with the red indicating those legs that belong to AL2.

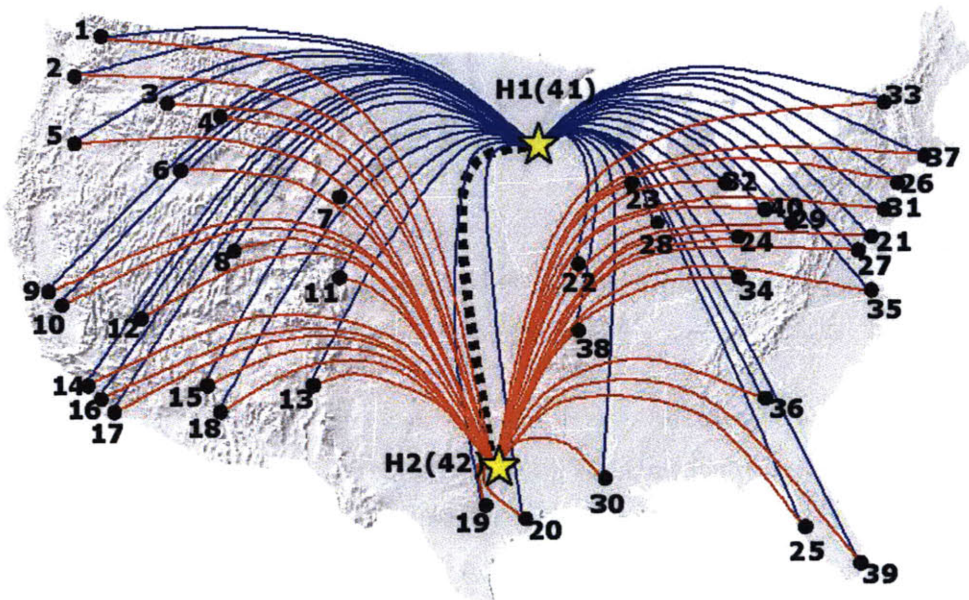


Figure 13: Network D6 Map

Throughout the PODS simulations, the experiments are conducted so that the airline of interest is AL1. Various revenue management methods are all tested on AL1, leaving AL2 as the control. With 40 spoke cities, AL1 has 126 legs, serving 482 O-D markets. In addition, AL1 offers six fare classes for all flights, which serve as the basis for distinction between the two Network D6s.

If both AL1 and AL2 offer a semi-restricted fare structure, the network is referred to as Network D6 Semi-restricted. Their fare structures are functions of whether or not advance purchase, Saturday night minimum-stay requirement, cancellation fee, or non-refundability is included for a specific fare class. The fare structure design is denoted in Table 13 below.

Fare Class	*AP*	*Min Stay*	*Cancel Fee*	*Non Refund*
1	0	NO	NO	NO
2	3	NO	YES	NO
3	7	NO	YES	YES
4	14	NO	YES	YES
5	14	NO	YES	YES
6	21	NO	YES	YES

Table 13: Network D6 Semi-restricted Fare Structure

In the baseline scenario in Network D6 Semi-restricted, both AL1 and AL2 use the EMSRb seat allocation heuristic with standard forecasting.

Contrary to the semi-restricted fare environment, both AL1 and AL2 may offer a fully-unrestricted fare structure. This network is referred to as Network D6 Unrestricted. As shown in Table 14, there are literally no differences in restrictions between the fare classes.

Fare Class	*AP*	*Min Stay*	*Cancel Fee*	*Non Refund*
1	0	NO	NO	NO
2	0	NO	NO	NO
3	0	NO	NO	NO
4	0	NO	NO	NO
5	0	NO	NO	NO
6	0	NO	NO	NO

Table 14: Network D6 Unrestricted Fare Structure

The baseline scenario for testing in Network D6 Unrestricted consists of AL1 using EMSRb (most likely with a form of Hybrid or Rational Choice Forecasting), while AL2 uses Adaptive Threshold Revenue Management, with a target load factor of 90 percent (AT90).

While Network D6 Semi-restricted is a more realistic representation of U.S. domestic fare structures, Network D6 Unrestricted captures the extreme case of the potential effects of spiral down. While this is an abnormal situation, where not even an advance purchase requirement segments demand, it is useful for the testing of revenue management methods focused on estimating passenger willingness-to-pay that seek to prevent spiral down. It is not completely unlikely, as these situations can arise in markets where a low cost carrier is present and offers extremely low-priced fares with virtually no restrictions.

Although Network D6 consists of only two airlines and two hubs, it serves as a good initial testing ground for revenue management methods. Its lack of complexity keeps competitive influences to a minimum, with AL1's performance more dependent on the methods that define its revenue management system.

6.2.2. Network T1 and Network T4

Network T serves as a more complex network for PODS simulations. Increasing the competition from Network D6, AL1 is now part of a four carrier network, serving 40 spoke cities. Each airline operates out of its own central hub (again traffic moving west to east), and each has its own baseline revenue management system. In the base case, AL1 is still the airline of interest, initially using a leg-based EMSRb revenue management system. Airline 2 (AL2) and Airline 4 (AL4) both use network-based DAVN, while Airline 3 (AL3) is an LCC, using Adaptive Threshold Revenue Management, with a target load factor of 90 percent.

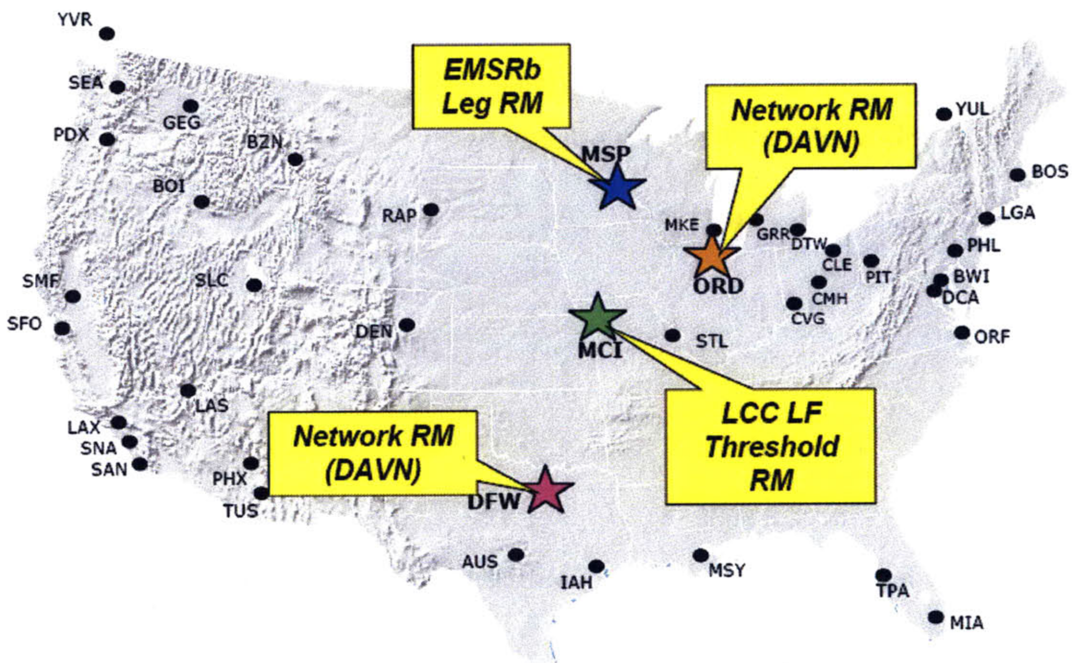


Figure 14: Network T Map

Out of the 572 markets created by the 40 spoke cities, the markets serviced by AL3 are known as LCC markets. Because of these markets, the other airlines operate with two fare structures, one for the 296 LCC markets and another for 276 non-LCC markets. Like Network D6, the fare structures will serve as the basis for distinction between the two versions of Network T.

In Network T1, the three airlines serving the non-LCC markets use a "More Restricted" Fare Structure as shown in Table 15.

Fare Class	*AP*	*Min Stay*	*Cancel Fee*	*Non Refund*
1	0	NO	NO	NO
2	3	NO	YES	NO
3	7	NO	YES	YES
4	10	YES	YES	YES
5	14	YES	YES	YES
6	14	YES	YES	YES

Table 15: More Restricted Fare Structure for Network T Non-LCC Markets

However, for all four airlines serving the LCC markets, a "Less Restricted" Fare Structure is used, as shown in Table 16 below.

Fare Class	*AP*	*Min Stay*	*Cancel Fee*	*Non Refund*
1	0	NO	NO	NO
2	0	NO	YES	NO
3	7	NO	NO	YES
4	7	NO	YES	YES
5	14	NO	YES	YES
6	14	NO	YES	YES

Table 16: Less Restricted Fare Structure for Network T1 LCC Markets

As an alternate to Network T1, in Network T4, LCC markets for all airlines have a fully unrestricted fare structure. It is important to note that the non-LCC markets will have the same fare structure that they had in Network T1, as shown in Table 15. The new unrestricted fare structure for Network T4's LCC markets is below in Table 17.

Fare Class	*AP*	*Min Stay*	*Cancel Fee*	*Non Refund*
1	0	NO	NO	NO
2	0	NO	NO	NO
3	0	NO	NO	NO
4	0	NO	NO	NO
5	0	NO	NO	NO
6	0	NO	NO	NO

Table 17: Unrestricted Fare Structure for Network T4 LCC Markets

In addition, there are other Network T characteristics worthy of noting. While the fare structures for non-LCC markets differ from LCC markets, fare ratios are also much different. With LCC markets ranging from \$105 for a Class 6 fare to \$366 for a Class 1 fare, the average fare ratio (highest to lowest fare) is equal to 3.5. However, in the non-LCC markets, fares differ from \$161 to \$804 for Class 6 and Class 1, respectively, creating an average fare ratio of 5.0.

Also, unlike Network D6, the business-leisure percent of passenger demand is not set to a constant 35-65 for all markets, respectively. Rather, cities are first classified as having high (90%), medium (60%), or low (40%) business demand. Then, depending on what two cities serve as the origin and destination, markets are classified into one of five percent business demand bins, creating a bell-shaped curve centered at a business demand of 36 percent, and ranging from 16 percent to 81 percent.

Looking at the four networks described, it is easy to notice that there are two distinct classifications of networks in terms of fare structure. Network D6 Semi-restricted and Network T1 both use fare structures that incorporate at least a few restrictions and advance purchase requirements in order to segment demand, with Network T1 being more complex and susceptible to many competitive feedback effects. Network D6 Unrestricted and Network T4 serve as the more extreme environments with fare structures that create a higher likelihood of spiral down. These two networks are better testing grounds for sell-up and passenger willingness-to-pay estimation. While it is impossible to create an environment completely reflective of a real-life network, these networks provide a reasonable, controllable platform for simulating airline revenue management methods, and statistical methods for sell-up rate estimation in particular.

6.3 Hybrid Forecasting with Data-based Sell-up Estimation

As presented in Section 3.4, data-based sell-up estimation methods provide a viable alternative to input FRAT5s. Both Direct Observation (DO) and Forecast Prediction (FP) make use of historical bookings in order to estimate sell-up probabilities for each class, which then lead to an average FRAT5 value for a given time frame for a market. Then, based on the FRAT5 values per time frame, either a logistic or regression-based cross-time frame fitter is applied. The FRAT5 curves may be determined on an aggregated system-wide basis, or kept at the single market basis, both of which are tested in this section. This section will set the baseline for testing the clustering process, which acts as the middle ground between system and market estimation.

The airline of interest for these experiments is Airline 1 (AL1), which will use the EMSRb seat allocation heuristic with Hybrid Forecasting, unless otherwise noted. In Network D6, AL2 uses EMSRb in the semi-restricted case or AT90 in the unrestricted scenario. In Network T, the competitor airlines' revenue management systems are described in Figure 14. If the experiment takes place in a network that has fully unrestricted fares, AL1 will use Q-forecasting instead of Hybrid Forecasting because no product-oriented demand exists with completely unrestricted fares. The simulation setup in PODS for testing in all networks is shown in Table 18 below.

	Aggregation Level	*Fitter*
FP	System	Regression
		Logistic
	Market	Regression
		Logistic
DO	System	Regression
		Logistic
	Market	Regression
		Logistic

Table 18: Data-based Sell-up Estimation Experiment Setup

Network D6 Unrestricted

In order to better understand the workings of sell-up estimation, Network D6 Unrestricted will serve as the first simulation environment. Much of the performance of AL1 in this network depends on sell-up estimation, which is clearly defined by the FRAT5 curves, as shown in Figure 15 and Figure 16 below.

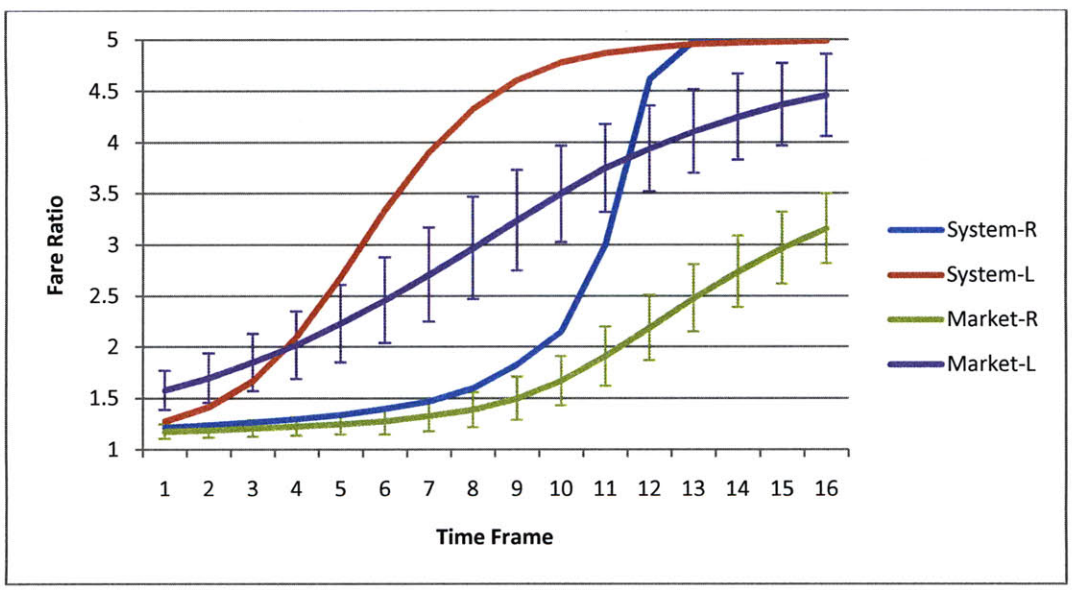


Figure 15: FP FRAT5 Curves for Network D6 Unrestricted

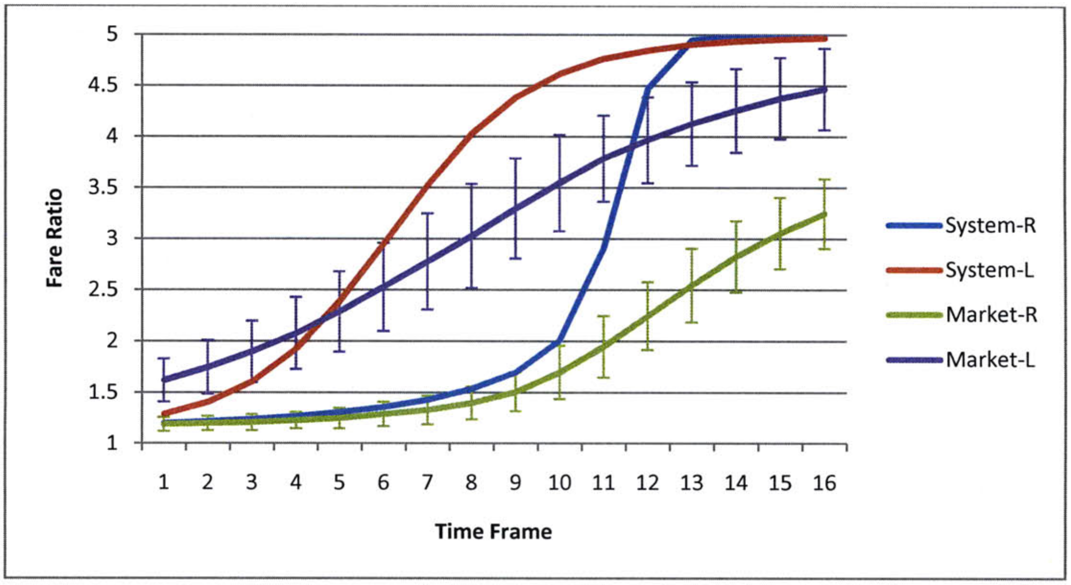


Figure 16: DO FRAT5 Curves for Network D6 Unrestricted

The solid curves without the error bars represent the FRAT5 curve developed over the entire system via both the logistic and regression cross-time frame fitters. The other two curves in each of the figures represent the mean of the FRAT5 curves developed for each of the 482 markets, with the error bar showing one standard deviation above and below the average. It is evident that estimating over the system creates much more aggressive FRAT5 curves, not only ending at a higher fare ratio, but also rising earlier and steeper than the market-based curves. This is largely due to the number of sell-up observations accumulated over the entire system as compared to a single market, in which fewer observations exist. Figure 17 provides good insights on how these curves impact revenue.

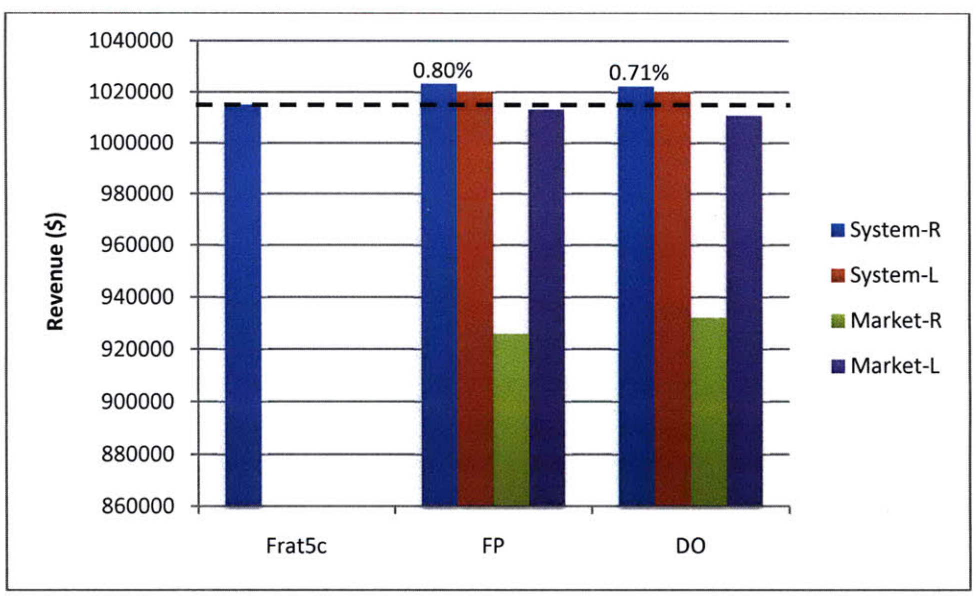


Figure 17: Data-based Sell-up Estimation Revenue in Network D6 Unrestricted

Compared to the input FRAT5c, the data-based estimation method over the whole system improves not only the revenue, but also the usability of the curve—it is more sensible to apply a curve that was developed from historical booking data rather than an arbitrary input FRAT5, especially if it improves the revenue. These revenue results also show that in this fare environment, lower, less aggressive FRAT5 curves perform worse. The market-based regression-fit curve performs the poorest out of all curves, about 8.5 percent worse than the input FRAT5c, and only reaches an average fare ratio of about 3.2. Meanwhile, the system regression and logistic-fitted curves perform remarkably well, staying at low FRAT5 values in early time frames and then rising to the maximum FRAT5 value of 5.0 in later time frames.

These trends are also evident in AL1's load factors and yields. Having a more aggressive FRAT5 curve implies that AL1 expects more passengers to be willing to sell-up to a higher fare class throughout the booking process. This will lead to more fare class closures beginning in earlier time frames and continuing throughout the booking process, causing a reduction in load factors and an increase in yield.

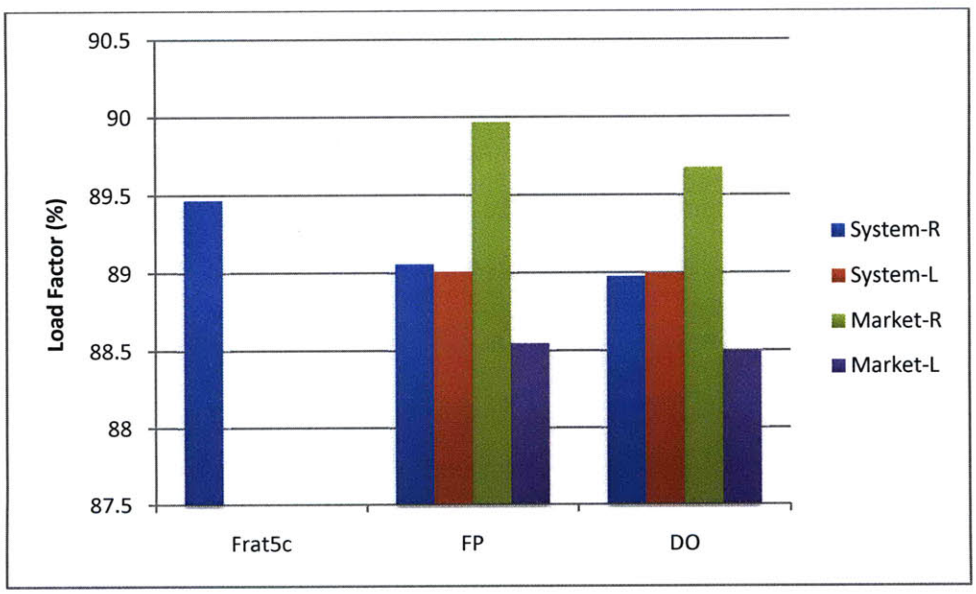


Figure 18: Data-based Sell-up Estimation Load Factors in Network D6 Unrestricted

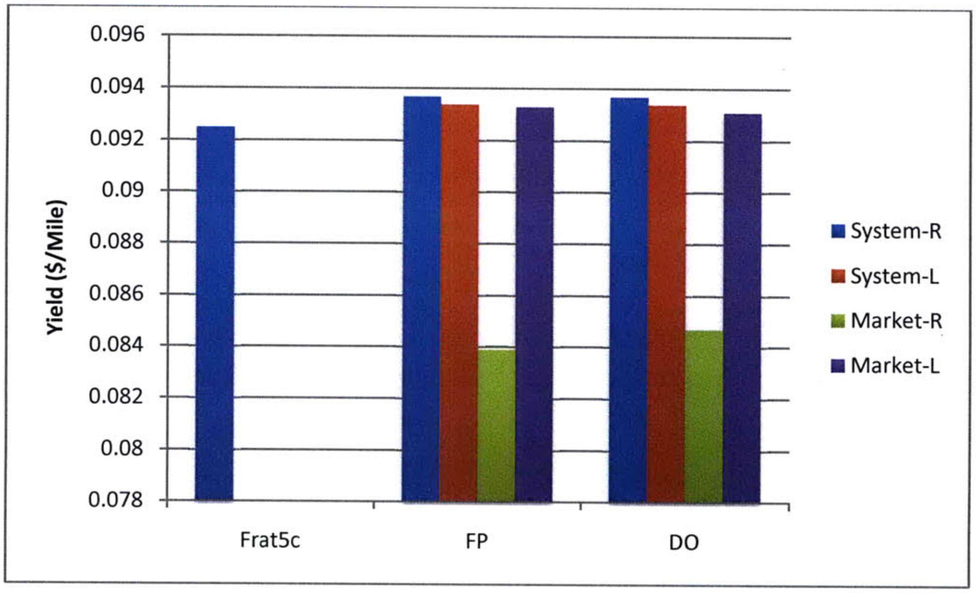


Figure 19: Data-based Sell-up Estimation Yields in Network D6 Unrestricted

Figure 18 and Figure 19 show that those methods with higher, more aggressive curves result in lower load factors with higher yields, as compared to market-based regression-fit case. To further analyze the effects of the FRAT5, consider a comparison of the fare class (FC) closure rates between the most aggressive curve (FP system-based logistic-fit) and the least aggressive curve (FP market-based regression-fit). Because the fare environment is fully unrestricted, without any advance purchase requirement, the FRAT5 curve's estimate of sell-up greatly impacts the class closure rates.

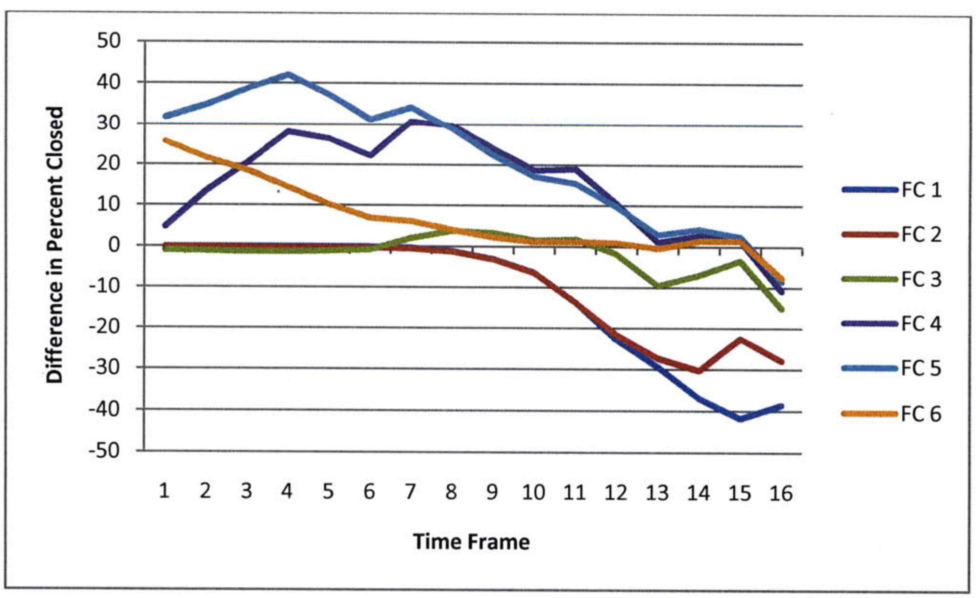


Figure 20: Difference in Fare Class Closure over Time: System-based Logistic-Fit minus Market-based Regression-Fit

To help visualize this, Figure 20 shows the difference in the percentage closed for each of the six fare classes over time. For positive values, the system-based logistic-fit has a greater percentage of the fare class closed. The more aggressive, system-based logistic-fit curve causes higher closure percentages in lower classes (FC 4, FC 5, and FC 6) compared to those of the market-based regression-fit. In later time frames, the market-based regression-fit has more of the higher fare classes closed. With only price-oriented demand, this results in a substantial loss in revenue, lower yield, and higher load factors when compared to the system-based logistic-fit case.

Another interesting impact of FRAT5 sell-up estimates is its ability to mitigate the harsh effects of spiral down, especially in an unrestricted fare environment like that of Network D6 Unrestricted. To visualize this, the forecast per path over the booking process from the baseline input FRAT5c serves as a good starting point.

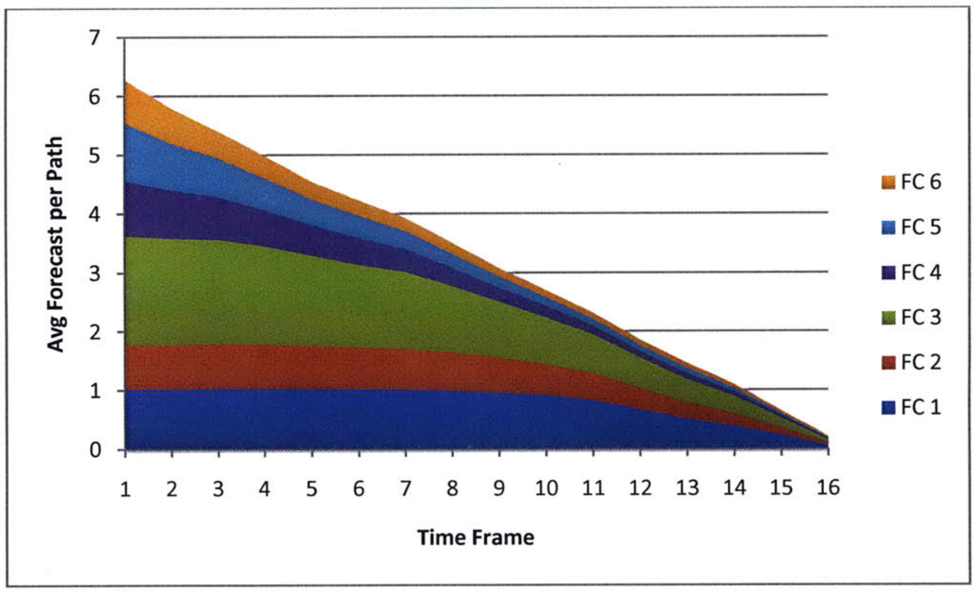


Figure 21: Average Forecast per Path: EMSRb with QF (FRAT5c) in Network D6 Unrestricted

Initially, Q-forecasting (QF) causes the average forecast per path distribution to place more weight on the middle to high fare classes. This is due to the FRAT5c curve suggesting that half of the people will be willing to sell-up to a higher fare class, up to three times the base fare, by the end of the booking process. To account for this, FC 6's forecast becomes smaller compared to what the actual fare class mix is by the end of the booking process. This is shown by the average cumulative bookings per path over the 16 time frames, as shown in Figure 22.

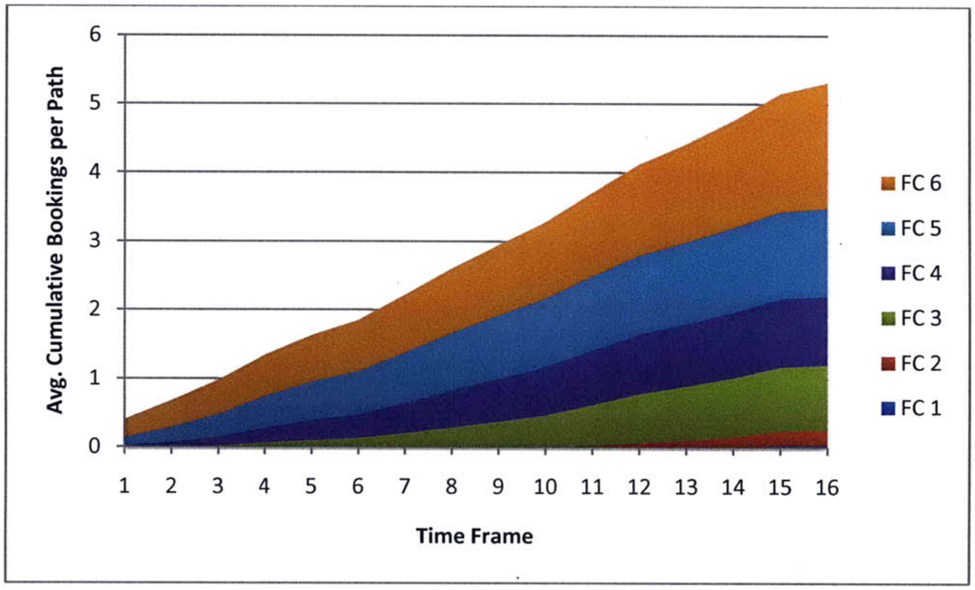


Figure 22: Average Cumulative Bookings per Path: EMSRb with QF (FRAT5c) in Network D6 Unrestricted

Despite having a large forecast for the middle to high fare classes, the fare class mix after 16 time frames is still dominated by the lower classes. However, keep in mind that there were no restrictions or advance purchase requirements to help segment demand, and therefore the sell-up estimate was the primary factor in keeping lower classes closed, forcing some bookings into higher fare classes. Figure 23 combines the average forecast with the average cumulative bookings over time, in order to create a picture of the evolution of turning a forecast into bookings.

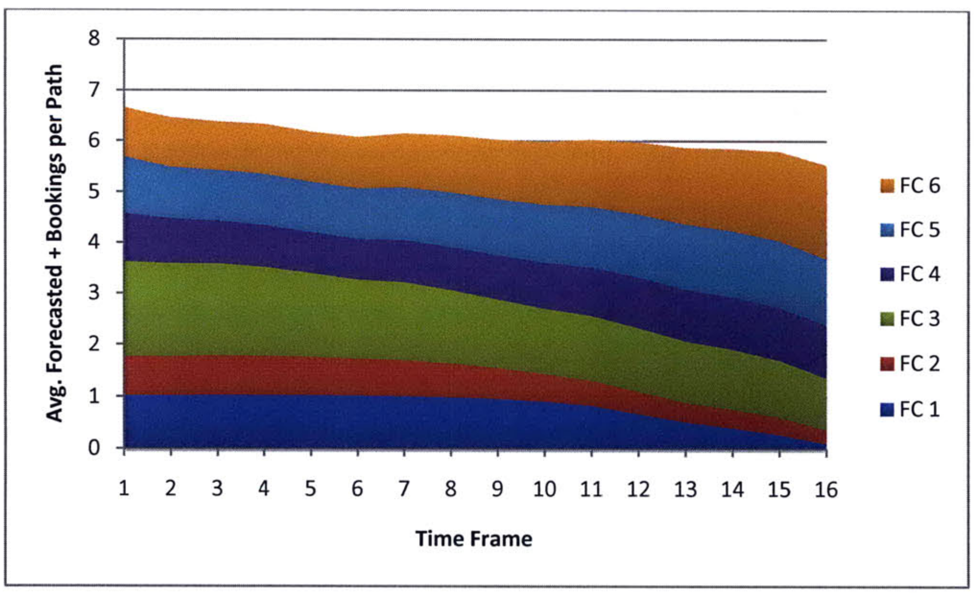


Figure 23: Forecast + Bookings per Path: EMSRb with QF (FRAT5c) in Network D6 Unrestricted

Despite the use of Q-forecasting, there still exists some evidence of spiral down. Overall, there is slight over-forecasting of total bookings; however it is important to see the growth of FC 5 and FC 6 from the initial forecast to the final bookings. Additionally, despite having a large forecast of one passenger per path in FC 1, the end share for FC 1 in the fare class mix is minimal. Using Figure 23 as a baseline, it is possible to see the positive effects of both Q-forecasting as a whole, as well as the use of the newer data-based sell-up estimation methods.

If AL1 only uses the EMSRb seat allocation heuristic, with no Q-forecasting, one should expect the complete spiral down of the forecast and bookings into all FC 6, as evident in Figure 24 below.

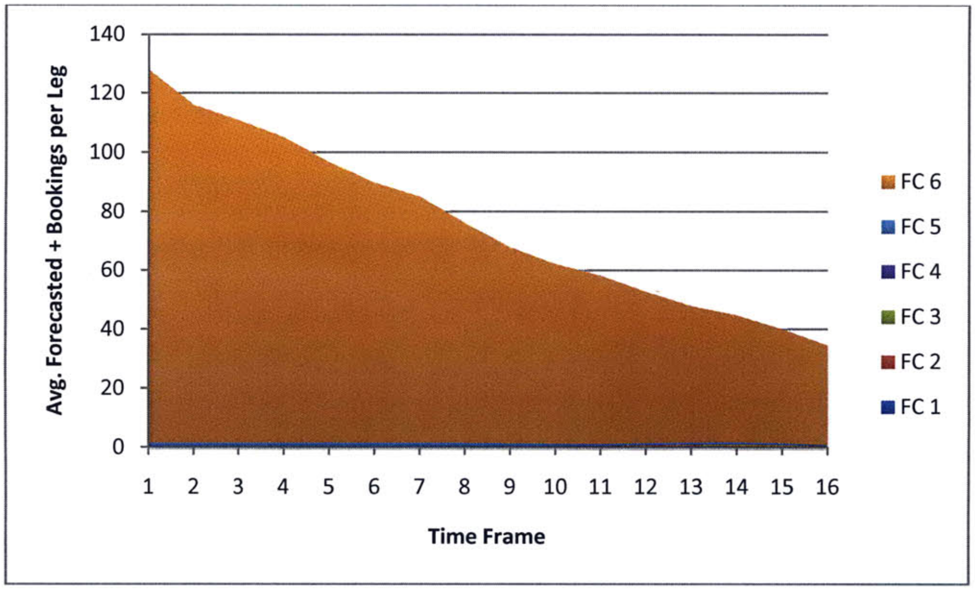


Figure 24: Forecast + Bookings per Leg: AL1 EMSRb with no QF in Network D6 Unrestricted

However, if AL1 employs the use of Q-forecasting with an even more aggressive FRAT5 curve compared to that of the input FRAT5c, then one should expect a higher forecast for the higher fare classes and an even greater prevention of spiral down. Figure 25 shows the forecasting and booking evolution from the use of the aggressive FP system-based logistic-fit curve.

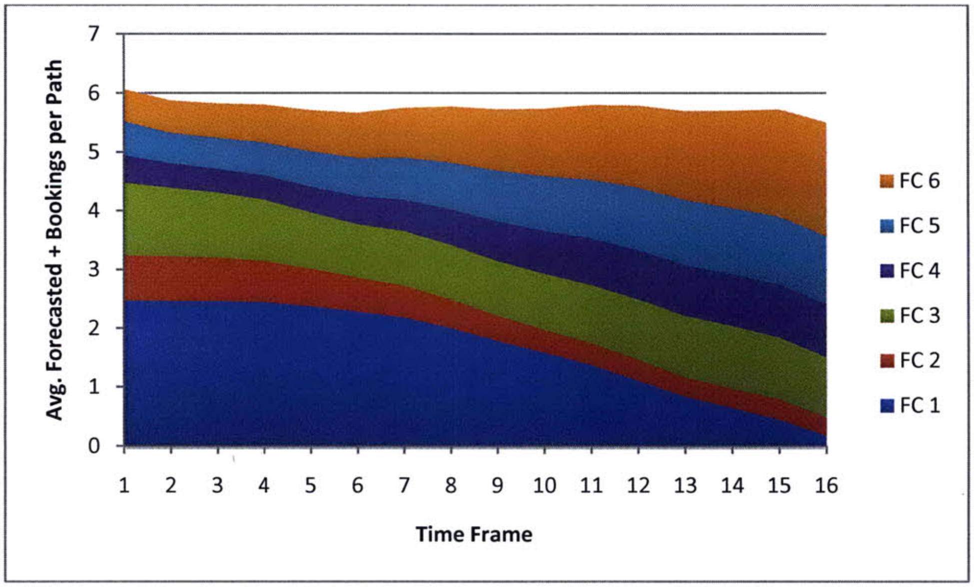


Figure 25: Forecast + Bookings per Path: AL1 EMSRb with QF (FP System Logistic-fit) in Network D6 Unrestricted

Compared to the input FRAT5c, the FP system-based logistic-fit curve created a much higher forecast for the higher fare classes. Maintaining these higher levels throughout the booking period caused the end fare class mix to include fewer lower class bookings than the input FRAT5c. Also, there is less evidence of spiral down, as well as over-forecasting compared the input FRAT5c.

While these data-based sell-up estimation methods worked well in Network D6 Unrestricted, they should also be tested in a more realistic and more competitive environment in Network T.

Network T4

Network T4 parallels Network D6 Unrestricted because a completely unrestricted fare structure is used for all LCC markets, while a more-restricted fare structure is used for the non-LCC markets. Airline 1 will again use EMSRb, now with Hybrid Forecasting instead of Q-forecasting because there now exist both price- and product-oriented demand. When AL1 uses system-based sell-up estimation, it distinguishes between non-LCC and LCC markets, creating a single FRAT5 curve for each, as shown in Figure 26 for the FP case.

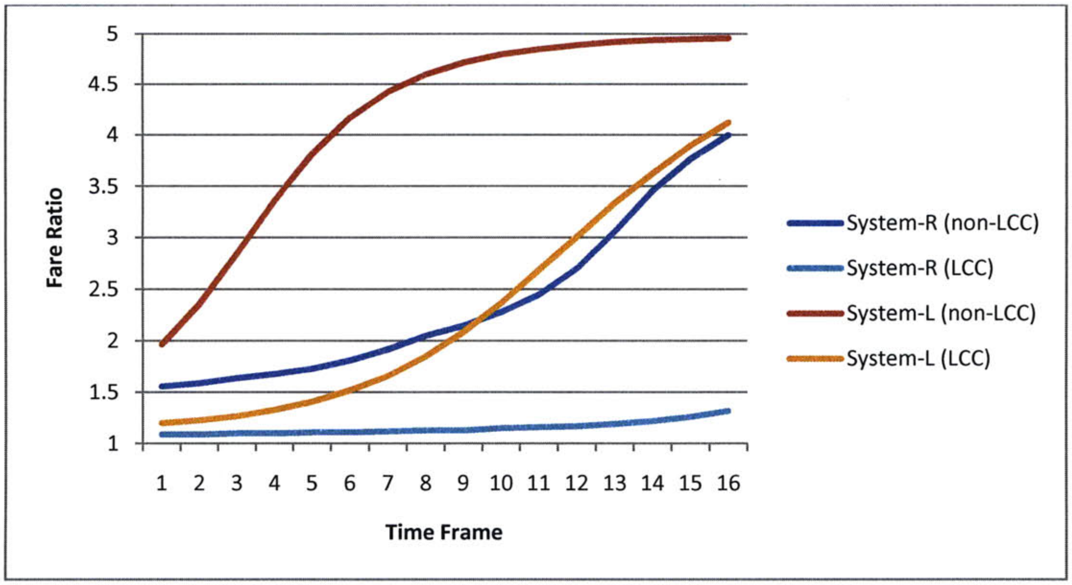


Figure 26: FP System-based FRAT5 Curves in Network T4

The system-based logistic-fit curves are much more aggressive than the regression-fit curves for both the non-LCC and the LCC markets. Also, it is worthy to note that this method indicates a higher likelihood of sell-up in non-LCC markets indifferent of the method. In Network T, recall that non-LCC markets have a average fare ratio of 5.0 while LCC markets have a average fare ratio of 3.5, which means that sell-up to higher fare ratios (above 3.5) will exist much more in non-LCC markets. Based on this chart, one would expect a higher yield for the system-based logistic-fit scenario, with a fare class mix favoring more towards the higher classes.

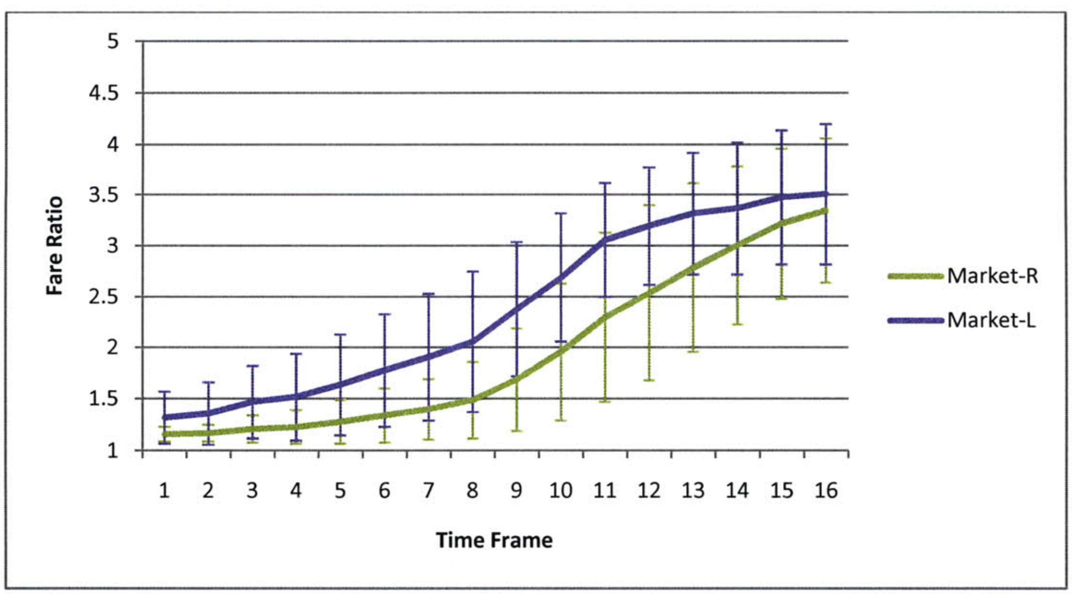


Figure 27: FP Market-based FRAT5 Curves in Network T4

Looking at the market-based curves, it is evident that on average the logistic-fit creates higher, more aggressive FRAT5 curves than the regression-fit, as it did in the system-based case. While only the FRAT5 curves from the FP method are shown, the DO curves behave similarly and generally follow the same pattern. To measure the effect of the different curves, Figure 28 shows the revenue performance for AL1 in Network T4.

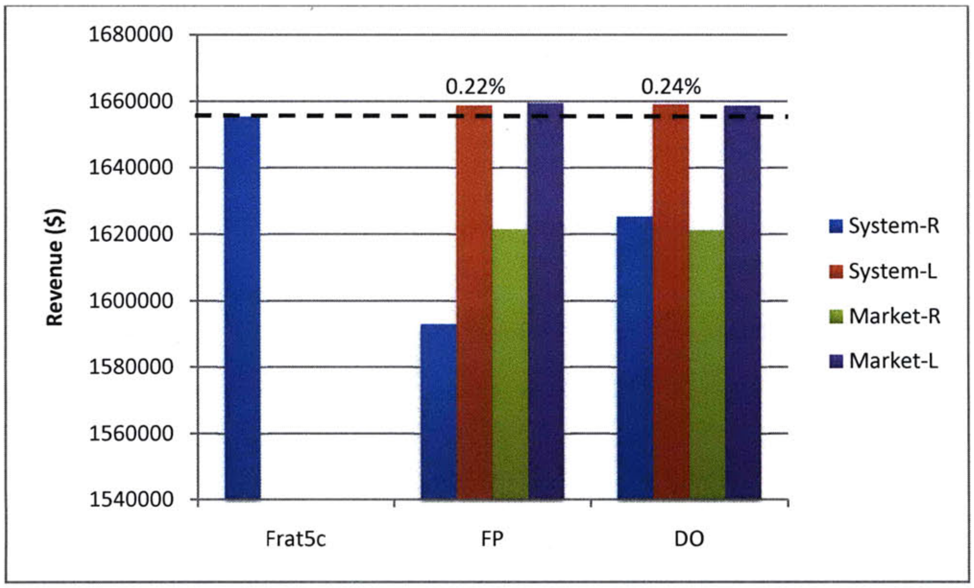


Figure 28: Data-based Sell-up Estimation Revenue in Network T4

Compared to the results in Network D6 Unrestricted, both logistic-fit methods continue to perform well with the market-based regression fit performing poorly. However, unlike in Network D6 Unrestricted, the system-based regression-fit also performs poorly, most likely due to the very low FRAT5 curves generated by the regression-fit. Here, both FP and DO logistic-fit methods create revenue gains over the baseline input FRAT5c curve of up to 0.24%, which is also 5.94% greater than AL1 using just EMSRb without Hybrid Forecasting. To take a closer look, the load factors and yields shown in Figure 29 and Figure 30, respectively, provide more detail.

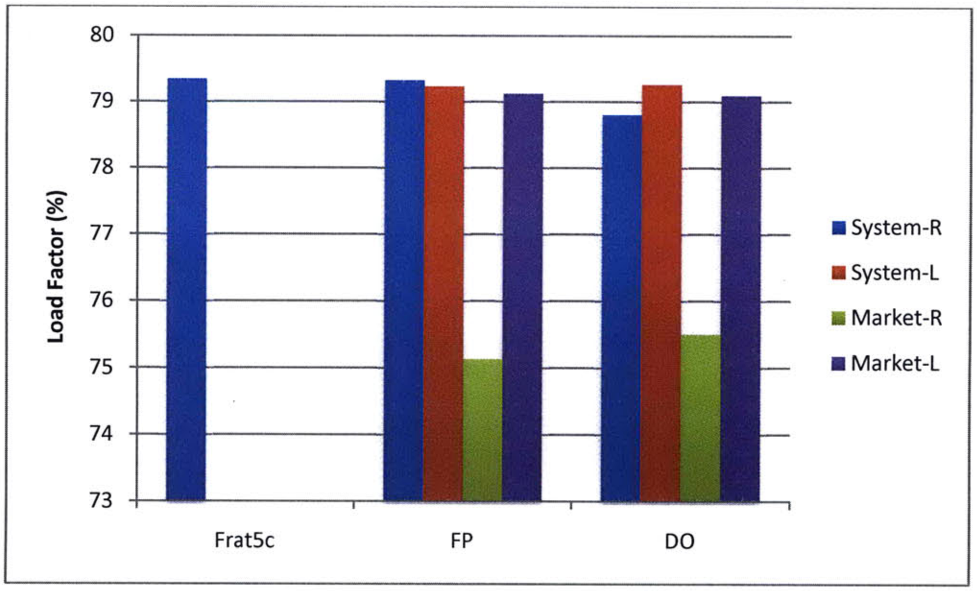


Figure 29: Data-based Sell-up Estimation Load Factors in Network T4

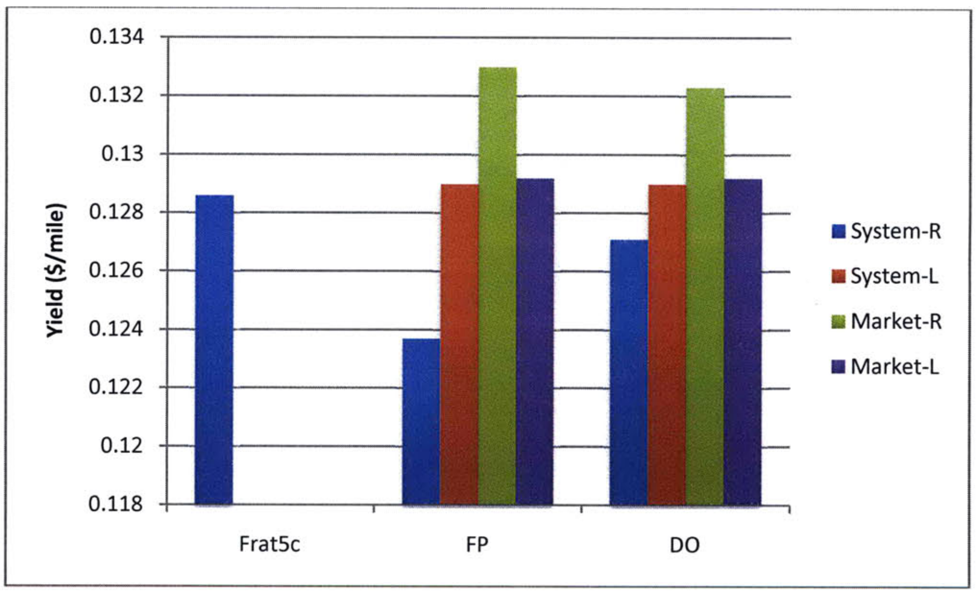


Figure 30: Data-based Sell-up Estimation Yields in Network T4

Compared to the input FRAT5c, the logistic-fit methods create similar load factors but slightly higher yields, in accordance with their higher, more aggressive FRAT5 curves. However, the regression-fit cases are a different case. With the system-based regression-fit curves being much lower than the logistic-fit curves, AL1's load factor was similar to that of the input FRAT5c scenario, but its yield suffered significantly, causing a large drop in revenue. However, the market-based regression-fit case presents an even more interesting scenario, where despite having on lower curves on average, its yield was very high compared to that of the higher, more aggressive logistic-fit curves. But upon further analysis of fare class mix, the increase in yield was not due to more high fare class bookings, but rather to a lack of FC 6 bookings when compared to the other methods. This is also shown by the very low load factors, resulting in an overall decrease in revenue.

Network T1

To further test the performance of system-based and market-based estimation for FP and DO, Network T1 provides another interesting test case. Network T1's more restricted fare structure, even for the LCC markets, will make sell-up observations rarer, but nevertheless it is important to see if there exists any improvement.

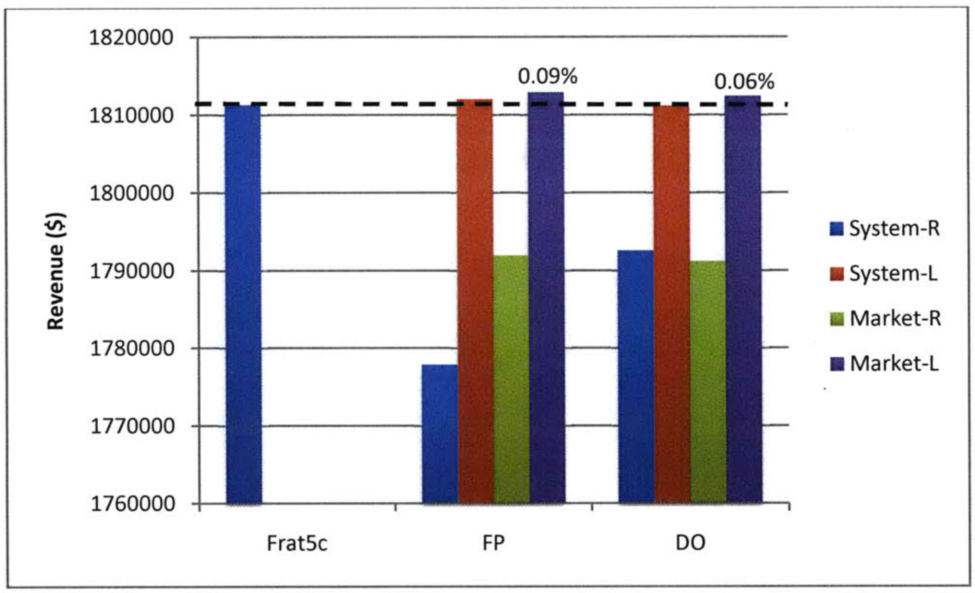


Figure 31: Data-based Sell-up Estimation Revenue in Network T1

Figure 31 shows similar performance trends for data-based sell-up estimation when compared to Network T4, but the increases in revenue over the FRAT5c for the logistic-fit scenarios are much smaller. This is primarily due to the fare structure that better segments demand in LCC markets. Again, regression-fit performance is well below that of the input FRAT5c and the logistic-fit performance. Also, compared to Network T4, the overall revenue is much higher (approximately $1.81 million to $1.66 million) with the more restricted fare structures.

Based on the results from Network D6 Unrestricted, and Networks T4 and T1, it is evident that FP and DO logistic-fit sell-up estimation methods improve revenue compared to that of an input FRAT5c. In addition, there is no clear advantage between system-based and market-based estimation, or between FP and DO. These results and information will serve as a good level of comparison for alternate market sell-up aggregation levels developed through clustering, presented in Section 6.4.2.

6.4 Alternative Methods for Estimating Sell-up

Current methods for estimating sell-up include input FRAT5s and the FP and DO data-based sell-up estimation methods aggregated on the system and market levels. This section attempts to both develop and test new methodology, designed to examine both the fundamentals of input FRAT5s, as well as the aggregation levels of the data-based sell-up methods. These new methods are first tested in Network D6 Unrestricted, the simplest network where the most sell-up may occur, making the use of a FRAT5 curve vital to performance. In order to gain a better understanding of how FRAT5 curves impact sell-up and revenue performance, we first turn to an analysis through the use of particular input FRAT5s.

6.4.1. Hybrid Forecasting with Piecewise Sell-up Estimation

The shape of a FRAT5 curve estimates the expected amount of sell-up over the booking period. In early time frames, there are more leisure passenger bookings. This means that the FRAT5 curve should be lower to accommodate a lower willingness-to-pay. However, in later time frames closer to departure, one would expect the value of the FRAT5 curve to be much higher with a greater number of business passenger bookings. Overall, with the progression of leisure bookings to business passenger bookings, one should expect the curve should be monotonically increasing. With this hypothesis, many questions arise. First, what area of the FRAT5 curve is more important to estimate correctly? Is it worse to overestimate willingness-to-pay in earlier time frames or to underestimate willingness-to-pay in later time frames? If the estimate is too high early on, one should expect fewer low class bookings, leaving too many empty seats at departure. However, if the estimate is too low, very low yields and high load factors will create a drastic decrease in revenue. It is clearly important to determine the best level of the FRAT5 curve for a given time frame within the booking period.

In order to analyze and determine the proper value of a FRAT5 for a given time frame, or set of time frames, testing flat FRAT5s may provide a good starting point. This may also offer insights into the FRAT5's effects on fare class bookings throughout the booking period. While a later time frame's booking limits are not independent of previous bookings, and are greatly dependent on the FRAT5 value, this method is still worthy of consideration. For this experiment, the following FRAT5s will be tested in Network D6 Unrestricted, as shown in Figure 32.

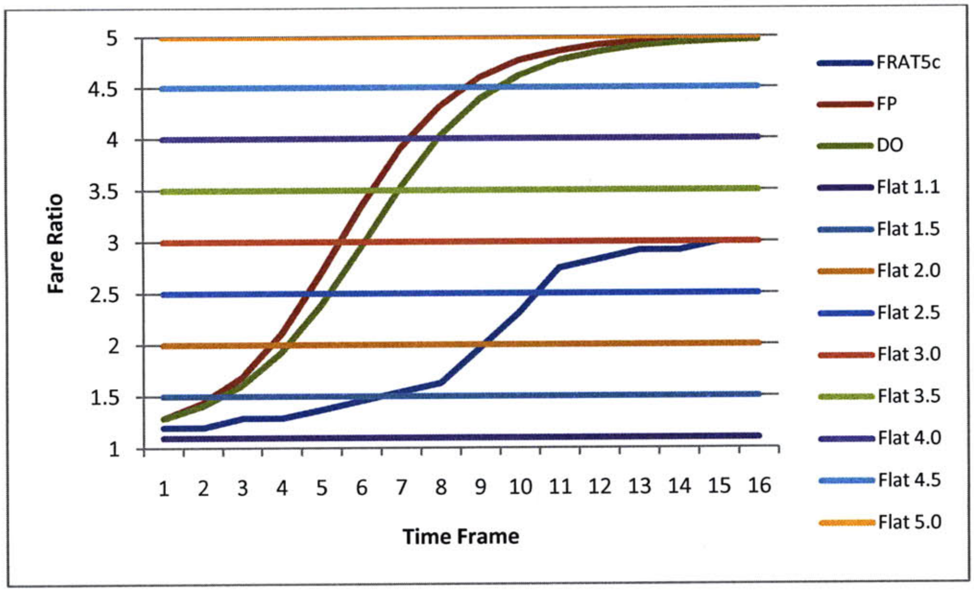


Figure 32: Flat FRAT5 Experiment in Network D6 Unrestricted

In addition to a flat FRAT5 between 1.1 (the minimum FRAT5 value) and 5.0 (the maximum value), the input FRAT5c and FP and DO system-based logistic-fit curves are tested for comparison.

Network D6

The flat FRAT5s produced very surprising results, with the higher FRAT5 curves almost matching the revenue of the data-based estimation methods, as shown in Figure 33.

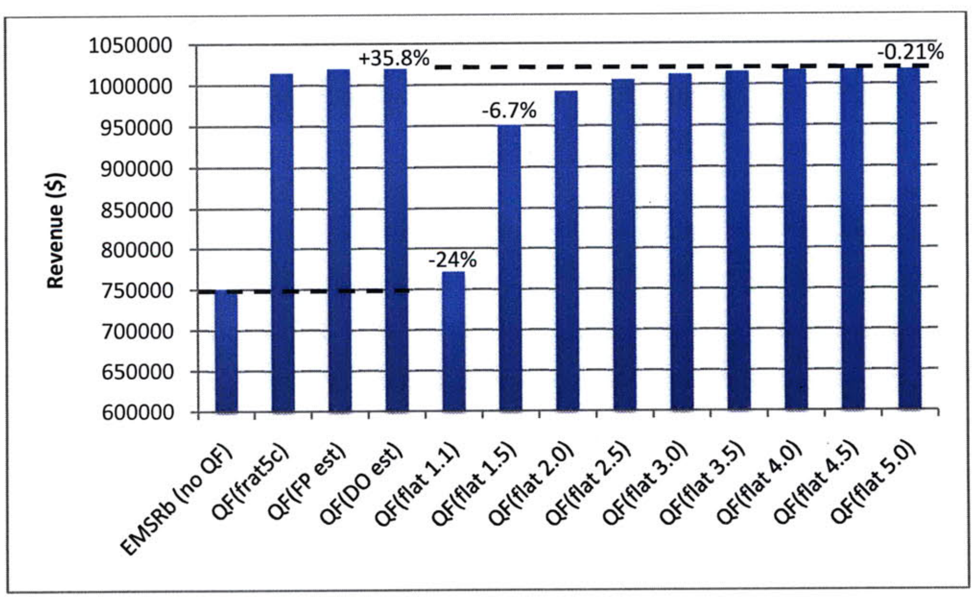


Figure 33: Flat FRAT5 Revenue in Network D6 Unrestricted

It is evident that a higher flat FRAT5 curve performs better than lower curves, meaning that it is better to be too high for early time frames than too low for later time frames, when the most sell-up will occur into the highest fare classes. For example, the flat FRAT5 set to 1.1 barely performed better than the baseline case of not using Q-forecasting at all, and was about 24% below that of the best data-based sell-up methods.

Looking at Figure 34, it is easier to comprehend how the FRAT5 curves impact bookings per class per time frame over the entire booking period.

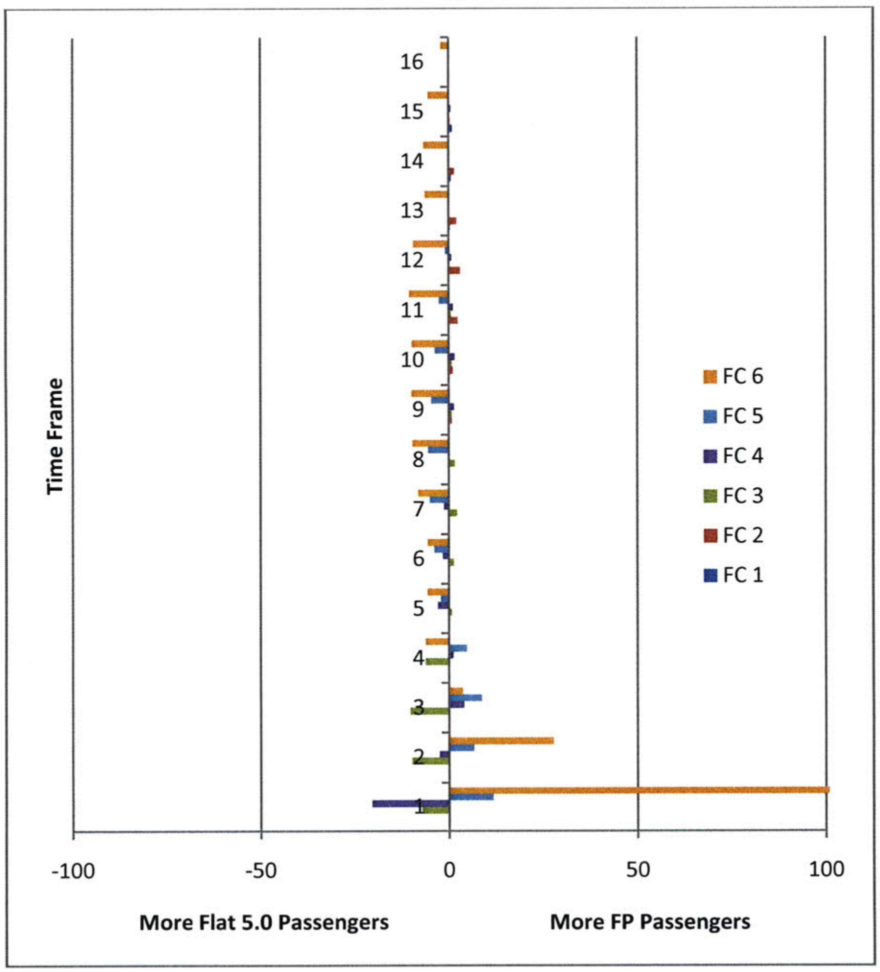


Figure 34: Passengers per Class per Time Frame: FP minus Flat 5.0 in Network D6 Unrestricted

The FP FRAT5 curve's low levels in early time frames result in many more bookings in lower classes, especially FC 5 and FC 6, when compared to the Flat 5.0 FRAT5. Because the Flat 5.0 FRAT5 has more open space due to its very high estimates of sell-up in all 16 time frames, some FC 6 bookings are seen throughout the rest of the booking period. Overall, there is not a large difference in bookings once the FP curve reaches the level of the Flat 5.0 curve. The only main difference between Flat 5.0 and FP estimation is the early lower class bookings, which in the end result in the small 0.21 percent advantage in revenue.

Another interesting comparison is between the Flat 5.0 and Flat 3.0 curves, as shown in Figure 35.

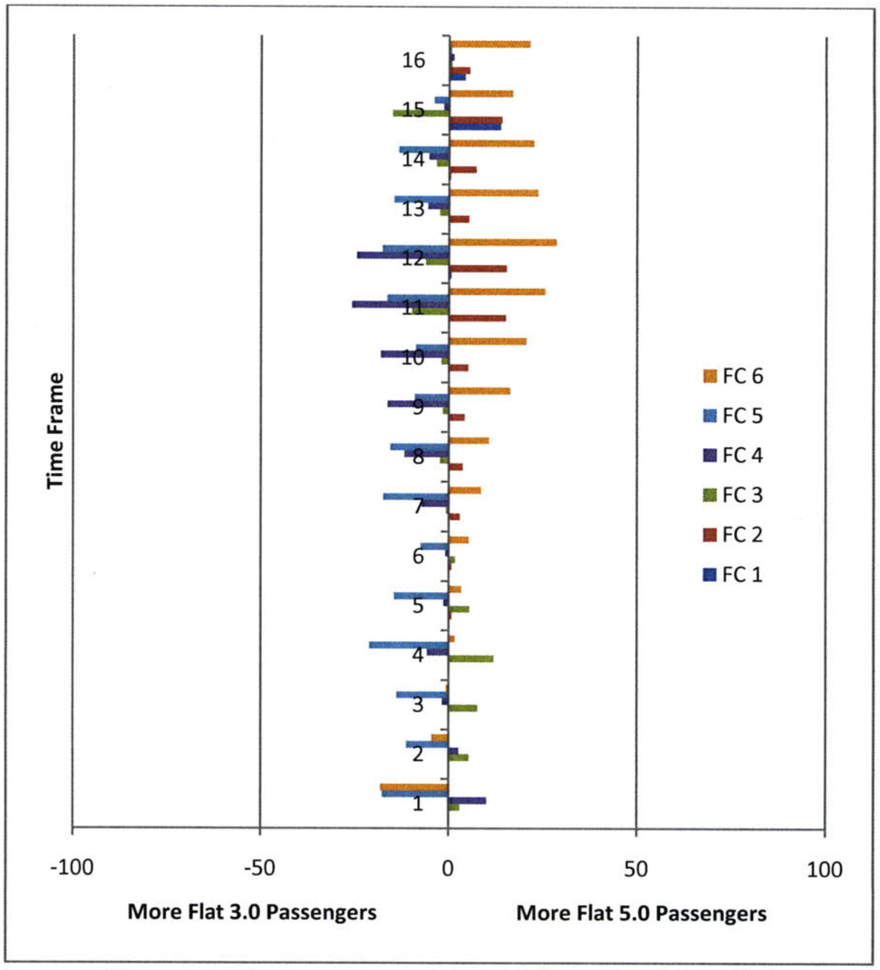


Figure 35: Passengers per Class per Time Frame: Flat 5.0 minus Flat 3.0 in Network D6 Unrestricted

The lower Flat 3.0 curve consistently creates more bookings in the middle classes, namely those of FC 3 through FC 5. However, it is interesting to note that the Flat 3.0 curve only creates more FC 6 bookings in the first two time frames (as expected with the lower curve), but has much fewer FC 6 bookings in the rest of the time frames, especially nearing departure. This is due primarily to Flat 5.0 being so aggressive early on that FC 6 is opened in later time frames. In addition, Flat 5.0 has more passengers in higher fare classes in later time frames, as expected with the higher FRAT5 curve.

To make sense of these various flat FRAT5s, it is important to realize that some curves may have an advantage over others for a particular time frame. For example, looking at how FP performed better than Flat 5.0 in earlier time frames suggests that a lower flat FRAT5 should be used earlier on. Repeating this process to find top performing flat FRAT5s in certain time frames, while under the realization that performance in a particular period is dependent on what happened in previous time frames, one can formulate a piecewise flat FRAT5 curve. This would essentially be a step function, made of certain levels of flat FRAT5s and divided at particular time frames. For this experiment, there are three pieces, or three steps, made from two divisions.

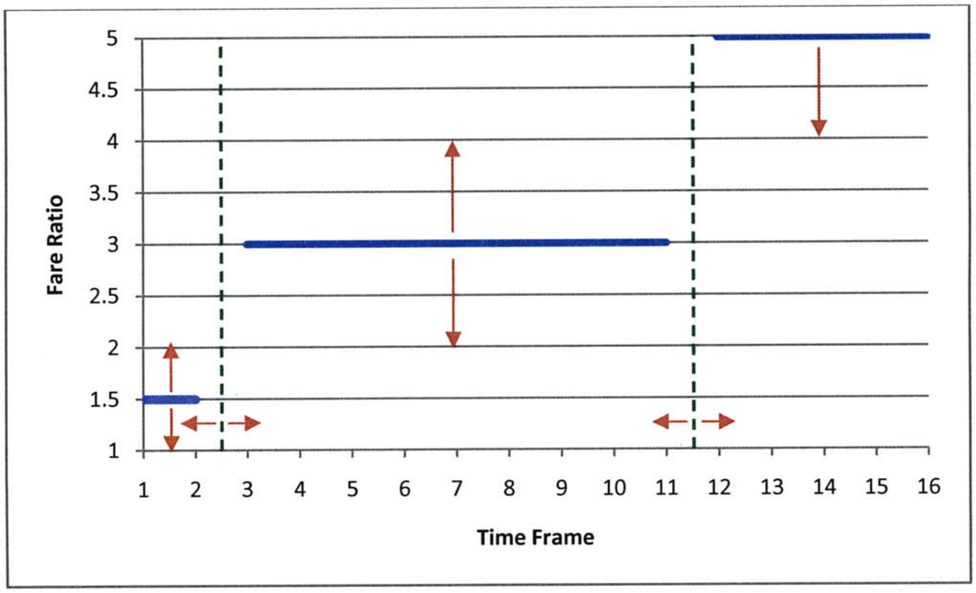


Figure 36: Piecewise Flat FRAT5 Development

The initial values of each FRAT5 level are set at 1.5, 3.0, and 5.0, with the time frame breaks between TF 2 and 3, and between TF 11 and 12. First, the time frame breaks are held constant, and the levels vary until a maximum revenue case is achieved. Then, using the three best FRAT5 levels, the time frame breaks are varied until the overall "best" piecewise FRAT5 curve is found. Using this method, the following piecewise FRAT5 curve was developed in Figure 37, with its comparison to the FP data-based estimation curve.

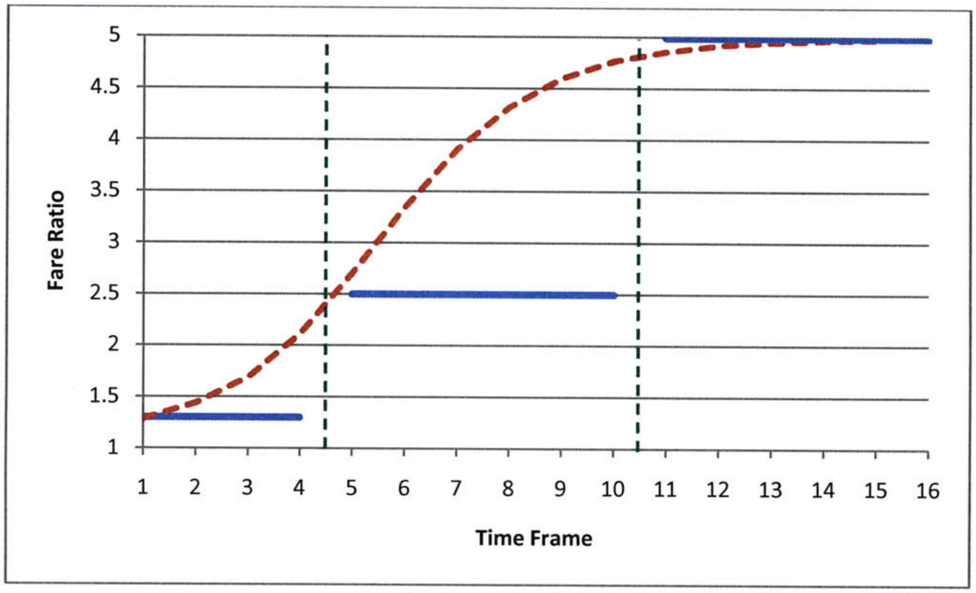


Figure 37: Best Piecewise FRAT5 versus FP FRAT5 Curve in Network D6 Unrestricted

The best piecewise FRAT5 has levels of 1.3, 2.5, and 5.0, with respective time frame periods of TF 1-4, TF 5-10, and TF 11-16. It is interesting to note how this relates to the PODS booking curve, specifically in regards to the percent of business versus leisure passengers in each time frame.

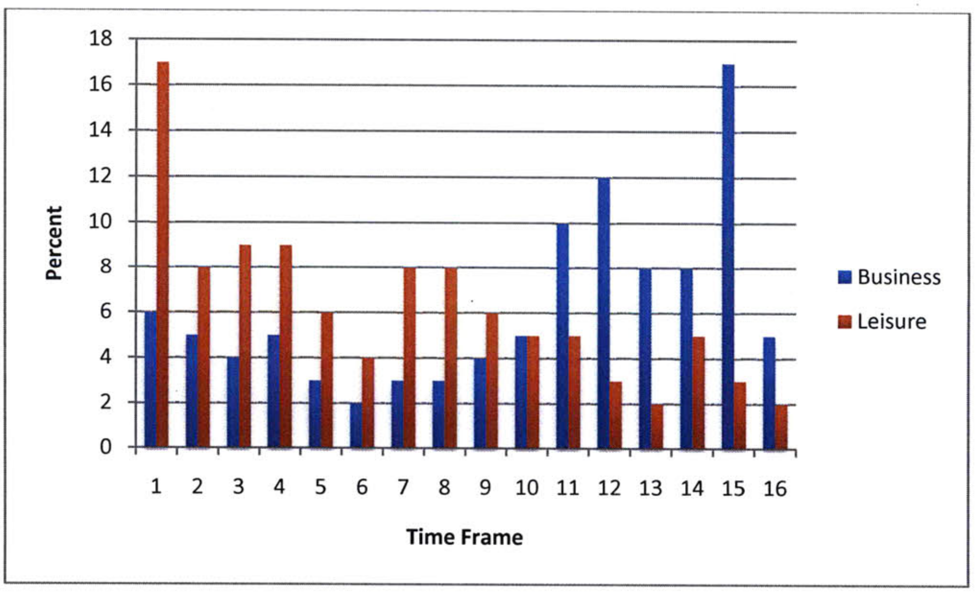


Figure 38: Percent Bookings per Time Frame in PODS

Using what was just determined as the "best" time frame breaks in piecewise FRAT5, one can see how they line up with the type of passenger booking in the particular time period. In the first period of TF 1-4, there are primarily leisure passenger bookings. In the middle period of TF 5-10, there is a mix of passengers, and the last period of TF 11-16 has primarily business passengers. With leisure passengers as the least likely to sell-up because of a lower maximum willingness-to-pay, a FRAT5 level of 1.3 makes sense. In later time frames, where mostly business passengers are booking, a higher FRAT5 level of 5.0 is used.

This informal method of choosing the best piecewise FRAT5 curve actually performed quite well, translating into the best revenue producing FRAT5 curve, creating almost a one-half percent revenue gain over FP system-based sell-up estimation.

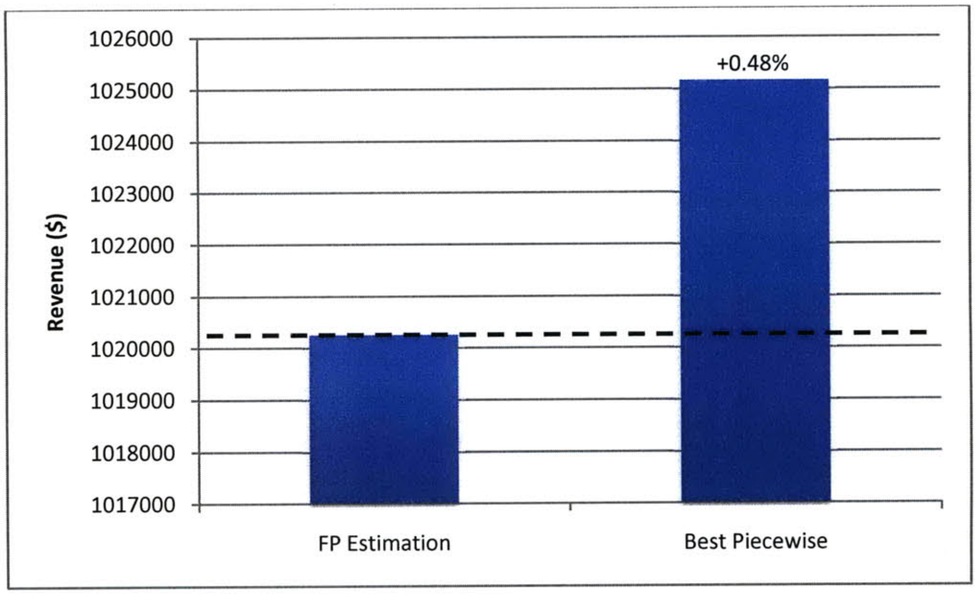


Figure 39: Best Piecewise versus FP FRAT5 Revenue in Network D6 Unrestricted

When comparing the FP estimated FRAT5 with the best piecewise FRAT5, the largest difference lies between TF 5 and TF 10. In that period, the FP curve rises from 2.5, nearly reaching 5.0, while the best piecewise curve maintains a level value of 2.5. To better understand the effects of this and the origin of the extra half percent of revenue, Figure 40 provides a closer look.

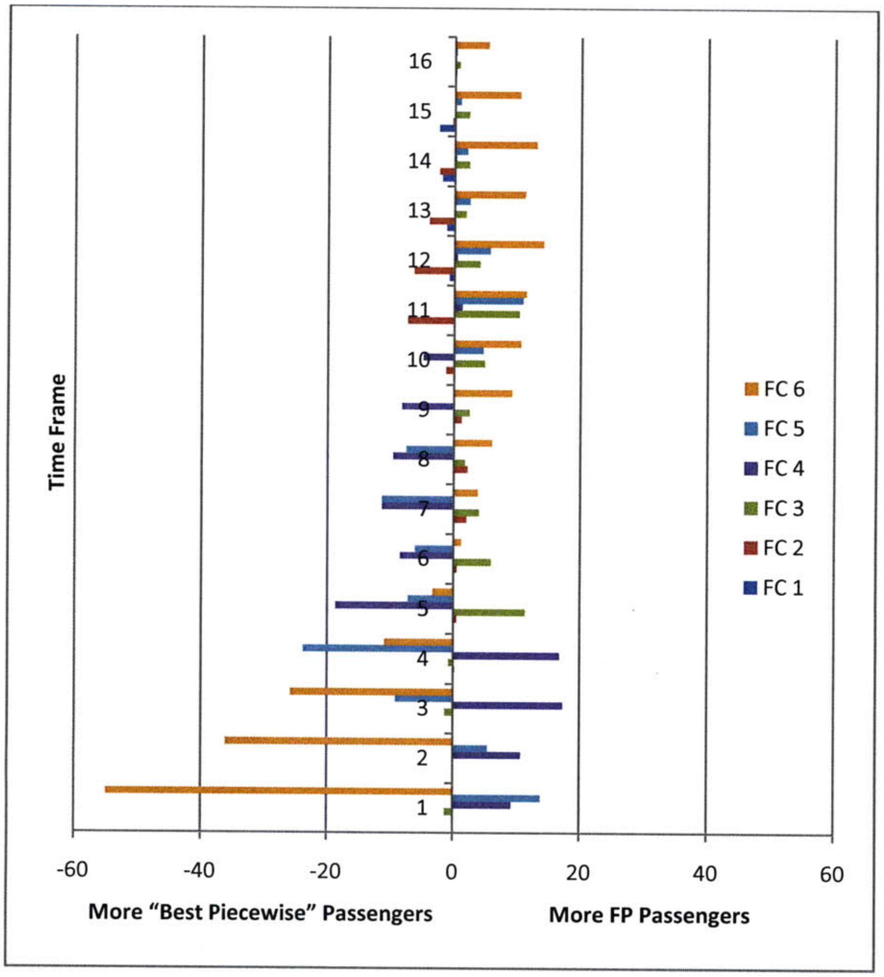


Figure 40: Passengers per Class per Time Frame: FP Estimation minus Best Piecewise in Network D6 Unrestricted

Compared to the FP estimated FRAT5, the best piecewise FRAT5 curve created more bookings in FC 5 and FC 6 in the first period from TF 1 to TF 4. This is due to the fact that the best piecewise curve maintained a level value at 1.3 while the FP estimated curve began to climb. In the middle time frames between TF 5 and TF 10, there is a shift to the piecewise curve creating more bookings in FC 4 and FC 5. In the final time frames, it is evident that the piecewise curve creates slightly more FC 1 and FC 2 bookings compared to the FP estimated curve. Because of FP's over-aggressiveness in the middle period between TF 5 and TF 10, FP does create more FC 6 bookings in later time frames, similar to what occurred with the Flat 5.0 FRAT5 curve.

There appears to be distinct advantages for different sets of fare classes in accordance with (1) where the time frame breaks are and (2) what the FRAT5 levels are for the given period. For example, between TF 1 and TF 4, there are primarily more FC 1 and FC 2 bookings for the piecewise FRAT5 because its FRAT5 value for that period is 1.3. This corresponds to most of the bookings in that period having a fare ratio less than that FRAT5 value, for the average fare ratios are 1.26 for FC 5 and 1.00 for FC 6.

With the piecewise FRAT5 performing very well in Network D6 Unrestricted, it is important to determine how well it functions in other environments. In Network D6 Semi-restricted, the same process for determining the "best" piecewise input FRAT5 curve was repeated, achieving similar results.

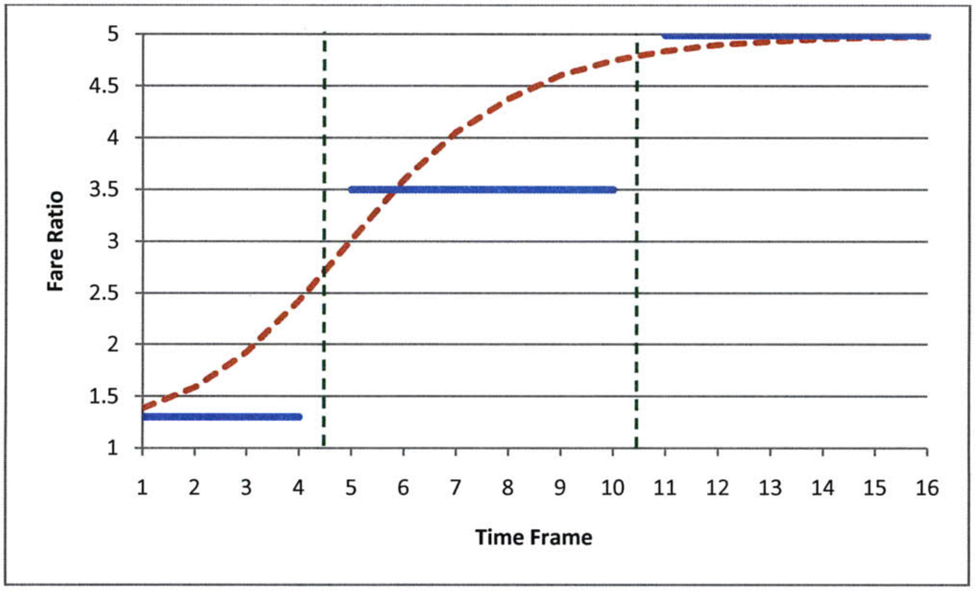


Figure 41: Best Piecewise FRAT5 versus FP FRAT5 Curve in Network D6 Semi-restricted

The best piecewise FRAT5 for Network D6 Semi-restricted has the same time frame breaks as the Network D6 Unrestricted, as well as the same values for the first and third periods. The only difference is that the middle level is at 3.5 instead of 2.5. Using the piecewise input FRAT5 proved beneficial for AL1 as it did in the unrestricted case, producing a 0.17% increase in revenue over the best data-based sell-up estimation method (FP). The revenue increase was not quite as great simply due to the increased segmentation of the semi-restricted fare structure and there being a smaller opportunity for sell-up to occur.

Network T4

To apply the piecewise FRAT5s created in both Network D6 Unrestricted and Network D6 Semi-restricted to Network T4, there exist several options. For example, using the fare structure parallels between Network D6 and Network T4, one option is to use the piecewise FRAT5 created in Network D6 Unrestricted and apply it to the LCC markets in Network T4, while using the piecewise FRAT5 created in Network D6 Semi-restricted and apply it to the non-LCC markets in Network T4. In addition, one may apply the same piecewise curve to both categories of markets in Network T4, and another option is to apply the average of the two piecewise curves to each kind of market. These test options are listed in the table below.

	LCC Markets			Non-LCC Markets		
	low	*medium*	*high*	*low*	*medium*	*high*
Piecewise #1	1.3	2.5	5.0	1.3	3.5	5.0
Piecewise #2	1.3	3.5	5.0	1.3	3.5	5.0
Piecewise #3	1.3	2.5	5.0	1.3	2.5	5.0
Piecewise #4	1.3	3.0	5.0	1.3	3.0	5.0

Table 19: Network T4 Piecewise FRAT5 Curves

Applying these input piecewise FRAT5 curves in Network T4 created outstanding results, better than the input FRAT5c and both FP and DO data-based estimation methods (system-based logistic-fit).

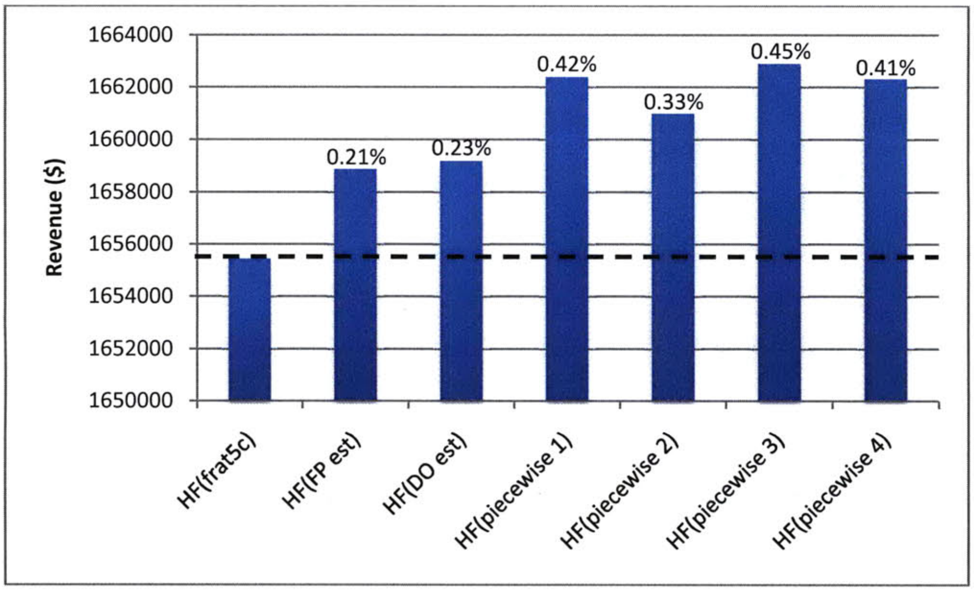


Figure 42: Piecewise FRAT5 Revenue: Network T4

Looking at Figure 42, it is evident that applying the piecewise FRAT5 developed from Network D6 Unrestricted to both the non-LCC and LCC markets in Network T4 created the highest revenue, 0.45% over that of the input FRAT5c and 0.22% greater than DO data-based estimation. This method is followed closely by assigning the piecewise FRAT5 curve from D6 Unrestricted to the LCC markets and the piecewise FRAT5 curve from D6 Semi-restricted to the non-LCC markets.

Despite being an arbitrary input FRAT5 not statistically estimated from booking data, piecewise methodology works well and outperforms the best data-based method in all three networks tested. Later in this thesis, a new methodology to transform the input piecewise step function into an actual data-based method for creating a piecewise FRAT5 will be presented. In short, this method will take the aggregation level advantage from clustering (presented in the following section) and combine it with the performance of the piecewise FRAT5 in hopes of creating a "clustered piecewise FRAT5."

6.4.2. Hybrid Forecasting with Clustered Sell-up Estimation

6.4.2.1. Clustered With Logistic-fit

Using the K-means clustering algorithm presented in Section 5.2, it is possible to cluster markets based on their FRAT5 curve, which defines each market's estimate of sell-up. Recall that the logistic fitter created the best-performing FRAT5s on both the system-wide and the per-market aggregation levels. The logistic-fit parameters, x1 and x2 for the two parameter version, and x1, x2, and x3 for the three parameter version, define the shape of the FRAT5 curve and serve as the basis for clustering. The goal of this process is to use these parameters, as well as a desired number of clusters, in order to place the markets into similar groups, estimate the FRAT5 for the given cluster, and assign all markets within a cluster the same estimated FRAT5 curve.

In PODS, the process begins with the simulation run of AL1 using an input FRAT5. From each trial, the logistic parameters for FP and DO from each of the 400 samples are estimated and recorded, but are not used. At the end of the simulation, there exist two or three logistic parameters averaged over the 400 samples for each market and over the number of trials in the simulation, as well as the average number of recorded "observations" used to determine the parameters. In this sense, the term "observations" means the number of "good" samples out of 400, multiplied by the number of trials run in the simulation (400 samples per trial, five trials in Network D6, two trials in Network T). Recall that there are rules for a sample to be considered "good:" for a single time frame there must be at least two occurrences of sell-up across the previous 26 historical observations in order to find a FRAT5 value for the time frame (estimated via FP or DO methods). If there are at least four time frames that have a FRAT5 for the sample, then the logistic fitter may be applied, and kept if the logistic parameter values are not at the upper or lower bounds. If the market does not meet these requirements, then it does not receive any logistic-fit parameters, causing it to be left out of the clustering process. If a market is not clustered, it instead receives an input FRAT5c for use in Hybrid Forecasting. If the market is clustered, it will receive the logistic-fit parameters from the cluster center (cluster mean). Following this process, the PODS simulation is re-run with the markets using their new clustered estimates of sell-up or with their continued use of the input FRAT5c.

Network D6 Unrestricted

Network D6 Unrestricted serves as the first network for testing the clustering process. In this environment, AL1 uses EMSRb with Q-forecasting, where sell-up is estimated via the 2-parameter logistic-fit for both FP and DO estimation methods. Airline 1 is in competition with AL2, which uses

AT90 revenue management. First, to better understand the clustering process, scatter plots of the two parameters for FP and DO, as well as the total within sum-of-squares provide a comparison of the spread of the data.

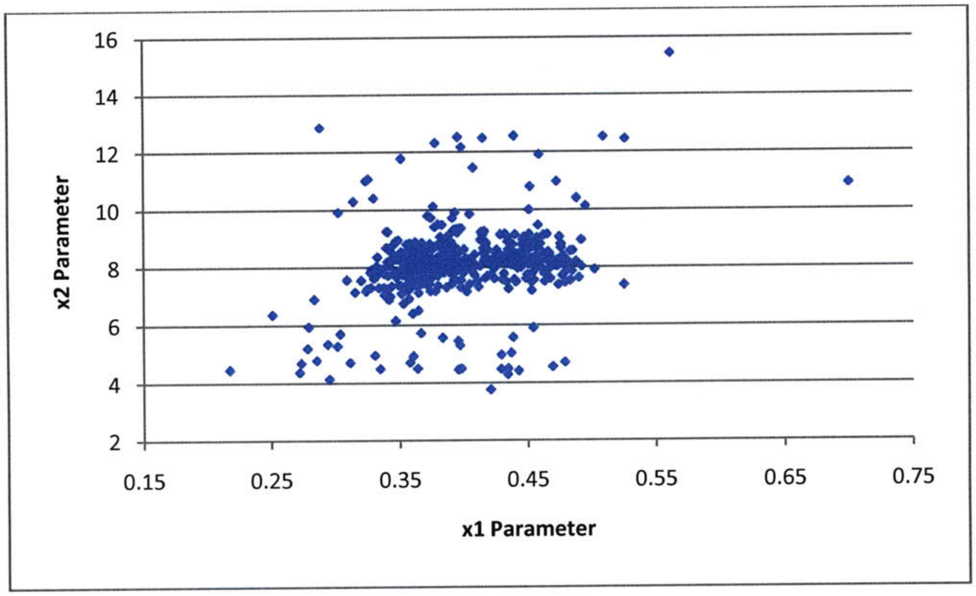


Figure 43: FP 2-Parameter Scatter Plot in Network D6 Unrestricted

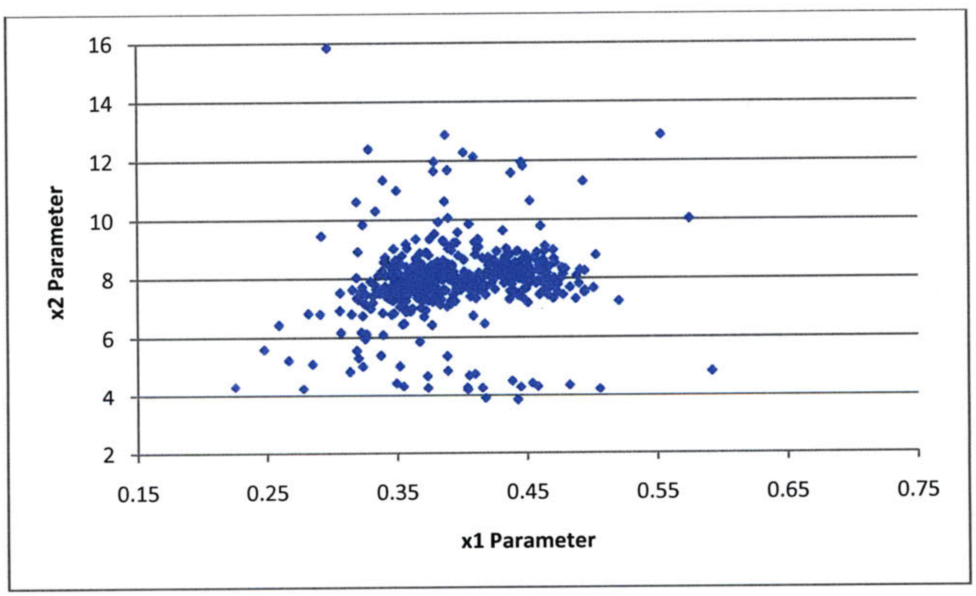


Figure 44: DO 2-Parameter Scatter Plot in Network D6 Unrestricted

Visually inspecting the scatter plots shows that in both cases, the data are mostly contained within a main grouping of points, with outliers surrounding it. Looking closely, it also appears that there are two sub-groupings within the main group of data. To see the impacts of clustering on the total cluster variance, a plot of the within sum-of-squares in Figure 45 provides good insights.

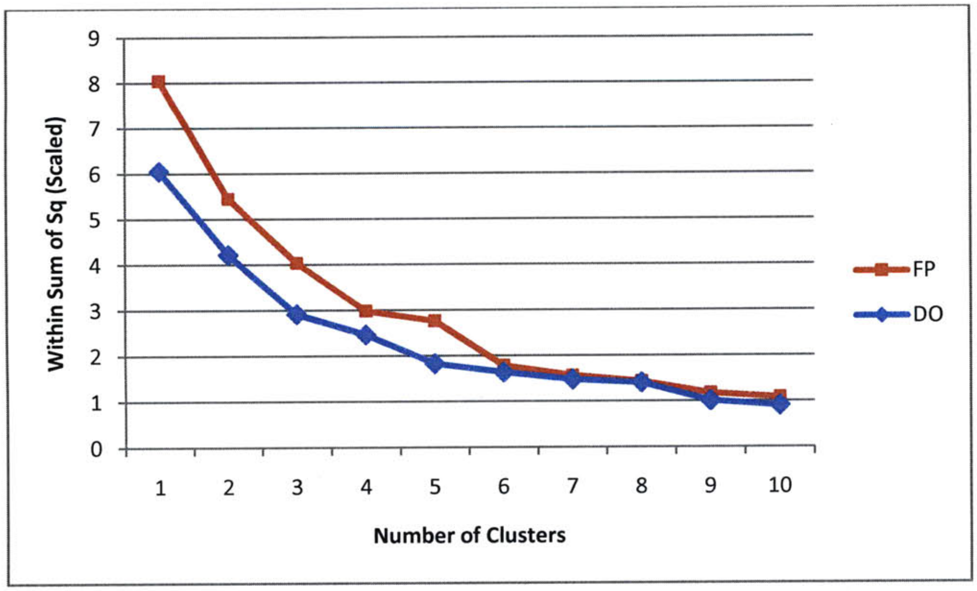


Figure 45: FP and DO Clustering Within Sum-of-Squares in Network D6 Unrestricted

To determine the total within sum-of-squares, the data are first scaled so that they all have a range of 1.0. Doing this creates a normalized representation of the spread, preventing one parameter with a larger range (such as x2 in this case) from dominating the total cluster variance. In the logistic-fit data, the lower DO total within sum-of-squares line suggests that DO creates more distinct clusters than FP, at least for up to and including five clusters. Beyond about seven clusters, little is gained by increasing *k*, the number of clusters. Another representation of this is the market distribution over the given number of clusters.

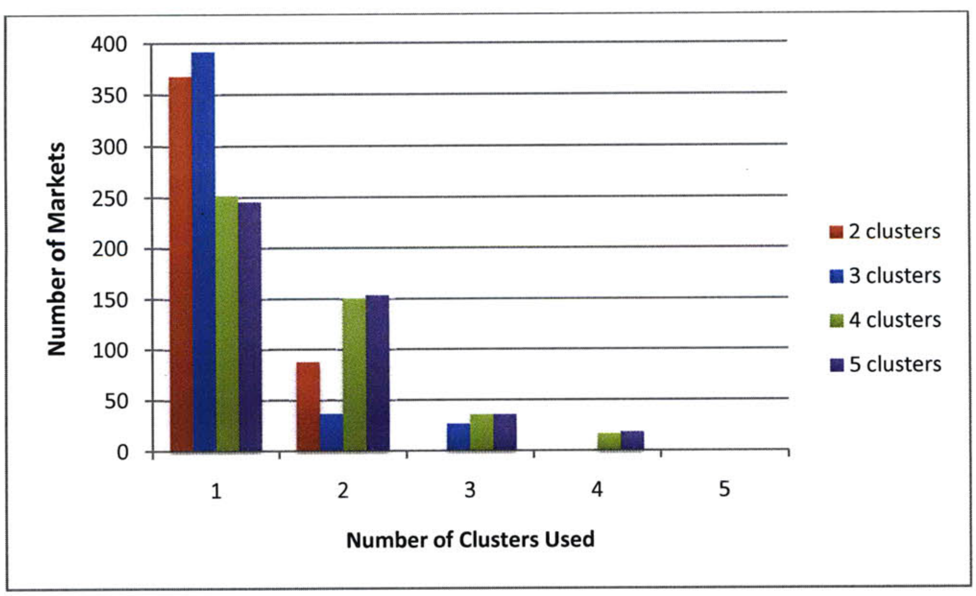


Figure 46: FP Estimation: Effects of Increasing *k* on the Number of Markets per Cluster in Network D6 Unrestricted

Looking at the FP case, it is evident that adding additional clusters beyond a certain point only breaks off a couple market data points and makes them their own cluster, minimally reducing the total cluster variance. In this case, increasing from four to five clusters only moved one market into the fifth cluster, which was previously the largest outlier causing the most damage to the total cluster variance. In order to determine the statistically proper number of clusters, the gap statistic, as presented in Section 5.3, provides another good reference point.

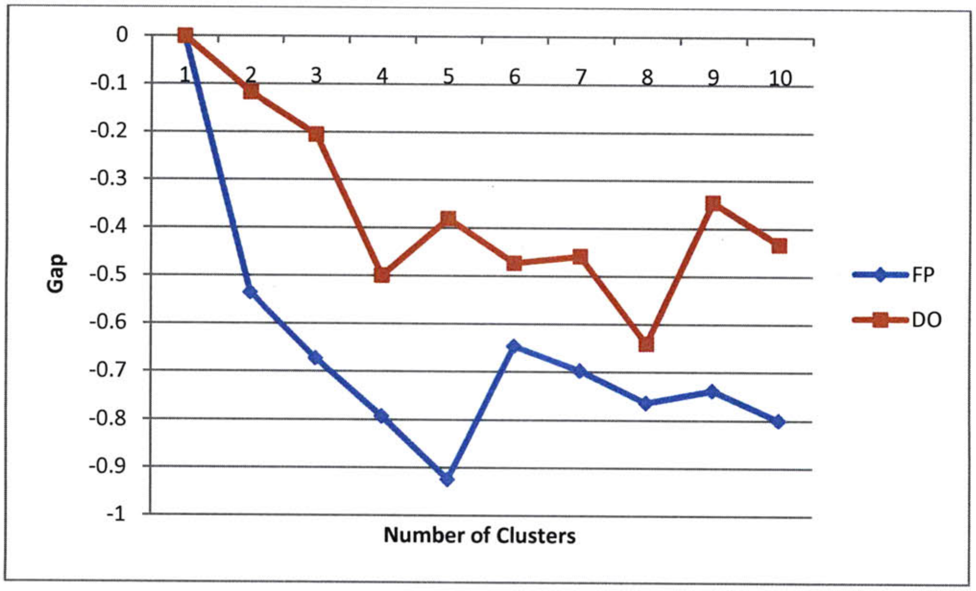


Figure 47: FP and DO 2-Parameter Logistic-fit Gap Statistic in Network D6 Unrestricted

According to the gap statistic methodology, the optimal number of clusters is the point at which the maximum value of the gap statistic curve occurs. In both FP and DO cases, this method suggests that k=1. Looking back at the scatter plots, this makes sense because most of the data appears to be confined in a single grouping. However, another important feature of the gap statistic is that any increase in the gap curve suggests a number of clusters where a given number of "sub clusters" may exist. Figure 47 suggests that a potential number of clusters other than k=1 may be at six for FP and five or nine for DO.

While there is no guarantee that the optimal number of clusters determined by the gap statistic is also the revenue-maximizing number of clusters, this method in addition to examining the market distribution and within sum-of-squares reinforces that a reasonable number of clusters will probably lie between two and seven. However, in Network D6 Unrestricted, more clustering options beyond seven clusters are tested, as shown in Figure 48 for FP and Figure 49 for DO.

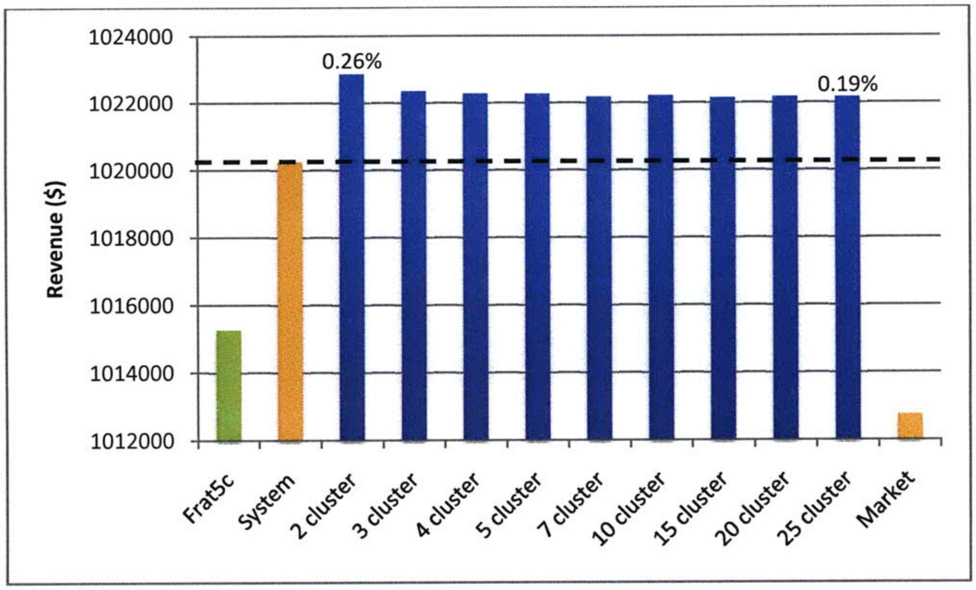


Figure 48: FP 2-Parameter Clustering Revenue per Cluster in Network D6 Unrestricted

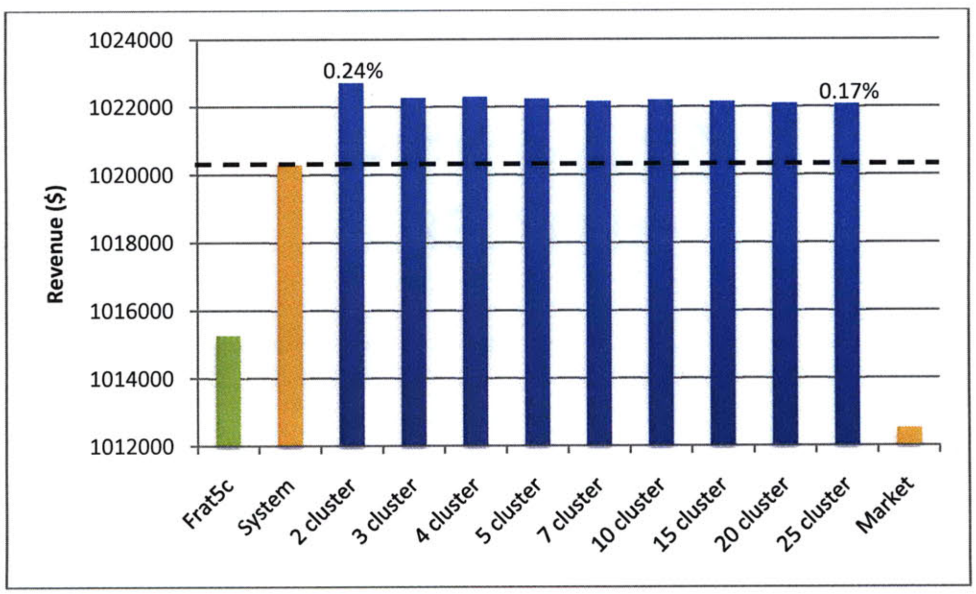


Figure 49: DO 2-Parameter Clustering Revenue per Cluster in Network D6 Unrestricted

The clustering method performed remarkably well for both FP and DO in Network D6 Unrestricted. Clustering outperformed the best previous aggregation level, system-based logistic-fit, by 0.26% for FP and by 0.24% for DO. Overall, FP slightly edges DO in terms of revenue gain, and for both cases, the revenues peak at k=2 clusters, while hovering around 0.20% before falling off at larger numbers of clusters. Compared to the input FRAT5c, the best clustering methods provide revenue increases of 0.75% and 0.73% for FP and DO, respectively. These are extremely promising results, showing that clustering not only successfully provides a middle ground between system and market-level aggregation, but grouping similar markets based on sell-up estimation improves the overall revenue.

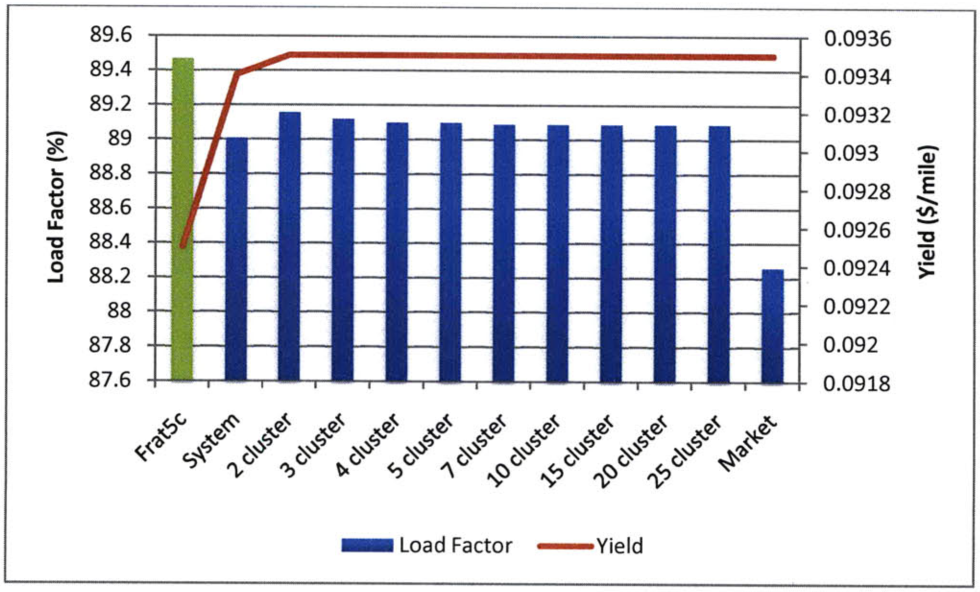


Figure 50: FP 2-Parameter Clustering Load Factors and Yields in Network D6 Unrestricted

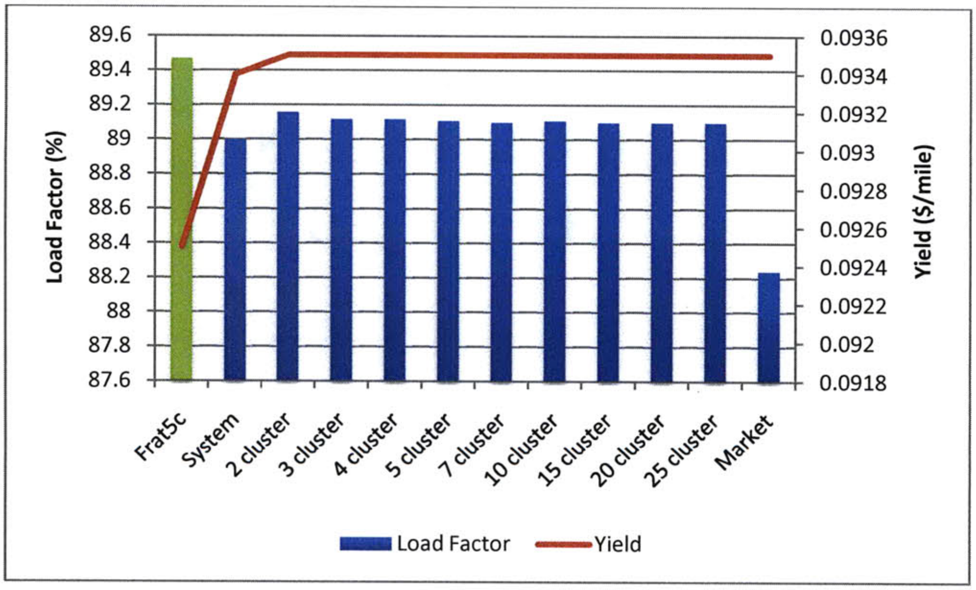


Figure 51: DO 2-Parameter Clustering Load Factors and Yields in Network D6 Unrestricted

Looking at the load factors and yields for the FP and DO clustering methods, most of the benefit gained by clustering results from achieving slightly higher load factors and obtaining the highest yields compared to all other data-based estimation methods. This is likely due to markets receiving a proper FRAT5 curve depending on their cluster membership, preventing the under-protection (high LF with FRAT5c) or over-protection (low LF with System and Market) of seats caused by other methods that were too broad or too specific in estimating sell-up. In the following analysis, focus is given to k=3, 5, and 7 clusters, both great performers in revenue and in the middle of the range of a sensible number of clusters.

These promising revenue results are highly dependent on the FRAT5 curves themselves. The clustering method created very reasonable FRAT5 curves, as shown in Figure 52 below.

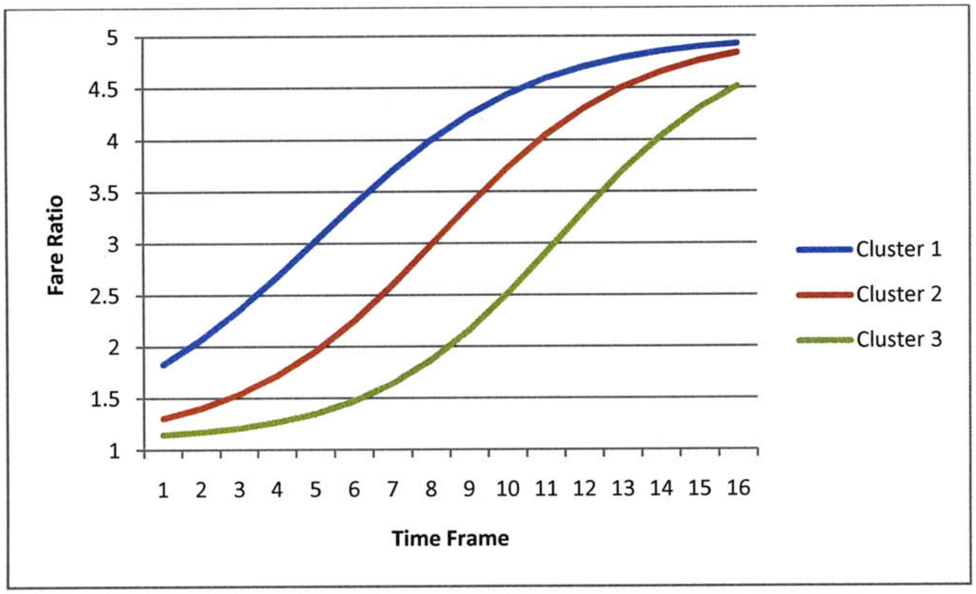


Figure 52: FP 2-Parameter 3 Cluster FRAT5 Curves in Network D6 Unrestricted

Figure 52 shows that the curves are well defined and spaced far enough apart to be distinguishable. This spread over the time frames and fare ratios is sensible because it creates three categories of markets, one that demands a very aggressive FRAT5 curve (Cluster 1 with 37 markets), one that demands a low unaggressive curve (Cluster 3 with 27 markets), and another that has a middle-of-the-road curve (Cluster 2 with 392 markets). Even when the number of clusters increases to k=5 and k=7, the curves continue to be well spaced and defined.

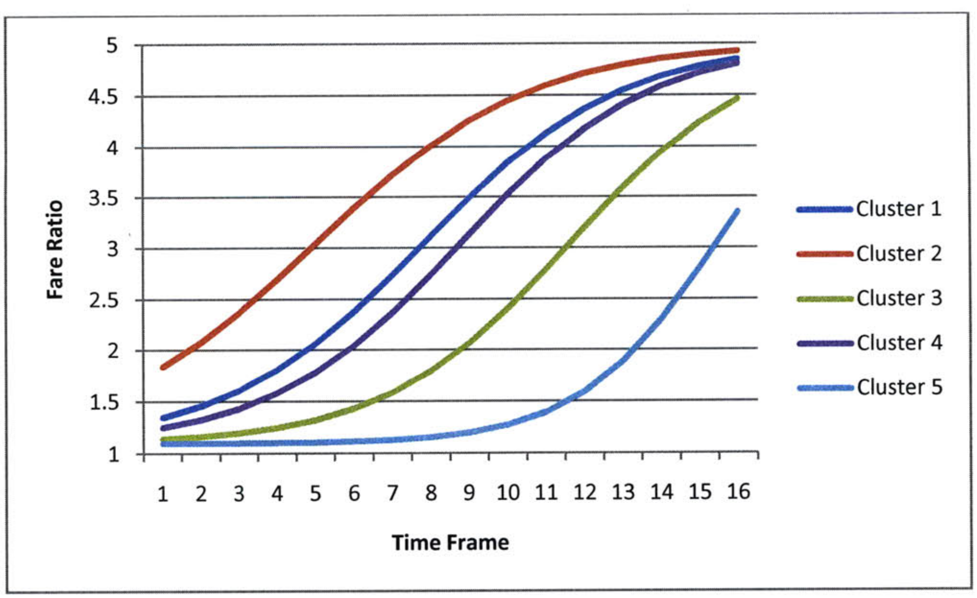


Figure 53: FP 2-Parameter 5 Cluster FRAT5 Curves in Network D6 Unrestricted

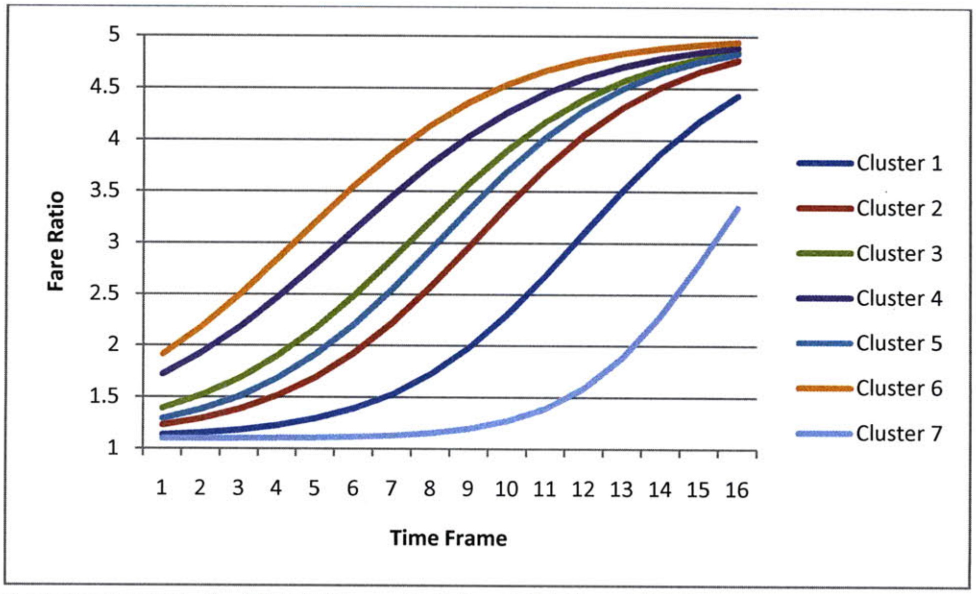


Figure 54: FP 2-Parameter 7 Cluster FRAT5 Curves in Network D6 Unrestricted

The seven cluster case shown above continues to produce well-defined FRAT5 curves, but in relation to the market distribution between clusters, some of the outlying curves receive only a minimal number of markets. For example, in Figure 54, Cluster #7 is by far the least aggressive FRAT5 curve and has only one market. Also, while only the FP curves are shown, it is important to note that the clustering method for DO creates extremely similar FRAT5 curves to FP. While the 2-parameter logistic-fit performed very well in regards to clustering, we also test the 3-parameter version.

To determine if the 3-parameter logistic-fit creates sensible results, the clustering method was applied for *k*=3, 5, and 7 clusters to both the FP and DO cases.

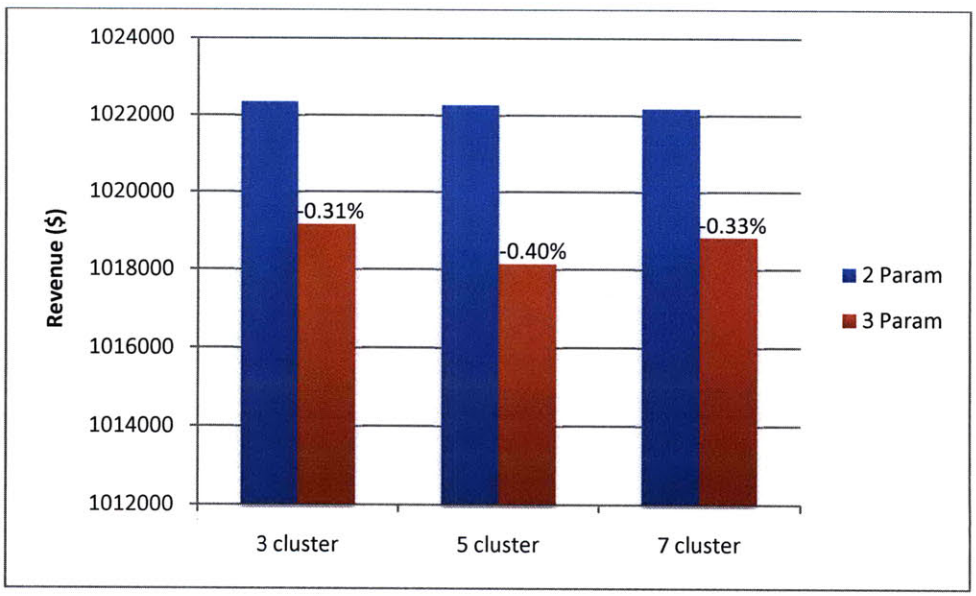


Figure 55: FP 2-Parameter versus 3-Parameter Revenue in Network D6 Unrestricted

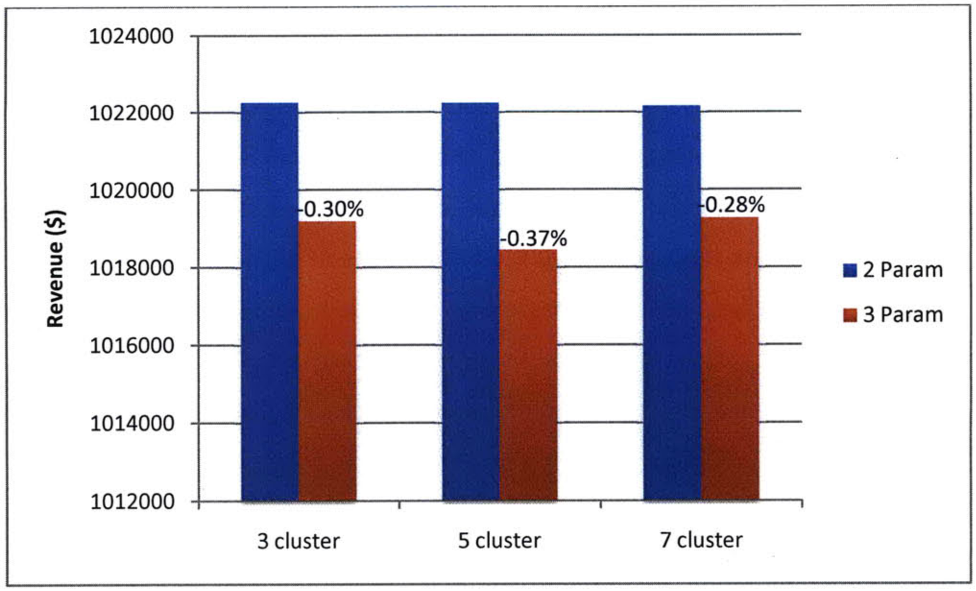


Figure 56: DO 2-Parameter versus 3-Parameter Revenue in Network D6 Unrestricted

The results indicate that the 3-parameter logistic-fit versions of FP and DO, while still better than the system or market-based methods, do not match the revenue levels of the 2-parameter fitter. To better understand why this occurs, it is evident that the 3-parameter method creates FRAT5 curves that are much less aggressive than the 2-parameter method, as shown in Figure 57.

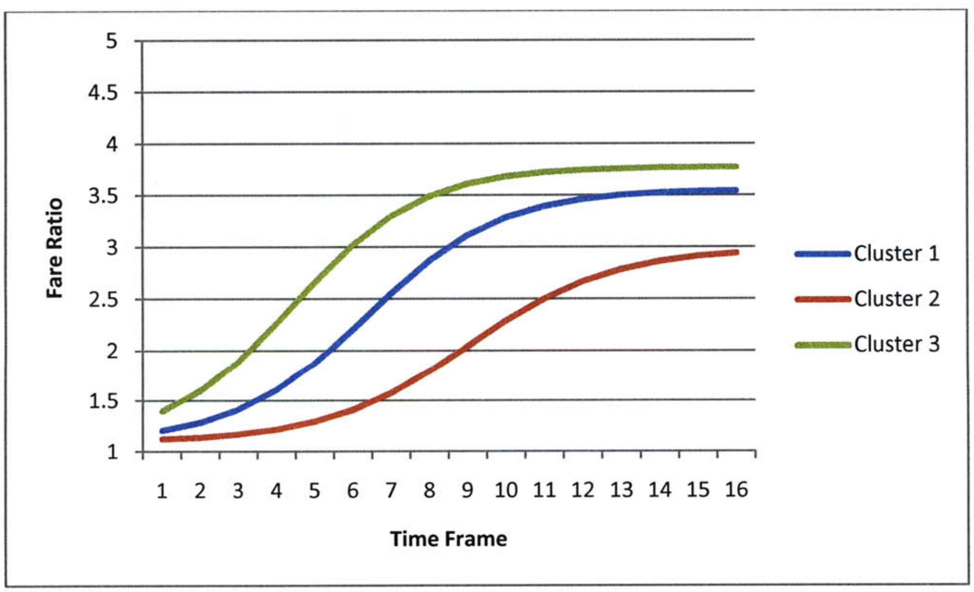


Figure 57: FP 3-Parameter 3 Cluster FRAT5 Curves in Network D6 Unrestricted

The DO 3-parameter method also created very similar curves to those of the FP method shown above. The 2-parameter clustered curves are much more aggressive than the 3-parameter curves, as shown by the 3-parameter curves barely reaching a FRAT5 level over 4.0.

Network T4

Network T4 provides a good test setting to determine how well clustering does in a more complex environment. In this network, clustering is performed irrespective of whether a market is classified as an LCC or non-LCC. Also, using what was learned in the Network D6 Unrestricted simulations, only 2-parameter logistic-fit methods are tested in Network T4 at *k* values of 3, 5 and 7 clusters. Because Network T4 has both a fully unrestricted fare structure as well as a more restricted fare structure for non-LCC markets, one would expect clustering to still perform well, but not to the extent of the performance in Network D6 Unrestricted, simply due to the smaller opportunity for sell-up to occur in the non-LCC markets.

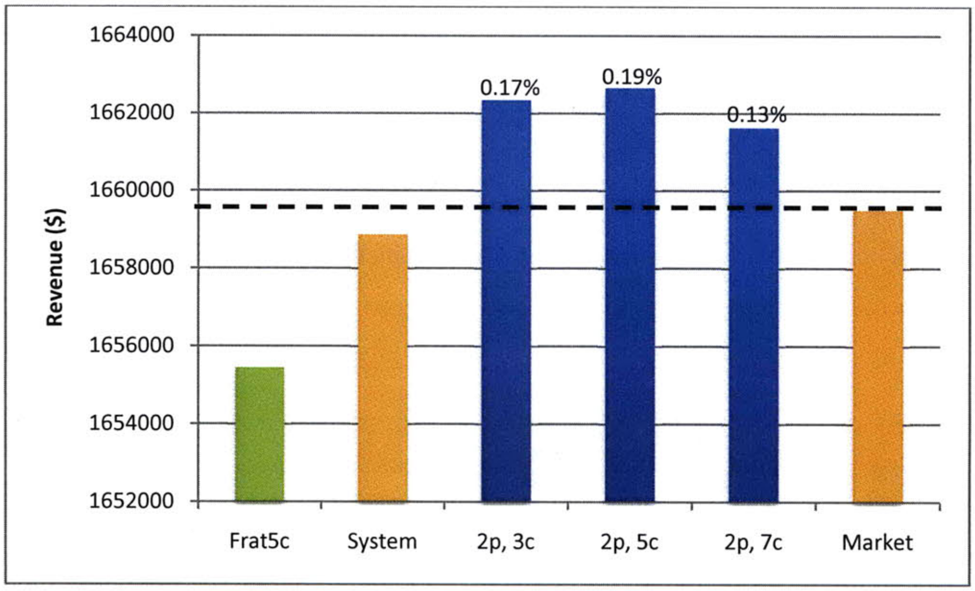


Figure 58: FP 2-Parameter Cluster Revenue in Network T4

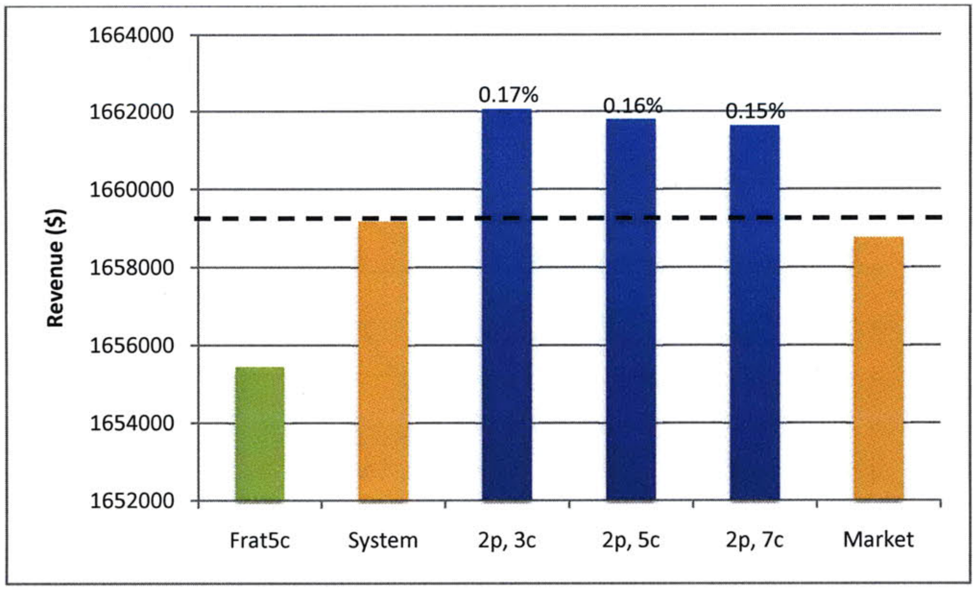


Figure 59: DO 2-Parameter Cluster Revenue in Network T4

Similar to Network D6 Unrestricted, the clustering method worked very well in Network T4, boasting a revenue increase of 0.19% over the best previous FP logistic-fit aggregation level, and 0.17% better than the best previous DO logistic-fit aggregation level. Compared to the input FRAT5c, these increases amount to 0.44% and 0.40%, for FP and DO, respectively.

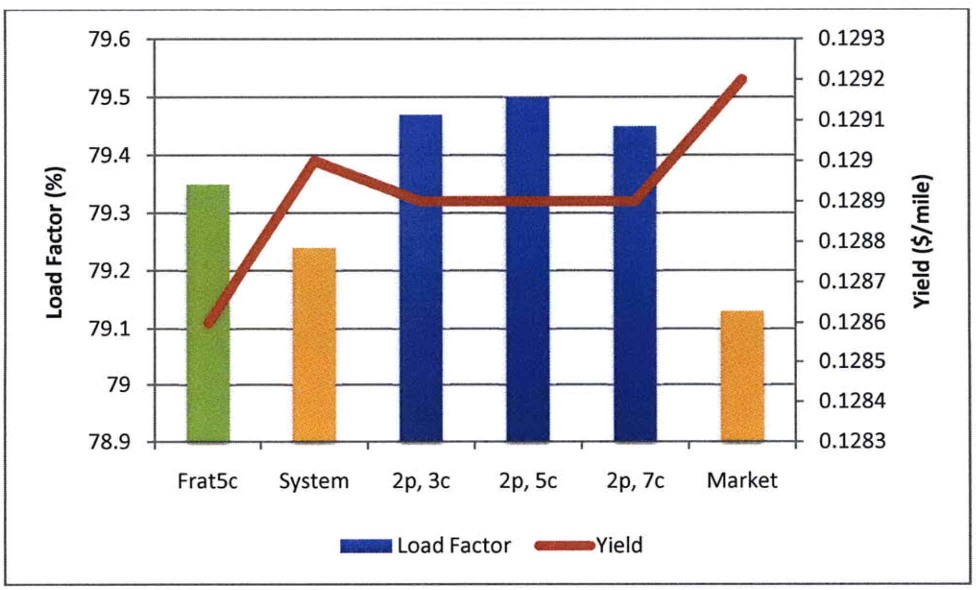


Figure 60: FP 2-Parameter Clustering Load Factors and Yields in Network T4

Additionally, the load factors and yields indicate that there are more passengers traveling under the clustering methods, while keeping yield at a moderate level. The results for the DO load factors and yields are quite comparable to those of FP estimation. Much like the FRAT5 curves for the clusters in

Network D6 Unrestricted, the curves for Network T4 are well-spaced, with analogous results for FP and DO.

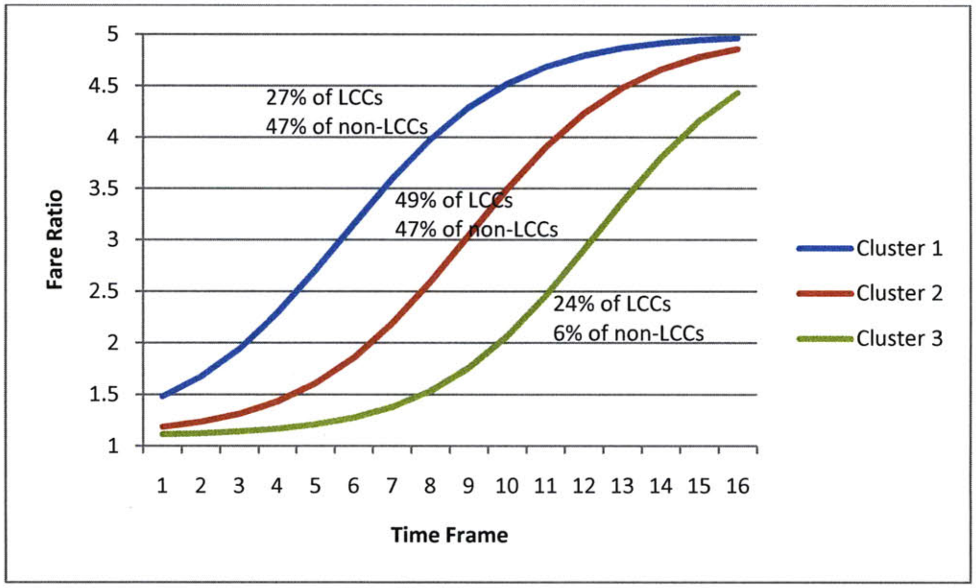


Figure 61: FP 2-Parameter 3 Cluster FRAT5 Curves in Network T4

As mentioned previously, clustering also provides an advantage in Network T by not forcing LCC and non-LCC markets to obtain separate FRAT5 curves, as it did in the system-based aggregation level. Figure 61 shows the importance of not separating the markets. While the middle FRAT5 curve has a similar mix of LCC and non-LCC clustered markets, the highest curve obtains most of the non-LCC markets, but also still has some of the clustered LCC markets. The opposite happens for the lowest curve. This means non-LCC markets generally receive more aggressive sell-up estimates compared to the LCC markets, but it is not impossible for an LCC market to have high degrees of sell-up. Keep in mind that LCC markets are often limited by the fare ratios of the fare classes within the market. Also, having a lot of sell-up (more in LCC markets with the unrestricted fare structure) does not necessarily imply a higher, more aggressive FRAT5 curve.

Network T1

Based on the performance of the clustering in Network D6 Unrestricted and Network T4, the method appears to be very strong in situations where sell-up is likely to occur. However, Network T1 provides answers about the clustering method's applicability in environments where less sell-up exists, ensuring that it is not detrimental to performance.

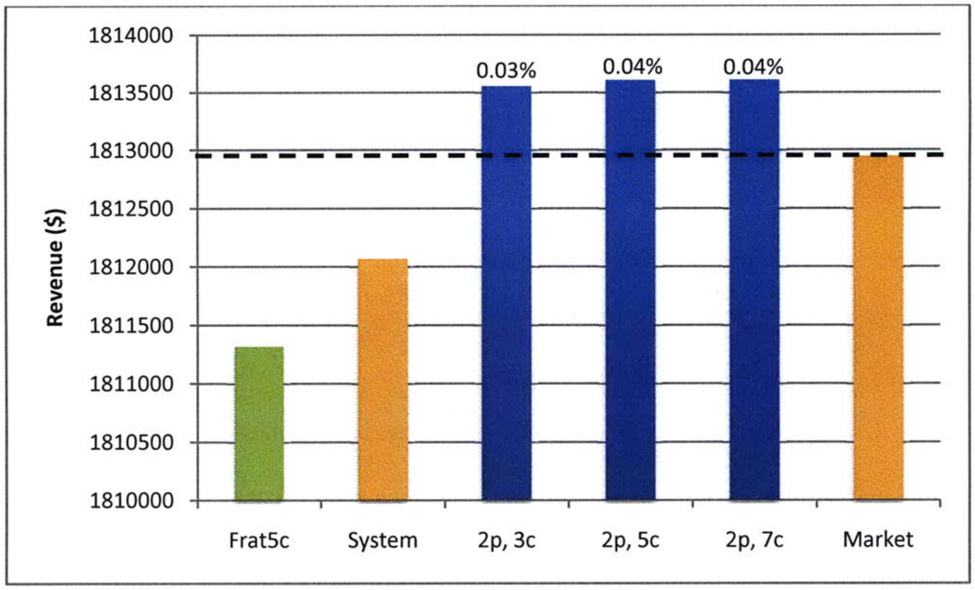


Figure 62: FP 2-Parameter Cluster Revenue in Network T1

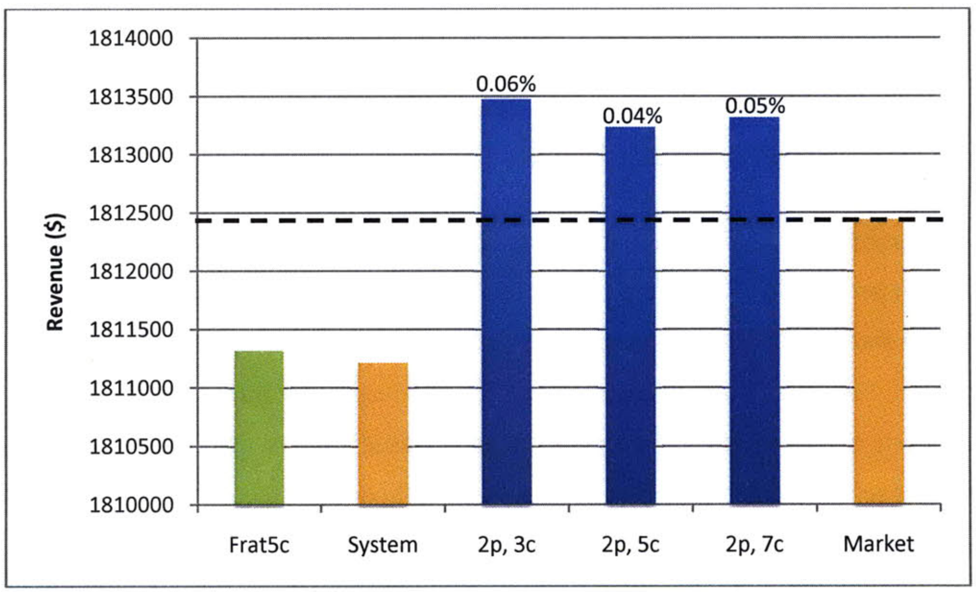


Figure 63: DO 2-Parameter Cluster Revenue in Network T1

Compared to the best previous aggregation level (market for both FP and DO cases), clustering provides a modest 0.04% and 0.06% increase in revenue for FP and DO, respectively. However, compared to the FRAT5c, clustering produced revenue increases of 0.13% and 0.12% for FP and DO, respectively. While these increases are not nearly as high as Network D6 Unrestricted or Network T4, it is important to keep in mind that the fare structures for the LCC markets are much more restricted compared to Network T4. In addition, despite being a small increase, the clustering method still improved revenue in Network T1.

Raising the Demand Multiplier to Increase Sell-up Observations

Despite the clustering method's increases in revenue in all three networks that were tested, the proportion of markets that did not estimate logistic parameters and therefore were not clustered is significant, as shown in Table 20.

		# Markets	% Markets
Network D6 Unrest.	FP	26	5.4
	DO	26	5.4
Network T4	FP	325	56.8
	DO	331	57.9
Network T1	FP	396	69.2
	DP	402	70.3

Table 20: Markets without Sell-up Parameters per Network

While only 5.4% of markets in Network D6 Unrestricted do not have sell-up logistic-fit parameters, the network has a completely unrestricted fare structure. More realistic markets that include more restricted fare structures, especially for the non-LCC markets, have many more markets without fits, about 57% in Network T4 and about 70% in Network T1.

The lack of observations with logistic-fit parameters could mean several things. First, if a market does not have logistic-fit parameters, there were not enough occurrences of sell-up between fare classes to create enough time frame FRAT5 estimates and therefore a logistic fit. However, because these markets still receive an input FRATc (a moderate level FRAT5 curve reaching a maximum FRAT5 value of 3.0), they still have an estimate of sell-up. Perhaps this is not an issue because the FRAT5c may actually be a good description of sell-up in the market, when it does actually occur. Second, if a market uses an input FRAT5c, this in a way ruins the data-based estimation of sell-up by adding an arbitrary curve. However, an argument against this is that if sell-up does not occur often in the market, then the estimate of sell-up does not really matter.

In an effort to reduce the number of markets not having logistic-fit parameters, thus increasing the number of markets with data-based sell-up estimation, one method is to increase the demand level in PODS. The demand level is simply a constant multiplied to the passenger generator in PODS. If there are more bookings, there will be more occurrences of sell-up, and thus more markets with a logistic-fit. Increasing the demand multiplier (DM) also obviously increases the load factors beyond what is reasonable in the real world. With the original DM for Network T4 set to 1.0, AL1 achieves load factors of approximately 79 percent. The methodology for this experiment is to increase the DM to a higher level and obtain new logistic-fit parameters for each market, cluster the markets based on those parameters, and rerun PODS *at the original DM of 1.0* with each market receiving the logistic-fit parameters from the new cluster means. Increasing the DM produces the following effect on the number of markets without logistic-fit parameters, as shown in Figure 64.

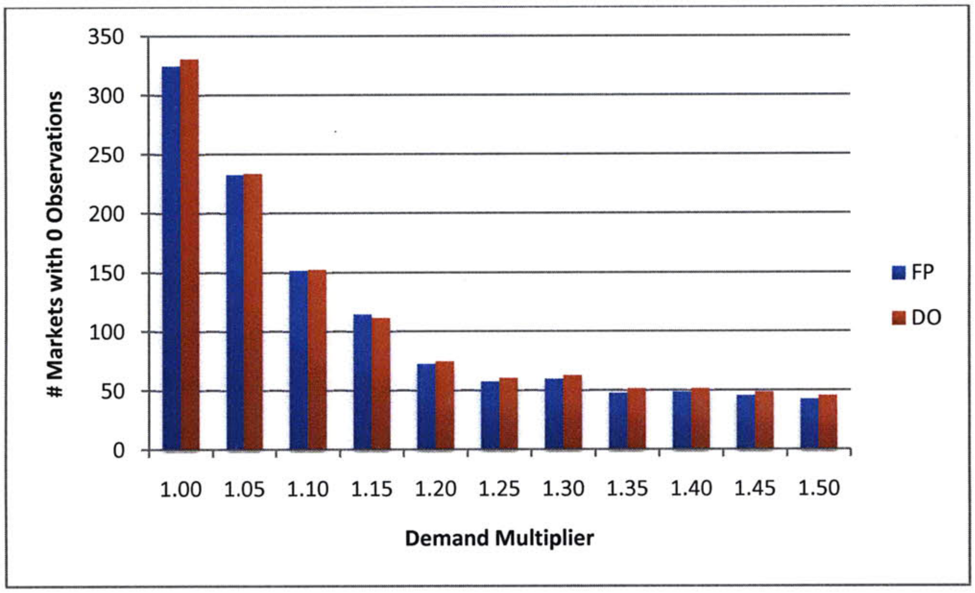


Figure 64: Effects of Increasing DM to Generate Sell-up Parameters in Network T4

By DM = 1.25, the number of markets with zero observations and thus zero logistic-fit parameters levels off to reach about 50 markets, or about 9 percent of the total markets in Network T. However, between DM = 1.25 and DM = 1.50, the load factors range between 86 and 89 percent, which is less realistic than the DM = 1.0 case. In addition, given that a market has sell-up parameters, the number of observations ("good" samples, as described earlier) that create those parameters increase with a higher DM, as shown in Figure 65.

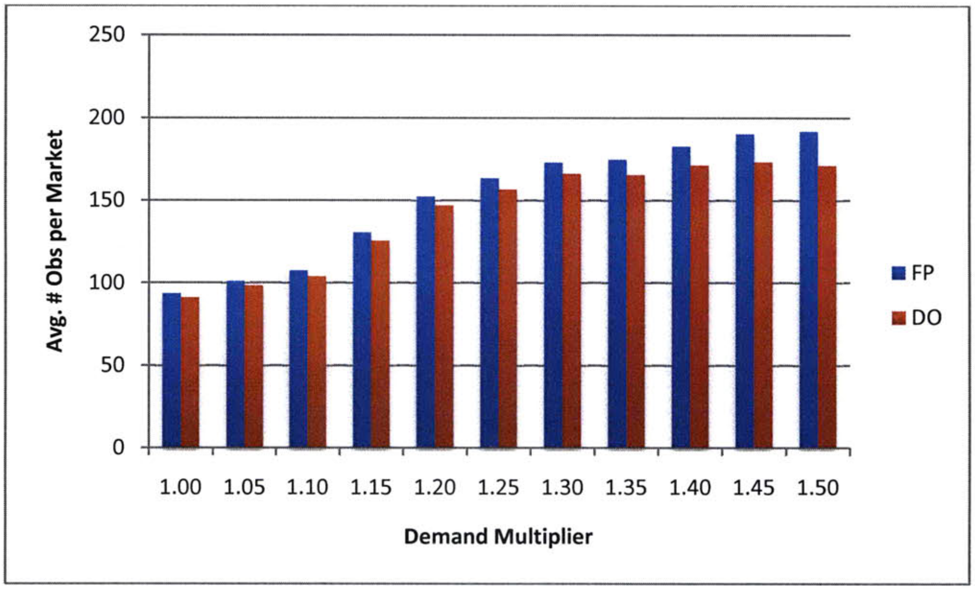


Figure 65: Effects of DM on Average Number of Observations of Sell-up per Market in Network T4

For FP, increasing the DM to 1.25 causes about a 75 percent increase in the average number of observations of sell-up per market, and increasing to a DM of 1.50 creates a 105 percent increase in the average number of observations. Because of this, not only will more markets have logistic-fit parameters, but the number of observations used to create the parameters increases, making the parameters more accurate.

One argument against this methodology is that the new sell-up estimates will not be accurate, and be too aggressive, thus overestimating a passenger's willingness-to-pay. To examine this further, the evolution of the FRAT5 curves for each cluster will provide more insight. For the three cluster case, Figure 66 and Figure 67 show that the corresponding cluster curves (relabeled as Cluster #1 for the highest curve down to Cluster #3 for the lowest curve) increase in height and aggressiveness as the DM increases. The only anomaly is the decrease in height of Cluster #1 for the DO case as DM increased. This is likely due to the fact that at DM = 1.0, only 88 markets belonged to this cluster, whereas more accurate curves were developed at higher DMs with 147 markets in the cluster at DM = 1.50.

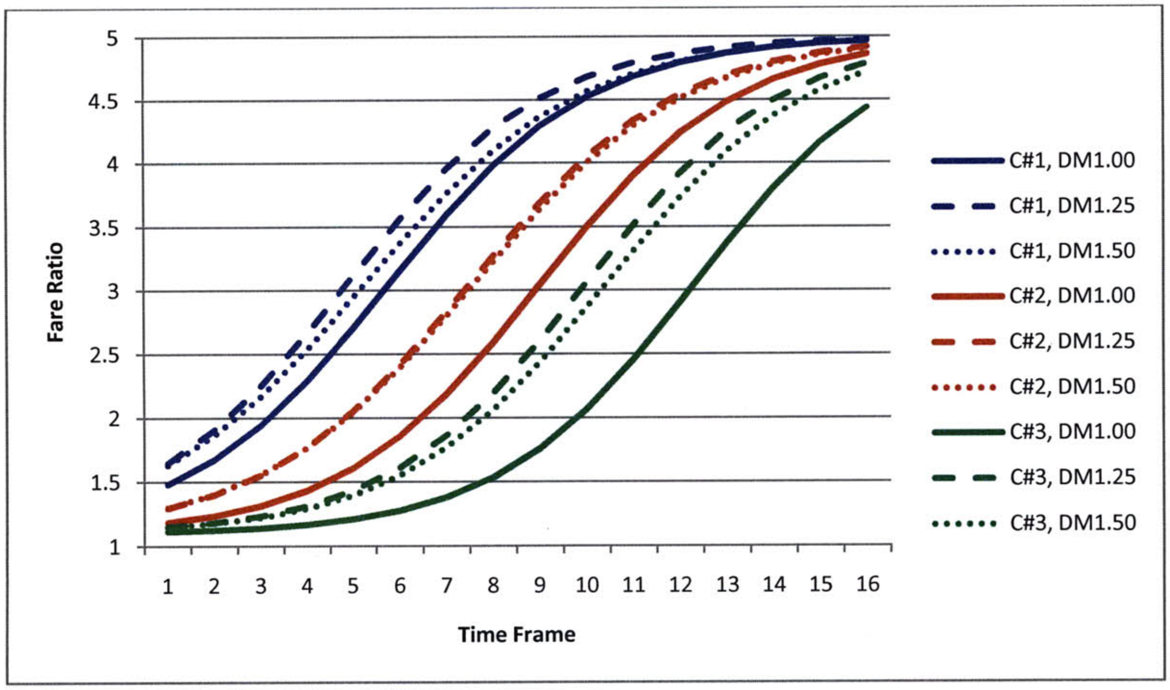


Figure 66: Effects of DM Increase on Cluster FRAT5s for FP in Network T4

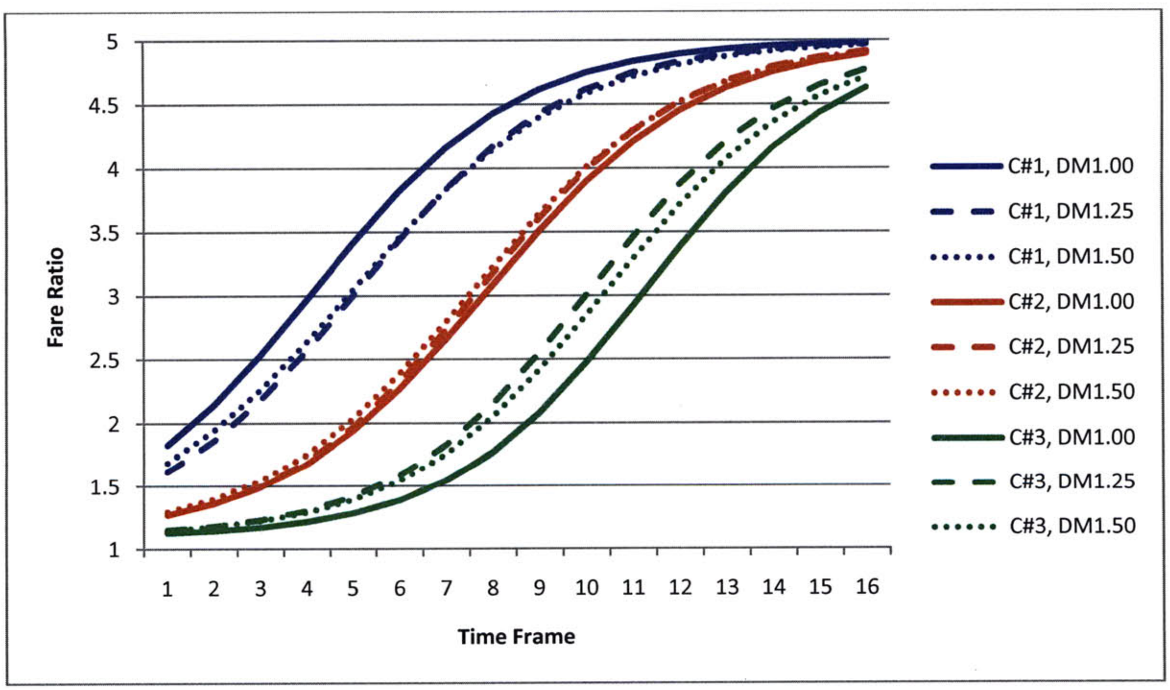


Figure 67: Effects of DM Increase on Cluster FRAT5s for DO in Network T4

Despite having fewer markets without sell-up observations, and increasing the average number of sell-up observations per market, the use of the new FRAT5 curves in PODS at the original DM level of 1.0 produces a small, but negative effect on revenue, as shown in Figure 68.

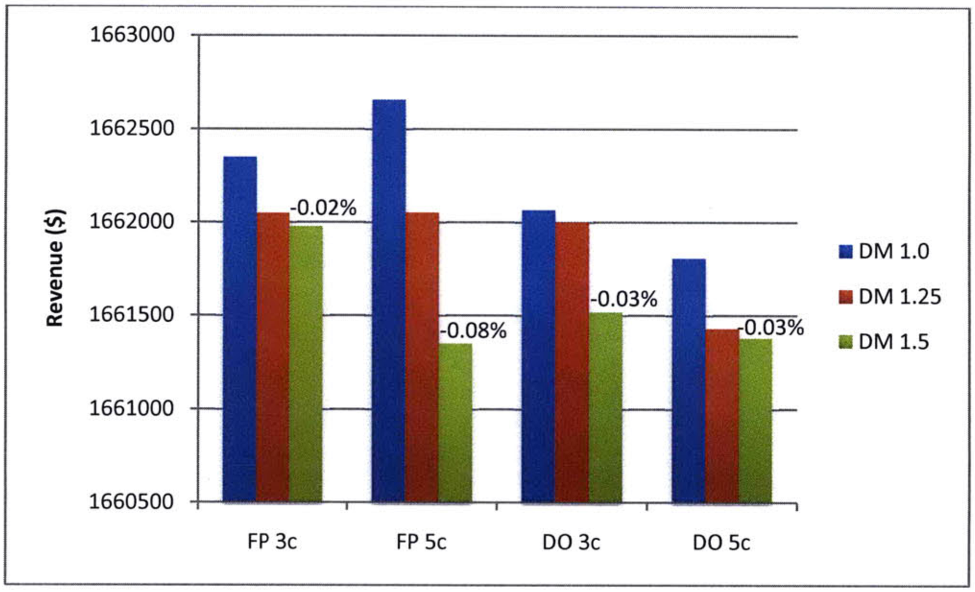


Figure 68: Revenue Impacts of the Change of Sell-up Parameter DM Level in Network T4

While the revenue losses are very small, the higher FRAT5 curves per cluster suggest that the sell-up estimates are too high, causing the overprotection of seats. In addition, this supports the hypothesis that markets with few sell-up observations do not need an accurate estimate of sell-up because there is not much sell-up to begin with. To better understand if this might be the case, it is important to look at which markets moved from having zero sell-up parameters, and to which cluster they moved to.

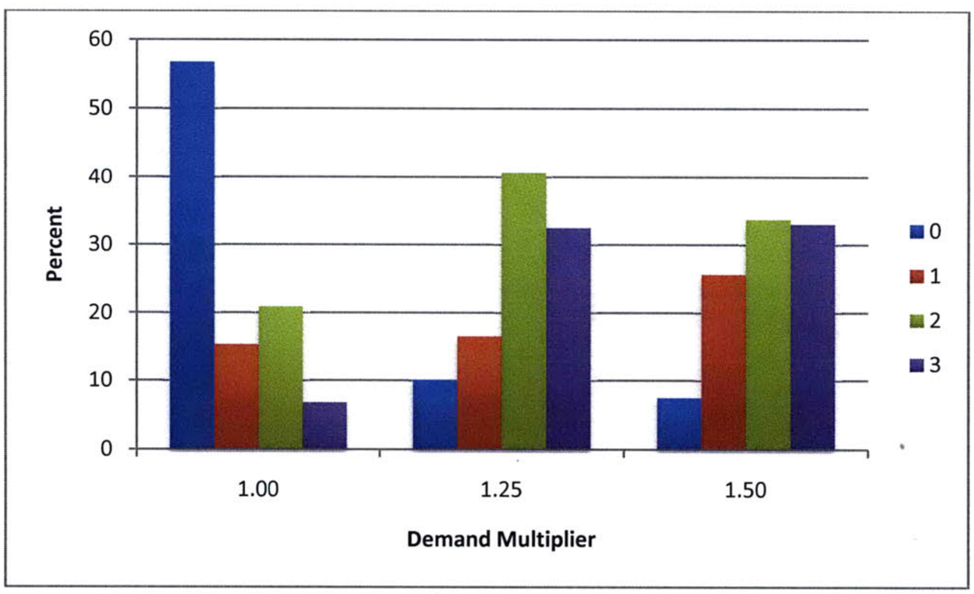


Figure 69: Cluster Distribution of Markets per DM for FP in Network T4

Just looking at the FP three cluster scenario, it is evident that increasing the DM to 1.25 caused most of the former Cluster "0" markets (those markets with no sell-up observations) to move into the lower clusters, mostly to Clusters #2 and #3. Increasing the DM further to 1.50 caused a greater upward shift of markets into clusters with higher FRAT5 curves. Similar results also occur for the DO three cluster scenario.

Table 21 provides further information about what specific clusters the markets move into upon gaining logistic-fit parameters at the new DM of 1.25. Those markets that originally had zero observations are identified as belonging to "Cluster #0".

Moving from Cluster #0 into:

Cluster	Total	LCC	% LCC	Avg Yield	Avg Fare Ratio	Avg % Business
0	57	41	71.9	0.202	4.033	39.6
1	32	9	28.1	0.271	4.693	43.2
2	124	36	29.0	0.208	4.505	45.6
3	112	75	67.0	0.150	4.038	35.8

Table 21: Market Location after Increasing DM for FP 3 Cluster Scenario in Network T4

Of the original 325 markets (56.8% of the total markets), only 57 of them remain in Cluster #0, still not receiving and logistic-fit parameters for sell-up. However, as mentioned before, most of the markets move into the lower two clusters. Of the markets that move to Cluster #3, about two-thirds of them are LCCs, while the markets that move into Cluster #2 are predominately non-LCCs. It is also interesting to see that there is a relationship between the cluster that the market moved into and the characteristics of the markets in those clusters. For example, those markets that moved into a higher cluster have a higher average yield, implying that a higher FRAT5 curve has a positive correlation with average yield in a market. In addition, the average fare ratio of the markets increases with higher FRAT5 curves. Finally, in relation to the percent of LCCs, the average percent of business passengers in the newly-assigned markets is generally higher for higher FRAT5 curves. This information provides a good transition into the next topic within clustering.

While the PODS simulations provide answers about the clustering method's applicability, many questions still exist about why a market is assigned to a specific cluster. What drives cluster membership beyond that of the simple logistic parameters that estimate sell-up? Are there any specific market characteristics that are common in markets belonging to a particular cluster? For example, to what degree do characteristics such as yield, whether or not the market is an LCC, business passenger percentage, fare ratio, load factor, or whether or not the market has a route advantage play a role in a market's cluster categorization? These questions will be addressed through detailed statistical regression methods in Chapter 7.

6.4.2.2. Clustered Piecewise

As discussed in Section 6.3, the input piecewise FRAT5 curve performed very well in all three networks that were tested. Combining the performance of the piecewise FRAT5 step function with the advantages of the aggregation levels of clustering, it is possible to create a data-based, clustered piecewise FRAT5. Another reason for pursuing this direction is that the piecewise step function FRAT5 offers some advantages over the logistic-fit curves.

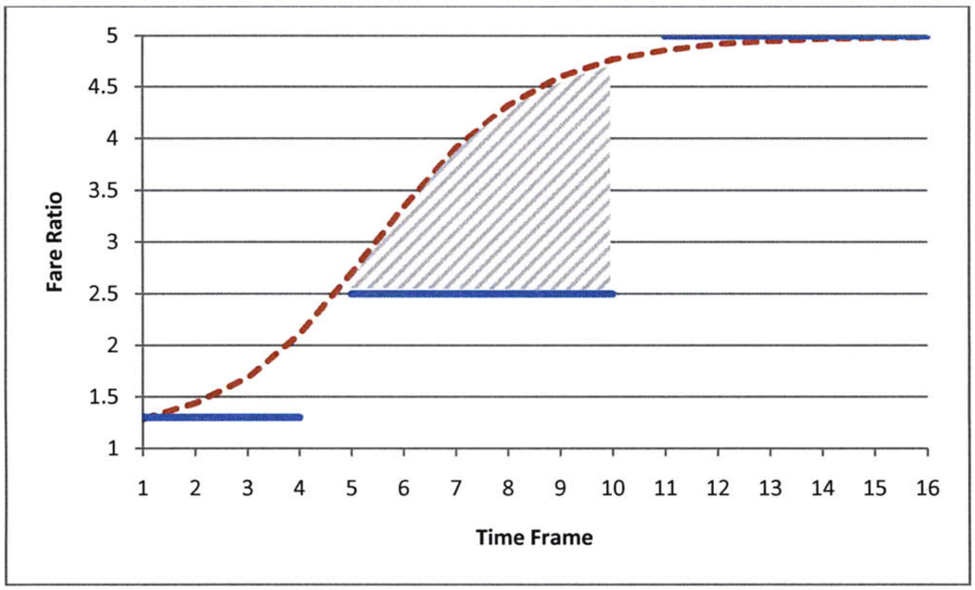


Figure 70: FP Logistic-fit versus Best Piecewise FRAT5 Curve in Network D6 Unrestricted

Compared to the FP system-based logistic-fit, the piecewise FRAT5 (which performed better than the logistic-fit curve) estimates much lower levels of sell-up between TF 5 and TF 10, where the logistic-fit curve is more aggressive. This is essentially the nature of any logistic-fit curve, as it is constructed on only two parameters. Therefore, instead of fitting up to 16 time frame observations of FRAT5 values to a logistic-fit curve, which may be over-aggressive in middle time frames, the same data points could be used to construct a piecewise step function FRAT5 curve. This fulfills the goal of constructing a data-based piecewise FRAT5 curve. In addition, employing the advantages of the clustering process ensures that markets will not be under or over generalized.

In a given PODS sample, a FRAT5 value is reported for a time frame for each market if there were at least two occurrences of sell-up in the previous 26 historical departures. To create a piecewise clustered FRAT5, one must first divide the 16 time frames into three periods. For this experiment, the same periods as the original piecewise curve are used—TF 1-4, TF 5-10, and TF 11-16. Because a market is not guaranteed to obtain a FRAT5 value for every time frame, the average value over the period is used.

In order to deal with outliers (mostly FRAT5 estimates derived from very few observations), if an observation is more than one standard deviation away from the mean for that period, then the observation is removed. However, for period one, this only applies to the upper standard deviation above the mean. Likewise, for the period three, this only applies to the standard deviation below the mean. This ensures that the FRAT5 curve is not influenced by outliers and that it generally increases over time, in accordance with the assumption that later booking passengers have a higher willingness-to-pay. In Figure 71 below, the data points in the shaded region are removed.

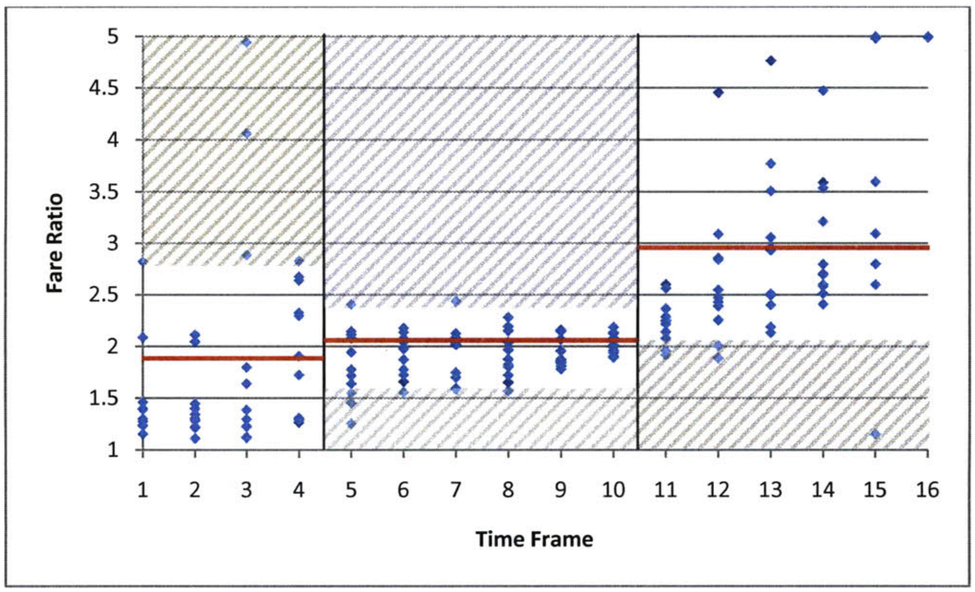


Figure 71: Removal of Outliers in Clustered Piecewise FRAT5 Curves

If more than four out of 16 time frame FRAT5 values are missing for a market, that market is removed from the clustering process and receives an input FRAT5c. For each market that may be clustered, there exist three average FRAT5 values—one per period—that serve as the basis for clustering. Then, following the same process for the logistic-fit parameter clustering, the markets are clustered, each given a new set of 16 FRAT5 values (based on the three cluster means), and re-run in PODS.

Network D6 Unrestricted

Compared to logistic-fit clustering, piecewise clustering over the average FRAT5 per time period creates much less aggressive FRAT5 curves, as shown for the FP estimation case in Figure 72. This is likely due to the fact that for period three, FRAT5 observations from early time frames in the period like TF 11 or 12, are much lower and bring down the average FRAT5 for the period. For logistic-fit clustering, the curve is able to adapt to lower values in earlier time frames and still reach the maximum FRAT5 value of 5.0, as shown by the gray curves in Figure 72.

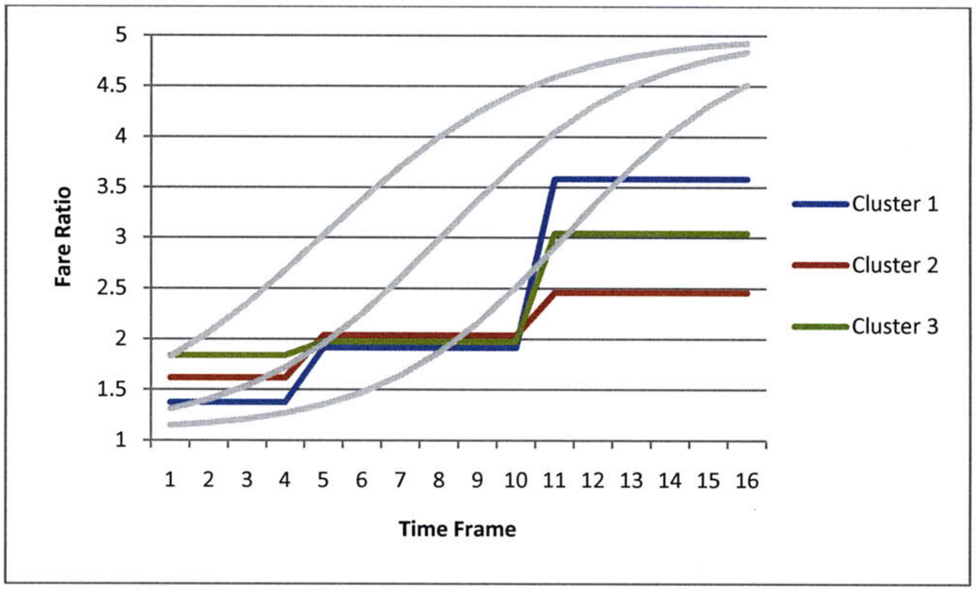


Figure 72: FP Clustered Piecewise FRAT5 Curves in Network D6 Unrestricted

These less aggressive FRAT5 curves in an unrestricted network negatively impact AL1's performance. In fact, the best clustered piecewise FRAT5 curves perform worse than the baseline input FRAT5c. This is due to the very high load factors and low yields, inherent to a case where sell-up is underestimated, causing high fare class seat protection to spiral down.

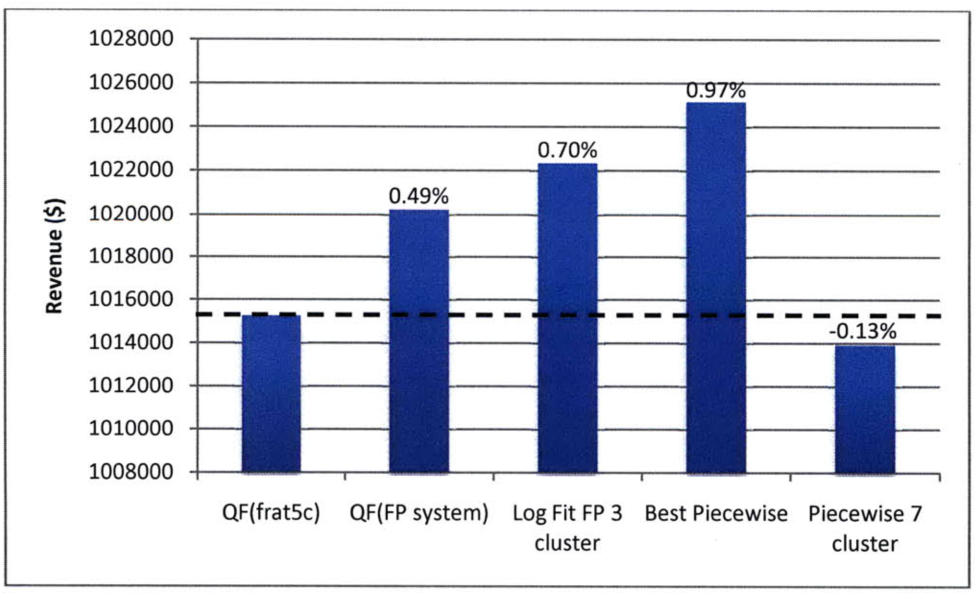


Figure 73: FP Clustered Piecewise Revenue in Network D6 Unrestricted

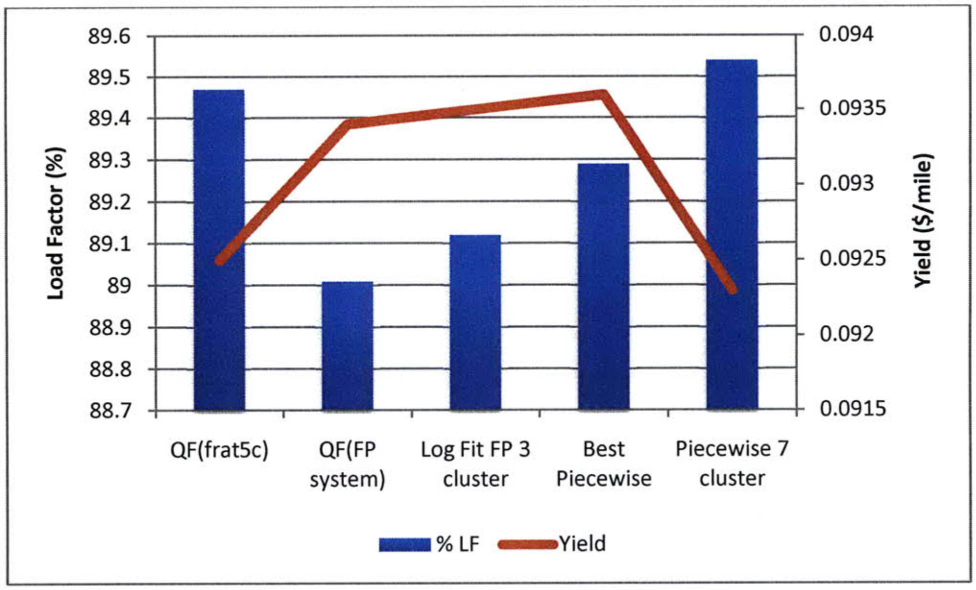


Figure 74: FP Clustered Piecewise Load Factors and Yields in Network D6 Unrestricted

Unfortunately the clustered piecewise FRAT5 curves performed much worse than FP system-based sell-up estimation, FP logistic-fit clustering, and the input piecewise FRAT5 curve.

Network T4

Applying the clustered piecewise FRAT5 methodology to the larger, more complex Network T4 produces similar results.

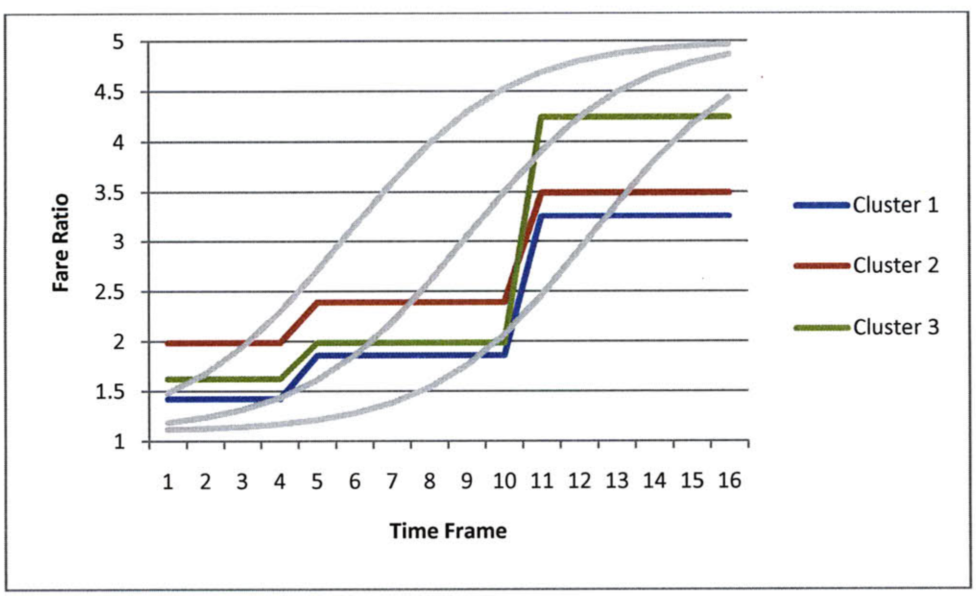


Figure 75: FP Clustered Piecewise FRAT5 Curves in Network T4

While the FRAT5 curves in the first two periods (TF 1 to TF 10) appear to be sensible, remaining lower than the logistic-fit curves (the goal of the clustered piecewise methodology), the curves in the third period (TF 11-16) are much lower than the logistic-fit curves. This is similar to the Network D6 Unrestricted curves, where the use of the average values in the third period prevents them from obtaining higher FRAT5 values.

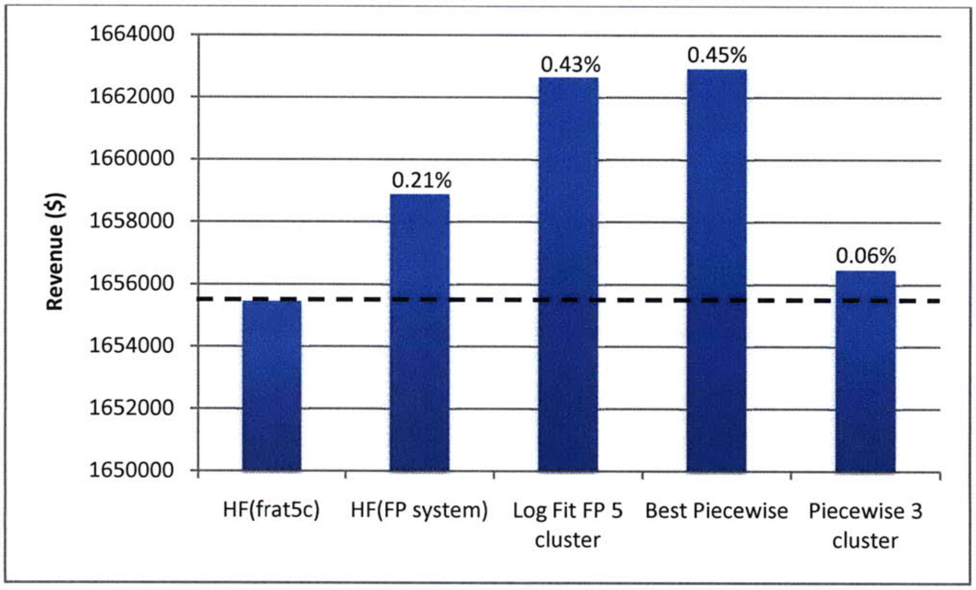


Figure 76: FP Clustered Piecewise Revenue in Network T4

As a result, the best piecewise clustering method (3 clusters) creates similar revenue as the input FRAT5c, which also has FRAT5 curves that remain lower than the logistic-fit clustering throughout the booking period.

Overall, creating a data-based piecewise FRAT5 for clustering does not improve revenue. A lot of data pruning must first occur to remove the effects of outliers, especially when using a method that depends solely on averages. Using the logistic fitter, outliers either had no effect on the curve, or they caused no fit to occur, resulting in a market using an input FRAT5c. Comparing the two methods, the logistic-fit more often enables the FRAT5 curve to reach its maximum value at 5.0 in later time frames. Having a curve in this area is crucial to the success of estimating sell-up in less restricted and unrestricted fare environments.

6.5 Rational Choice Forecasting

Contrary to the previous sections focusing on sell-up estimation in conjunction with Q and Hybrid Forecasting, the section focuses on Rational Choice (RC) Forecasting, which incorporates the sell-up estimates within each forecast. As discussed in Chapter 4, RC Forecasting is a regression-based forecasting method that creates partitioned passenger type forecasts based on the historical bookings and the lowest open class at the time of the bookings. Unlike Q and Hybrid Forecasting, RC Forecasting allows a booking in the lowest open fare class to not automatically be classified as a price-oriented booking—rather customer types based on passenger willingness-to-pay create other possibilities. Applying RC Forecasting in PODS, though based on a completely different methodology than QF and HF, still creates competitive results for an entirely data-based forecasting method.

6.5.1. RC Regular and Expanded

Hybrid Forecasting (with both input and data-based FRAT5s) serves as the benchmark for testing the performance of RC Forecasting in PODS. Not only are the revenue, load factor, and yield results important for comparison, but also the class forecasts themselves. In the following simulations, AL1 uses EMSRb while all other airlines use their baseline optimization and forecasting methods. To begin, Network D6 Semi-restricted serves as a basic environment for forecast comparison.

Network D6 Semi-restricted

In Network D6 Semi-restricted, RC Forecasting performs reasonably well, but not quite to the level of Hybrid Forecasting, as shown in Figure 77.

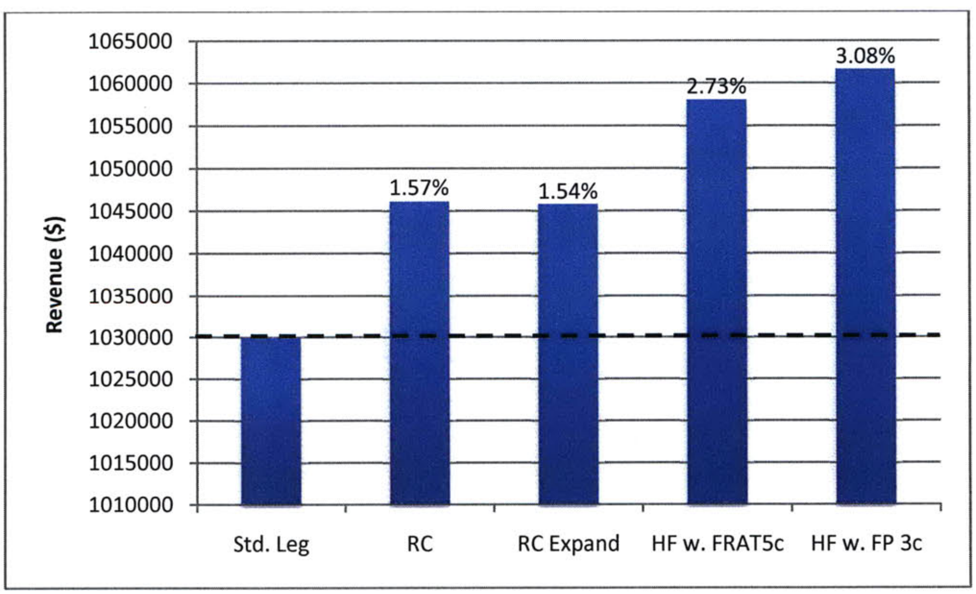


Figure 77: RC Revenue in Network D6 Semi-restricted

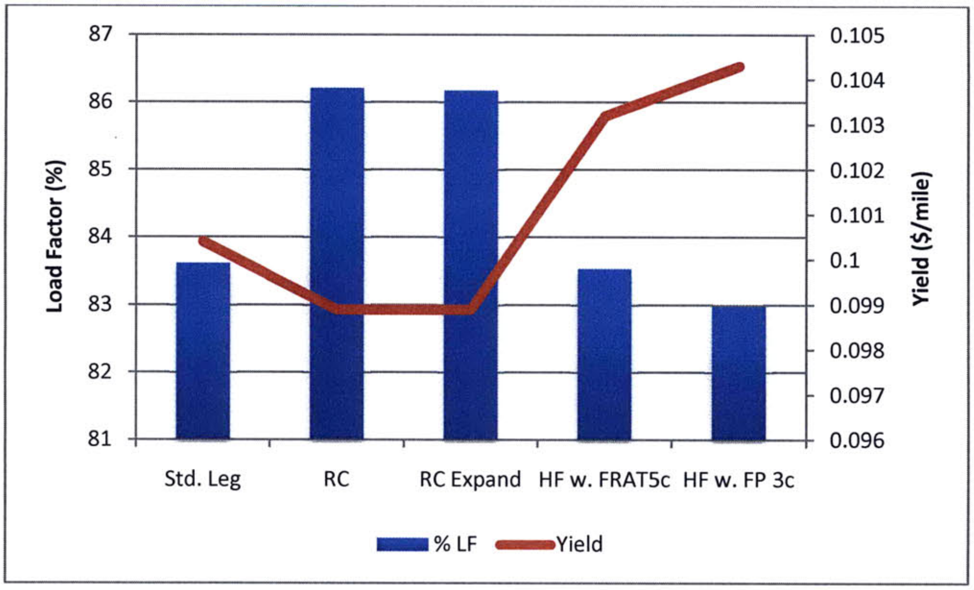


Figure 78: RC Load Factors and Yields in Network D6 Semi-restricted

Rational Choice Forecasting creates an increase in revenue over standard leg forecasting that is about half as much as Hybrid Forecasting with FP logistic-fit clustered sell-up estimation. Also, there is a minimal difference between the RC Expanded method and the regular RC method. It is evident that RC Forecasting creates much higher load factors and lower yields than Hybrid Forecasting, which is most likely due to high forecasts for the lower classes. To examine if this is the case, the actual RC forecasts and bookings provide some insight.

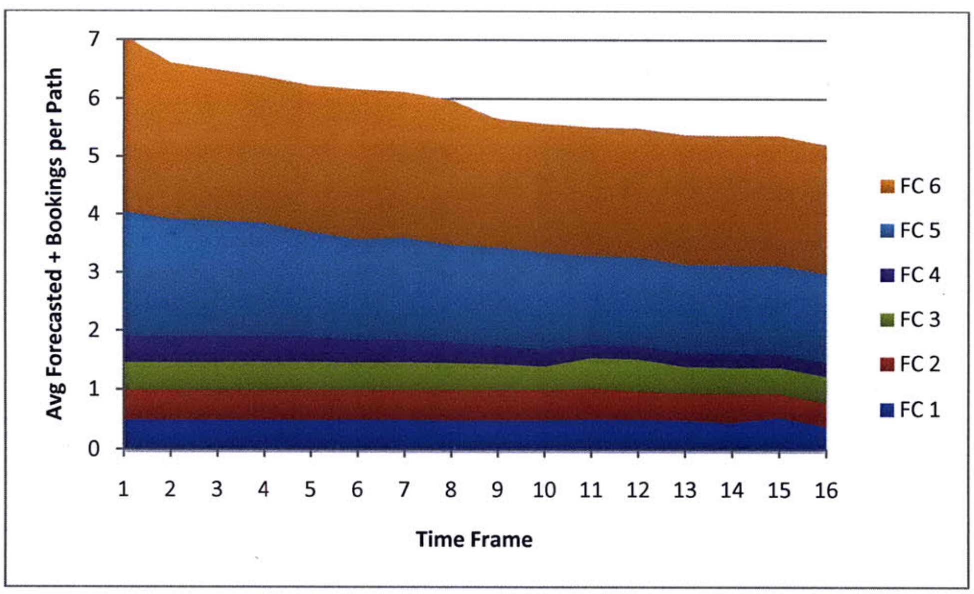


Figure 79: Forecast + Bookings per Path: EMSRb with RC in Network D6 Semi-restricted

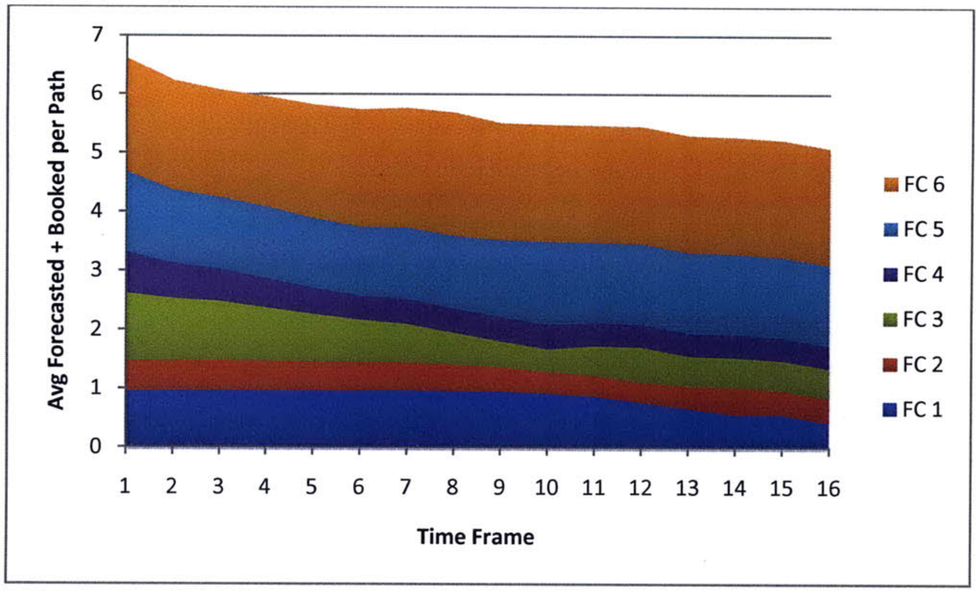


Figure 80: Forecast + Bookings per Path: EMSRb with HF (FRAT5c) in Network D6 Semi-restricted

Comparing the initial forecasts between RC and HF, it is clear that RC Forecasting has much higher forecasts for the lower classes, especially for FC 5 and FC 6. In addition, the forecasts for higher classes, especially FC 1 and FC 3 are much lower throughout the booking period for RC than for HF. Having this higher proportion of FC 5 and FC 6 passengers shows that RC is susceptible to spiral down, and it is the reason for the higher load factors, lower yields, and thus overall lower revenue when compared to Hybrid Forecasting.

Network T1

Network T1 provides the next test environment to determine how well RC Forecasting performs with more competition, but similar fare restrictions. In this scenario, RC again performs better than standard forecasting, but still not as well as Hybrid Forecasting.

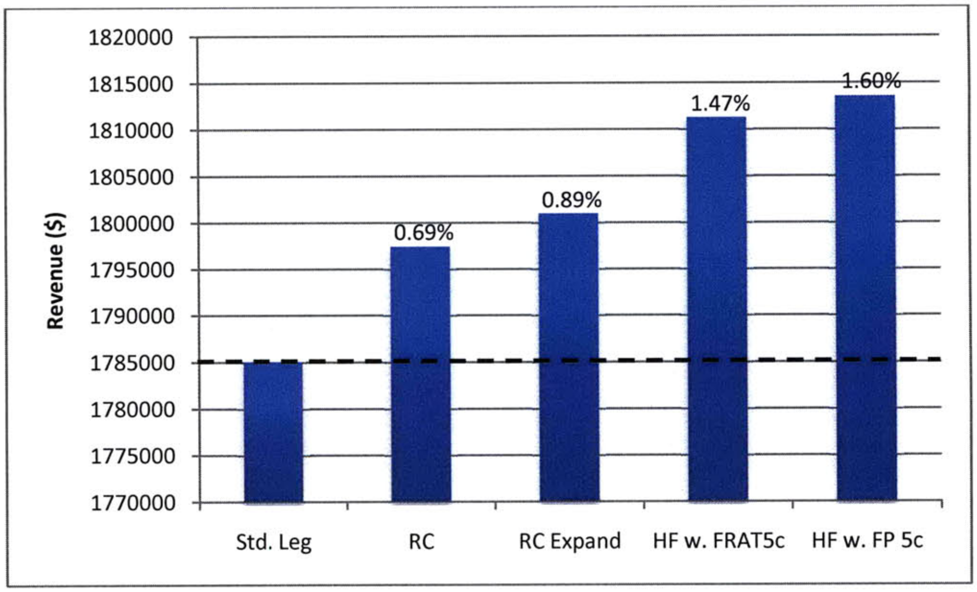


Figure 81: RC Revenue in Network T1

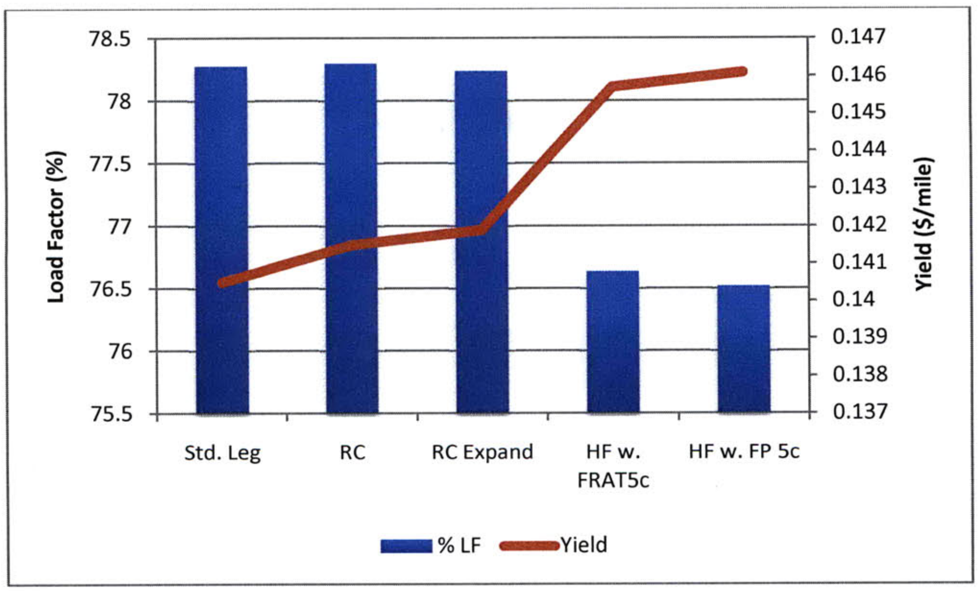


Figure 82: RC Load Factors and Yields in Network T1

Similar to the results in Network D6 Unrestricted, RC Forecasting produces about half of the revenue increase over standard forecasting as Hybrid Forecasting. Again, RC Forecasting produced very high load factors and lower yields compared to Hybrid Forecasting.

Looking at the fare class mix, it is clear that RC Forecasting allows more bookings mainly in FC 6, and creates fewer bookings in the higher FC 1 and FC 2 classes. This contributes largely to the higher overall load factors and lower yields for RC Forecasting compared to Hybrid Forecasting.

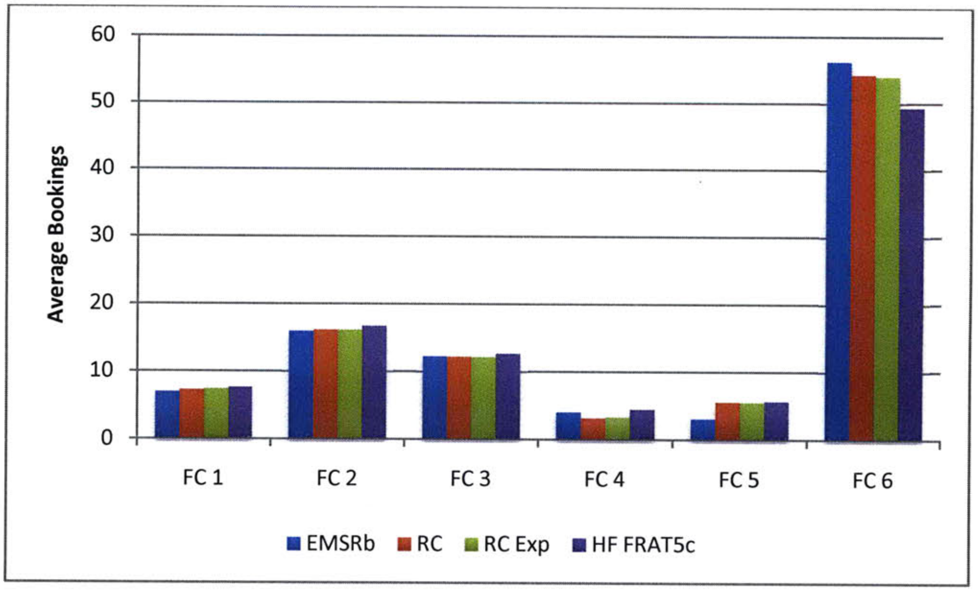


Figure 83: RC Fare Class Mix in Network T1

However, one interesting item to note is that while Hybrid Forecasting performs better in terms of revenue, it is not because of its accuracy in forecasting, but rather its inaccuracy.

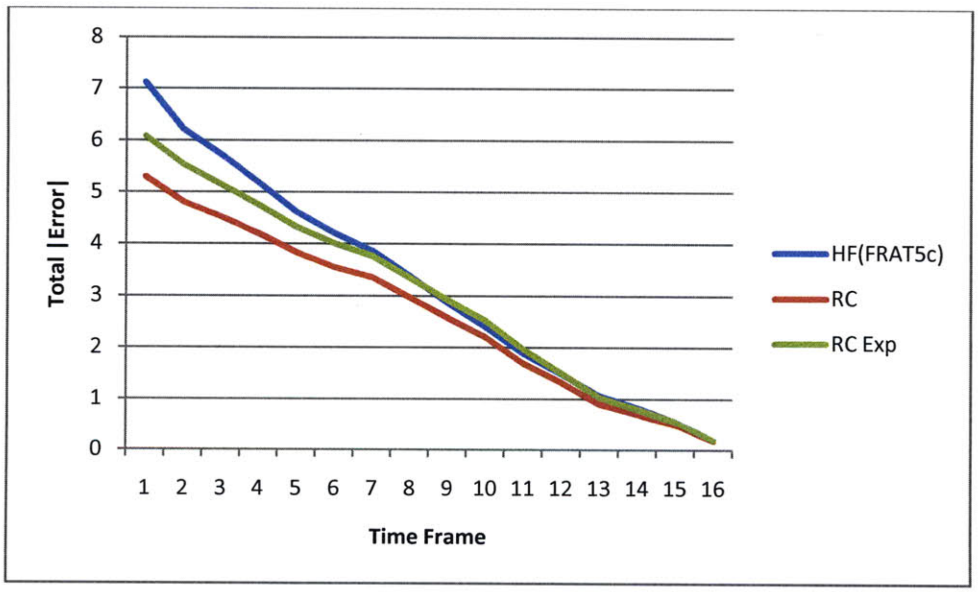


Figure 84: RC Total Absolute Forecasting Error in Network T1

According to Figure 84, Hybrid Forecasting has the most total absolute error between forecasting and actual bookings throughout the entire booking period. Upon further examination, most of this error is due to the over-forecasting of the higher classes, especially FC 1, and the under-forecasting of FC 6. This is extremely important to note because accuracy in forecasting does not always imply better revenue. Because Hybrid Forecasting's sell-up estimates increase the forecasts for the higher classes, it is able to keep them open longer, closing lower classes earlier.

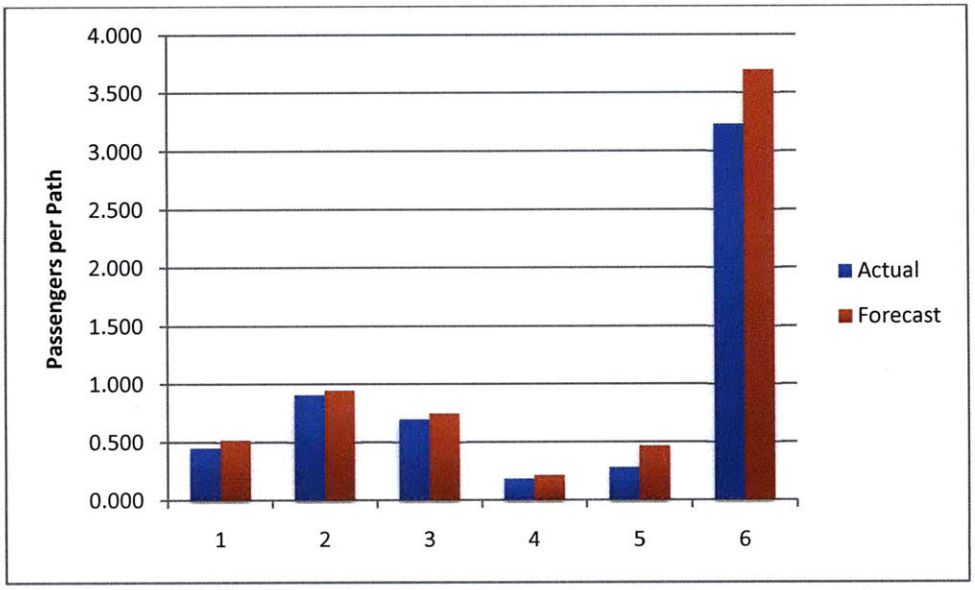


Figure 85: Rational Choice Forecasting Forecast vs. Actual Bookings in Network T1

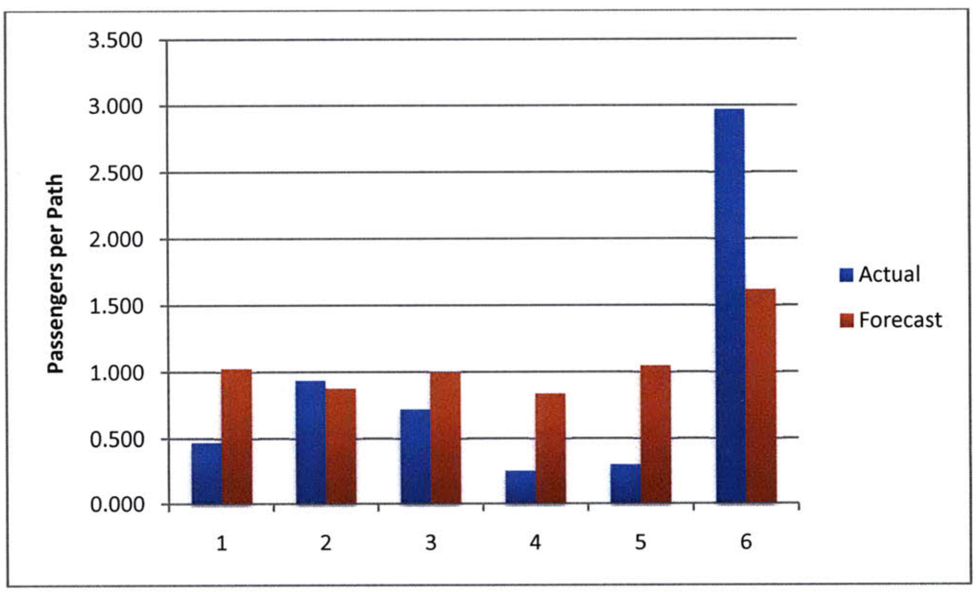


Figure 86: Hybrid Forecasting (FRAT5c) Forecast vs. Actual Bookings in Network T1

Looking at Figure 85 and Figure 86, it is clear that most of RC Forecasting's initial forecast is allocated to FC 6, whereas Hybrid Forecasting's forecast is more level across all fare classes, with a large percentage devoted to FC 1. While this does not actually occur with actual bookings, the forecast is enough to force more bookings into higher classes, reducing the load factor and increasing yield.

Network T4

To determine how Rational Choice Forecasting performs in a network with less restricted fares and more opportunity for sell-up, Network T4 serves as a viable test environment. Keep in mind that for RC

Forecasting, sell-up is already incorporated in the forecast by the structure of the linear regression, whereas for Hybrid Forecasting, sell-up probabilities are external to the process and must be applied to the Q-class equivalent bookings in order to repartition into a fare class forecast. Looking at the results, it appears that Hybrid Forecasting, despite its complexity, still outperforms RC Forecasting in a high sell-up environment.

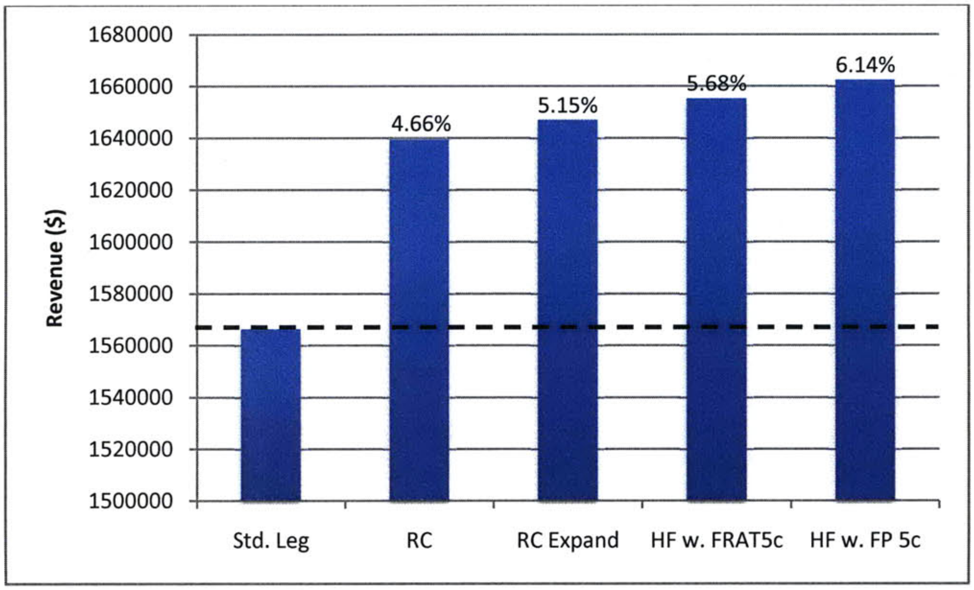


Figure 87: RC Revenue in Network T4

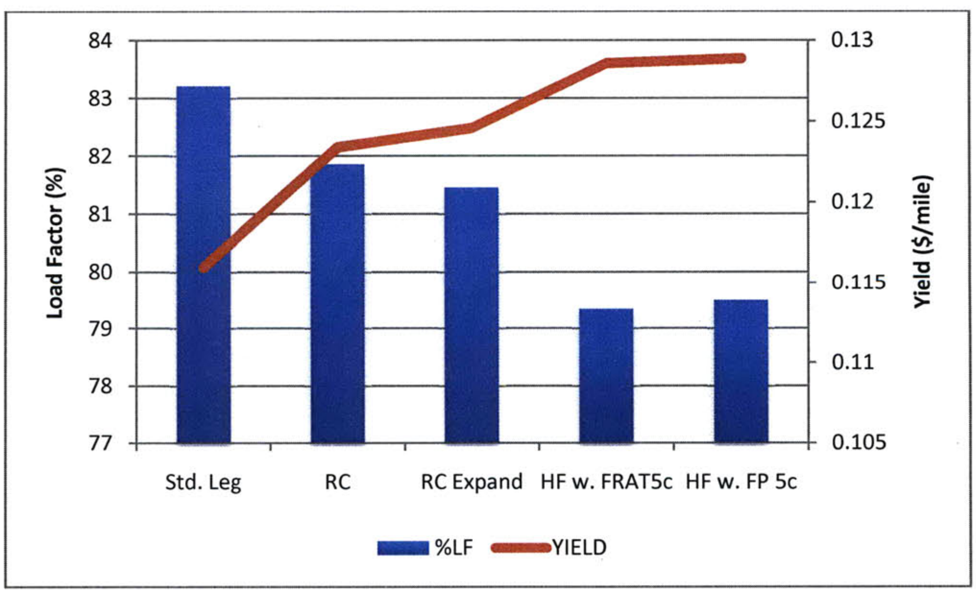


Figure 88: RC Load Factors and Yields in Network T4

RC Forecasting performs slightly worse than Hybrid Forecasting, with RC Expanded about one percent lower than the best clustering method for Hybrid Forecasting with FP sell-up estimation. Similar to previous results, RC Forecasting creates slightly higher load factors and lower yields when compared to

Hybrid Forecasting. To better understand why this occurs, the fare class closure percentages provide good insight.

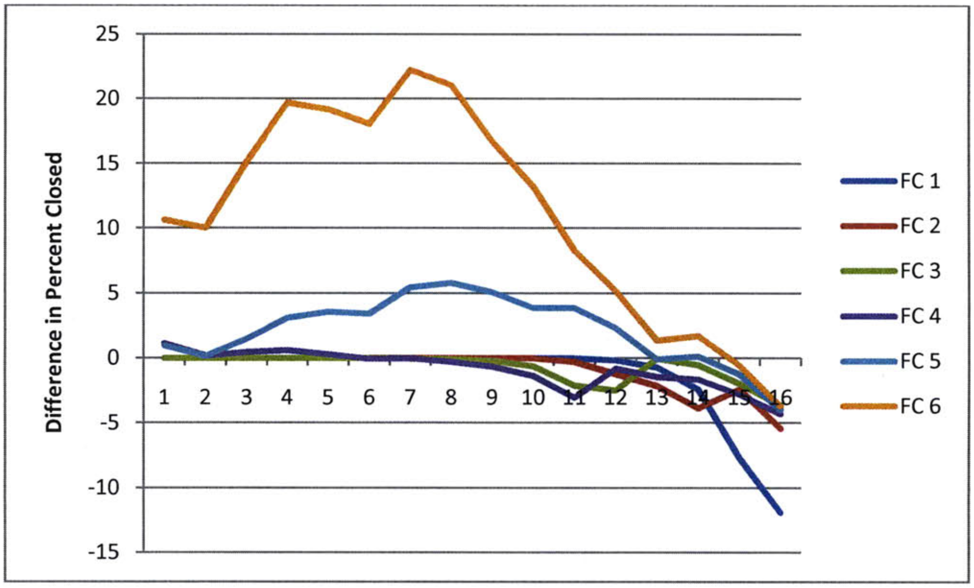


Figure 89: Difference in Fare Class Closures over Time: HF with FP (5c) minus RC Expanded in Network T4

According to Figure 89, the best Hybrid Forecasting option (FP sell-up estimation with 5 clusters) closes much more of the lower classes (FC 5 and FC 6) than RC Expanded. In addition, Hybrid Forecasting has FC 1 and FC 2 open longer at the end of the booking period, which is essential to gaining last minute business passengers, thus increasing the yield. The fare class closure explains why RC generally has higher load factors and lower yields compared to Hybrid Forecasting.

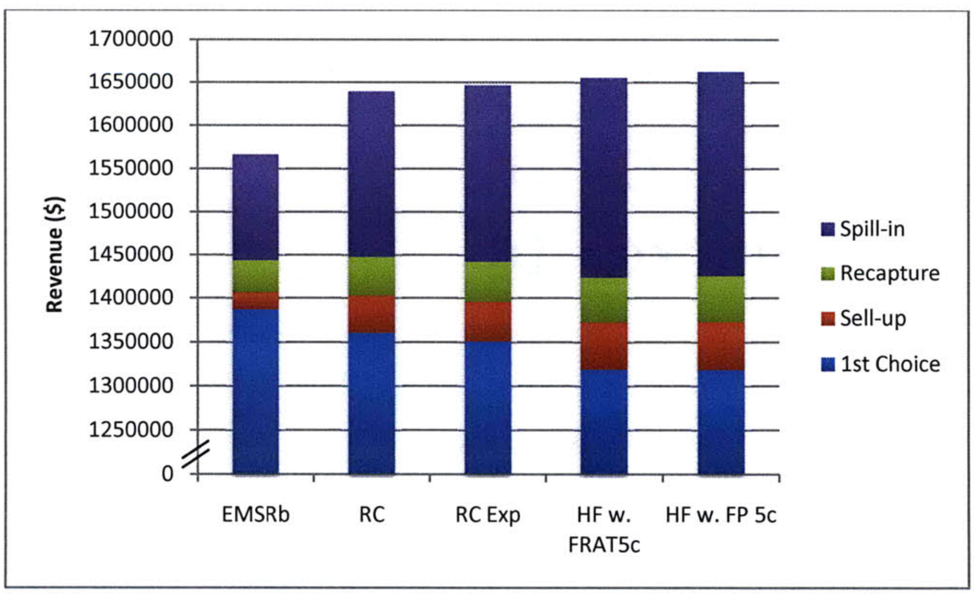


Figure 90: RC vs. HF Revenue Breakdown in Network T4

	EMSRb	RC	RC Exp	HF w. FRAT5c	HF w. FP 5c
Spill-in	7.8	11.7	12.5	14.0	14.3
Recapture	2.4	2.8	2.8	3.1	3.2
Sell-up	1.2	2.6	2.7	3.2	3.3
1st Choice	88.5	83.0	82.0	79.7	79.3

Table 22: RC vs. HF Revenue Breakdown Percent in Network T4

The fare class closure over time has many effects on the type of revenue gained by AL1 in each scenario. First, it is important to see the effects that both RC and HF have on sell-up revenue compared to EMSRb with standard pick-up forecasting, creating two to three times of standard forecasting's amount through the use of sell-up estimation. Compared to Hybrid Forecasting, more of RC Forecasting's revenue comes from "first choice" revenue, where a passenger initially planned and actually did purchase a particular fare class. This is due to Hybrid Forecasting's more aggressive fare class closures and high estimates of sell-up. In addition, this not only causes Hybrid Forecasting to gain more sell-up revenue than RC Forecasting, but also more spill-in and recapture revenue. Hybrid Forecasting has more seats available in the higher fare classes later in the booking period that not only permit sell-up, but allow space for recapture and spill-in passengers from other airlines.

Increasing the Forecast Multiplier

After comparing RC Forecasting to Hybrid Forecasting, it is apparent that Hybrid Forecasting is more aggressive in its forecasts. As a result, lower fare classes are closed down earlier and higher fare classes are kept open longer, lowering load factors and increasing yield through sell-up, spill-in, and recapture passengers. In an effort to increase the aggressiveness of RC Forecasting, this experiment uses a forecast multiplier (FM) to arbitrarily increase the forecast in hopes of matching Hybrid Forecasting. (Increasing the FM is only possible in PODS and would not be done in the real world.) In Network D6 Semi-restricted, AL1 uses EMSRb with standard RC forecasting with a forecast multiplier of 1.1, 1.2 and 1.3, and produces interesting results.

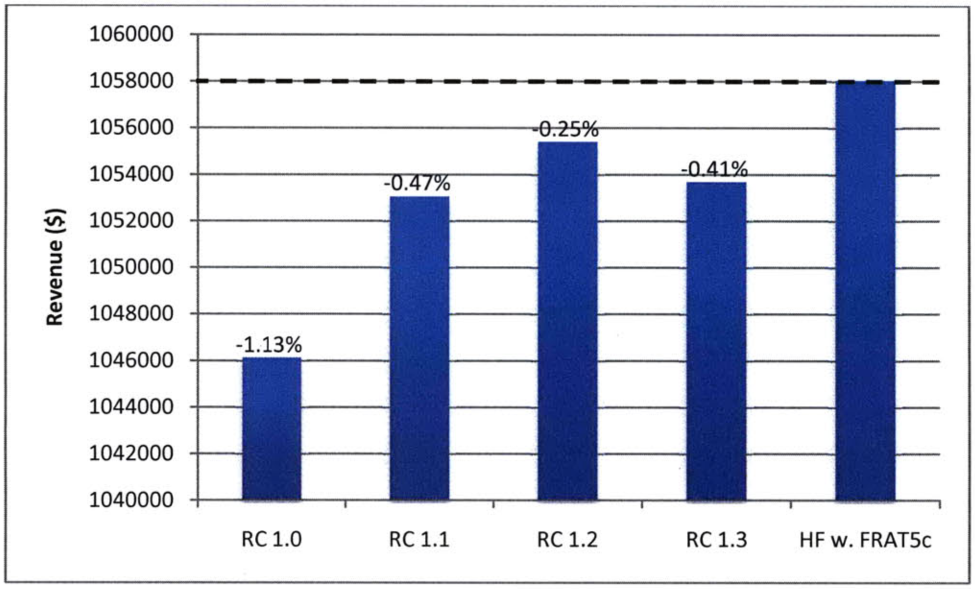


Figure 91: RC with Forecast Multiplier Revenue in Network D6 Semi-restricted

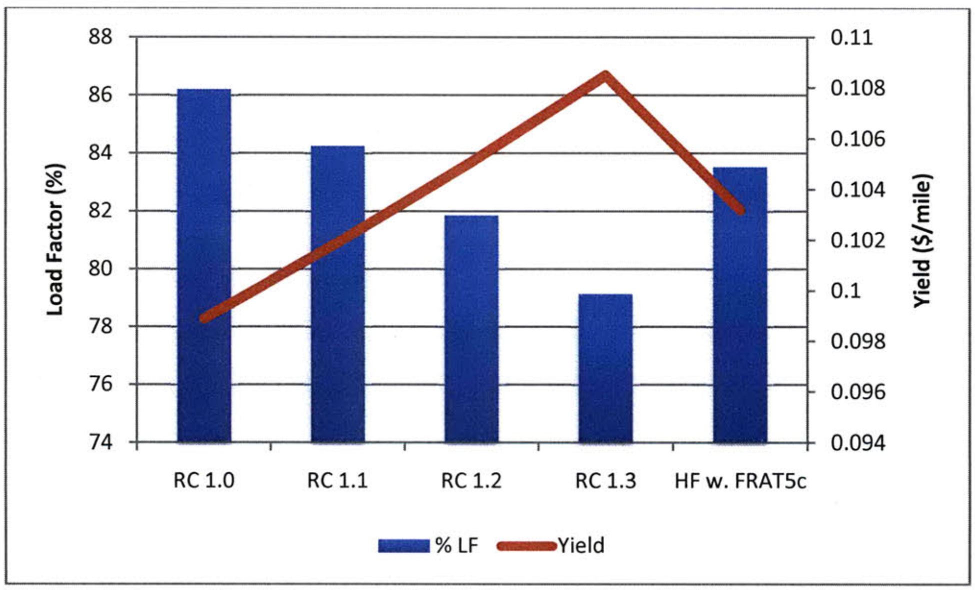


Figure 92: RC with Forecast Multiplier Load Factors and Yields in Network D6 Semi-Restricted

The forecast multiplier creates a bell-shaped curve for revenue, eventually reaching a point of being too aggressive at FM = 1.3. Increasing the FM caused load factors to drop and yields to increase, beyond that of the Hybrid Forecasting baseline. Comparing the revenue breakdown provides further understanding of why RC Forecasting with a FM still does not surpass the revenue performance of Hybrid Forecasting.

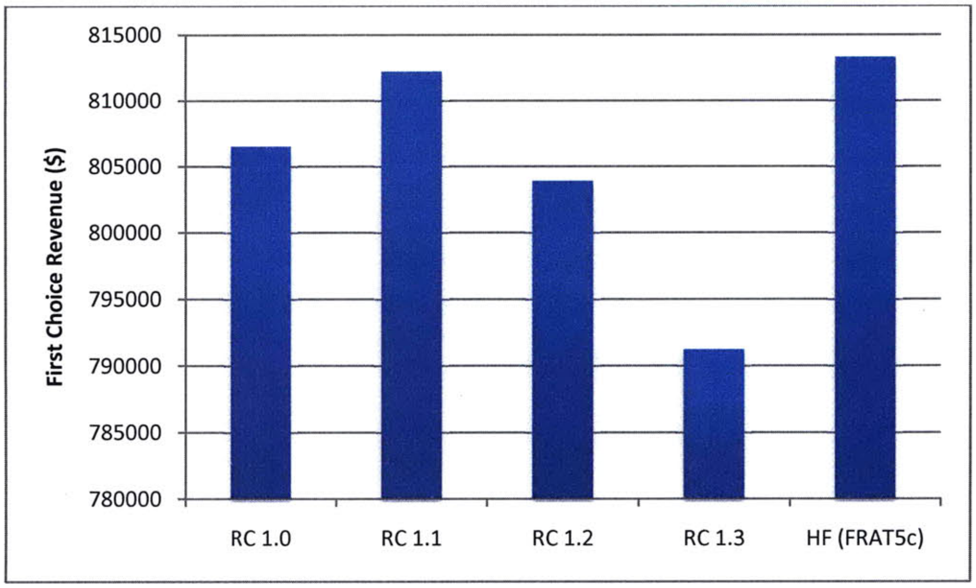


Figure 93: RC with Forecast Multiplier First Choice Revenue in Network D6 Semi-restricted

The biggest impact of the FM was on first choice revenue, where it peaked at FM = 1.1, but never surpassed Hybrid Forecasting.

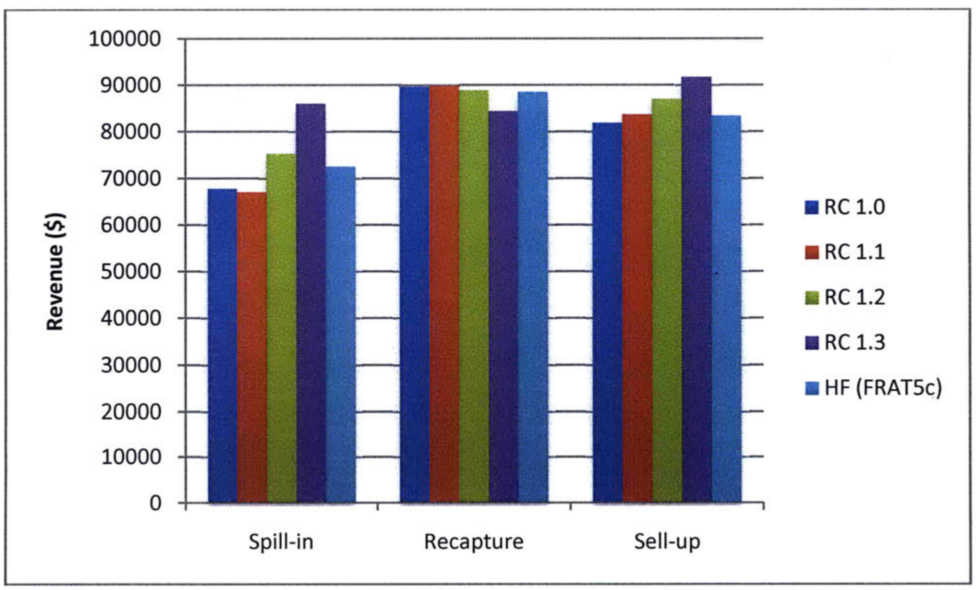


Figure 94: Spill-in, Recapture, and Sell-up Revenue for RC with FM in Network D6 Semi-restricted

Comparing the rest of the revenue types, it is evident that while the FM caused spill-in and sell-up revenue to increase well beyond the level of Hybrid Forecasting, the recapture revenue, in addition to the first-choice revenue, eventually decrease. All of these factors combined produce a best revenue at FM = 1.2 for Rational Choice Forecasting, still 0.25 percent below that of Hybrid Forecasting.

6.5.2. RC with Fare Adjustment

Using the sell-up probabilities derived from the price-oriented passenger types in RC Forecasting (Section 4.4), implementing fare adjustment (FA) is a simple process. Recall that the two options for FA in regards to RC Forecasting are to either keep the sell-up estimates on the path level, or to aggregate them on the market level (denoted by "p" or "m" in the analysis). Also, because the sell-up probabilities are available (and thus an elasticity constant), a FRAT5 may be obtained in order to compare sell-up estimates against the Hybrid Forecasting FRAT5s.

Network T1

In Network T1, fare adjustment performs quite well for AL1 using EMSRb with standard RC Forecasting, improving on the previous RC methods without FA, and nearly reaching the revenue level of Hybrid Forecasting.

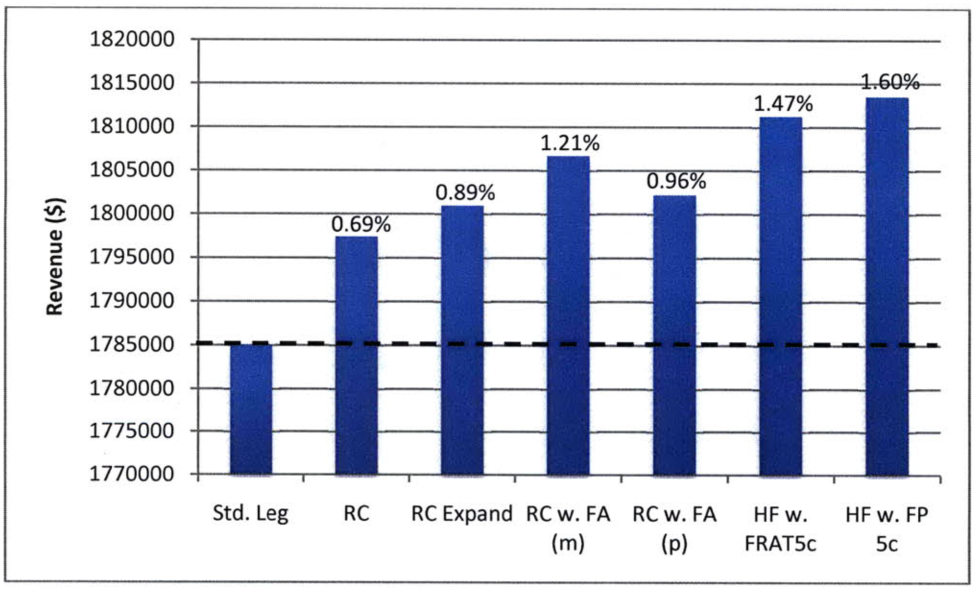


Figure 95: RC with FA Revenue in Network T1

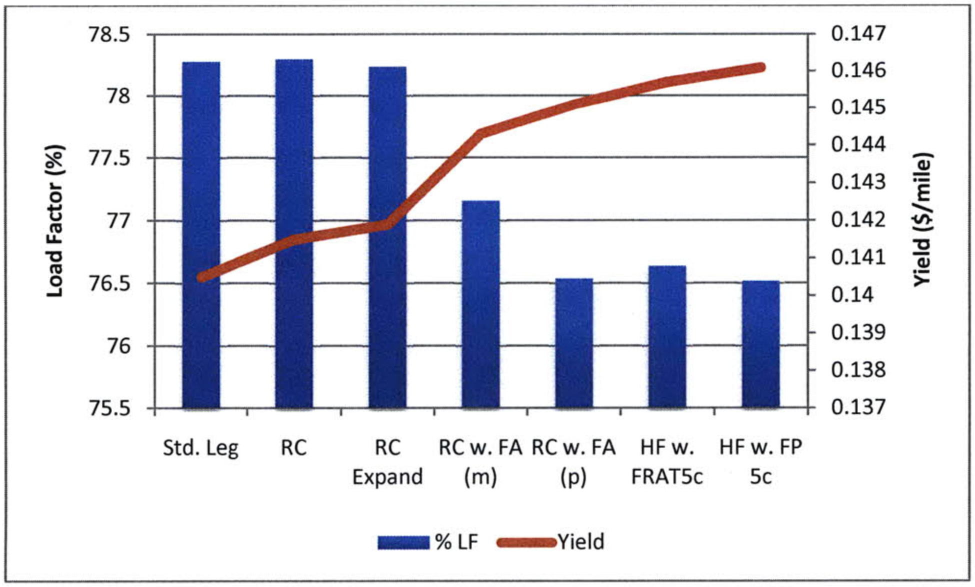


Figure 96: RC with FA Load Factors and Yields in Network T1

For Rational Choice Forecasting, using FA decreases load factors and increases yields, almost to the levels of Hybrid Forecasting. In addition, the market level of sell-up estimate aggregation for FA, as opposed to the path level, creates a lower yield but higher load factor, resulting in an overall gain in revenue. While RC with FA (market) performs slightly worse than HF with FP (5 clusters), it is largely due to the aggressiveness of the forecast.

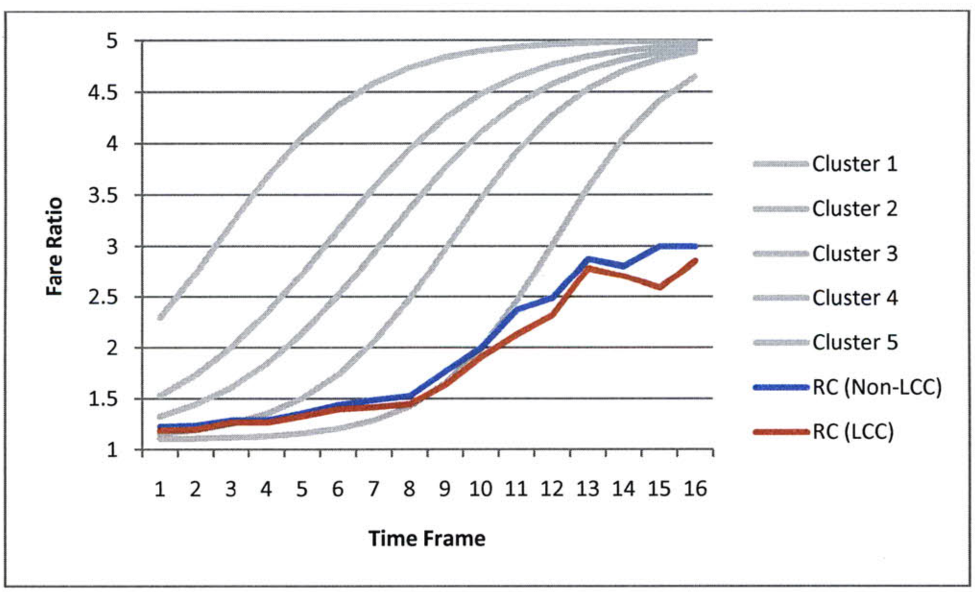


Figure 97: RC with FA (m) FRAT5 Curves in Network T1

The sell-up estimates for RC with FA are clearly not as aggressive as the HF method that uses FP sell-up estimation with a logistic-fit. Because of this, the fare class closures are more aggressive for HF when compared to RC Forecasting, especially in early time frames.

Network T4

Network T4 provides better insights into how RC Forecasting with Fare Adjustment performs in comparison to Hybrid Forecasting in a network with more price-oriented demand, thus making sell-up estimates even more important.

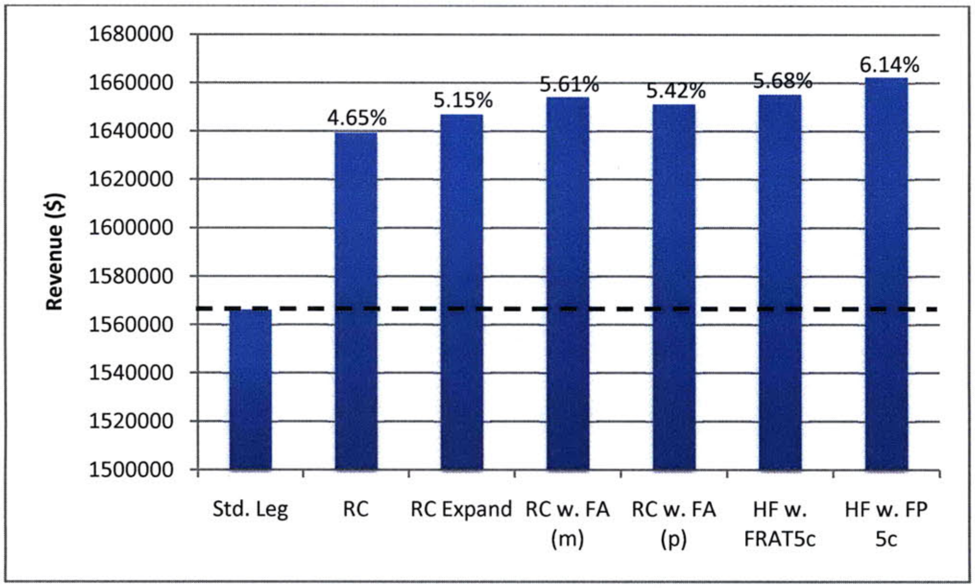


Figure 98: RC with FA Revenue in Network T4

Like Network T1, fare adjustment increases revenue for RC Forecasting, with market aggregation performing better than path. However, RC Forecasting with FA still falls short of the best Hybrid Forecasting method (FP 5 cluster), due to the aggressiveness of Hybrid Forecasting.

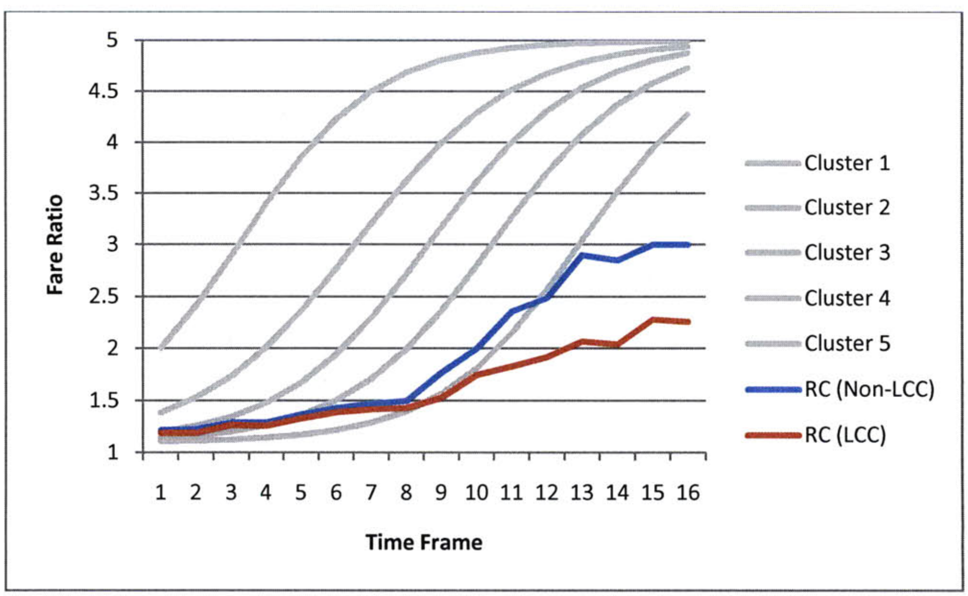


Figure 99: RC with FA (m) FRAT5 Curves in Network T4

The five cluster logistic-fit for Hybrid Forecasting has much more aggressive estimates of sell-up as compared to RC Forecasting with FA. Like Network T1, this has a direct effect on fare class closures and the overall aggressiveness of the forecast.

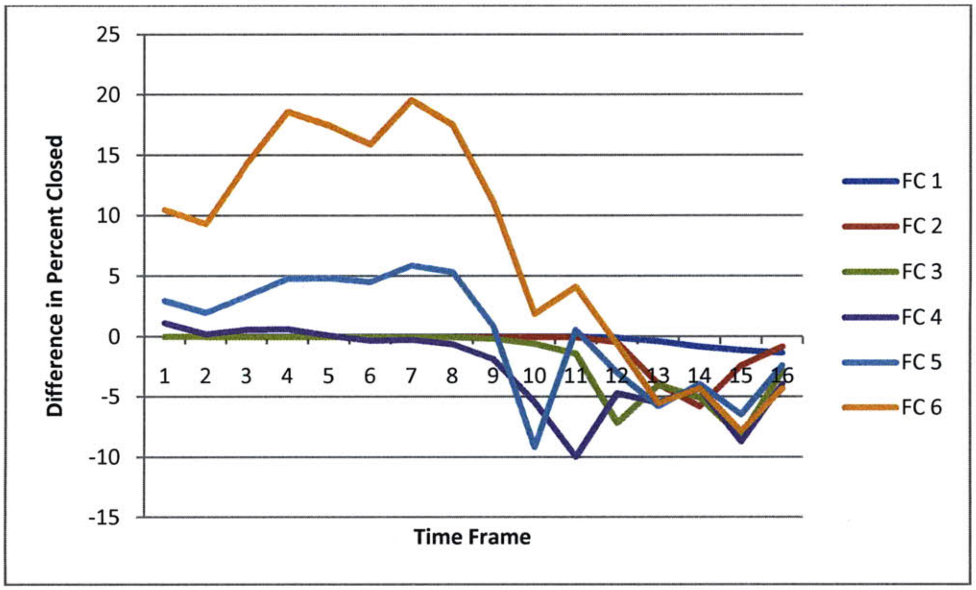


Figure 100: Difference in Fare Class Closures over Time: HF with FP (5c) minus RC w. FA (m) in Network T4

Comparing Hybrid Forecasting (FP with 5 clusters) to RC Forecasting with FA (market), the HF method closes FC 5 and FC 6 much earlier than RC with FA, as shown in Figure 100. In the last six time frames, however, it is evident that RC with FA is more closed in all fare classes, likely due to the abundance of lower class passengers. In contrast, this is where HF benefits from having more open seats in higher classes at later time frames, resulting in higher yields, lower load factors, and overall higher revenues.

The benefit of fare adjustment for RC Forecasting is still sizeable compared to not using it. It still increases the aggressiveness of fare class closures, as shown in Figure 101.

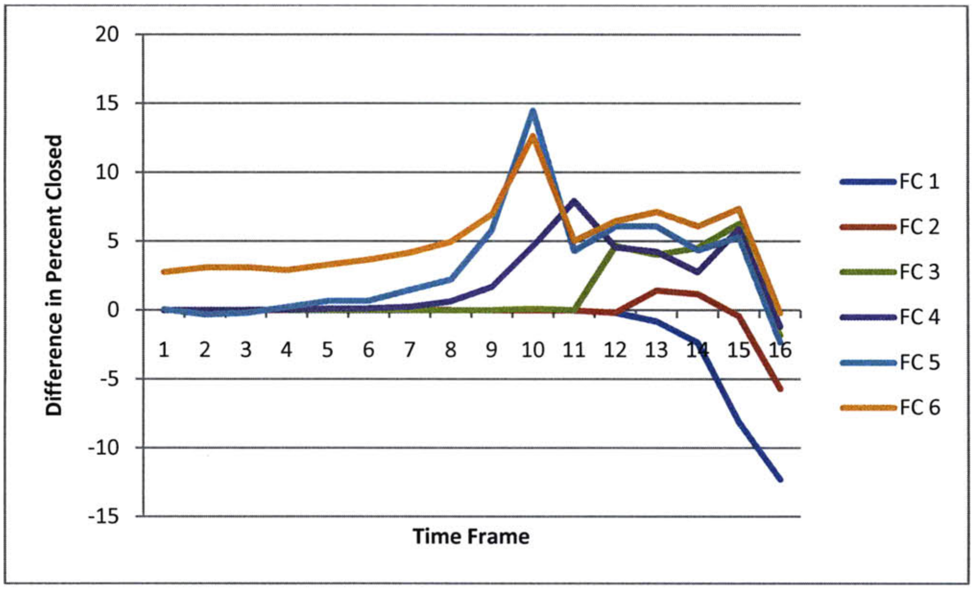


Figure 101: Difference in Fare Class Closures over Time: RC with FA (m) minus RC in Network T4

Throughout the entire booking process, fare adjustment causes FC 5 and FC 6 to be more closed than standard RC without FA. In addition, fare adjustment allows more of FC 1 and FC 2 to be open in the last four time frames, resulting in greater overall revenue. While not as aggressive as Hybrid Forecasting, fare adjustment for RC Forecasting does provide an increase in revenue.

6.6 Chapter Summary

Throughout Chapter 6, the Passenger Origin-Destination Simulator provided a tool for measuring the benefit of new forecasting, sell-up estimation, and clustering methods. PODS is capable of creating environments that simulate real-world scenarios in regards to networks, fare structures, competition, and passenger generation. In relation to this thesis, PODS enabled a focus on networks that use less-restricted fare environments, as seen in Network D6 Unrestricted and more complex Network T4.

Sell-up estimation is crucial to preventing the spiral down of forecasts and revenues in a less-restricted fare environment. Estimating the FRAT5 curve proves to be an important, but difficult task. While input FRAT5s perform well in the simulations, they lack real-world usability because they are arbitrarily developed and not based on historical booking data. (The best performing input FRAT5 is the piecewise step function FRAT5.) When developing a data-based method to estimate sell-up, several factors are important to consider.

First, what should be done with missing FRAT5 estimates for time frames? Using a logistic or regression-based fitter appears to solve the problem, with logistic-fit generally outperforming all regression-based methods. However, this comes with some drawbacks. As shown by the performance comparison between piecewise input FRAT5s and the logistic-fit method, the logistic-fit FRAT5s may over-estimate sell-up in middle time frames. Also, when using the logistic-fitter or regression-based fitter, sell-up may either be estimated by the Forecast Prediction (FP) or Direct Observation (DO) methods. Comparing the results between the two methods, it appears that FP consistently, but only slightly, outperforms DO.

Next, to combat the sell-up aggregation question about how general or how specific sell-up estimates should be for a given market, clustering provides a viable answer. Clustering creates a middle ground between over-specific sell-up estimation, where each market receives its own FRAT5 curve, and too broad of sell-up estimation, where every market in the entire system receives the same FRAT5 curve. Using clustering increases the number of sell-up observations used to develop a single curve, classifying each of the markets into one of k clusters. While the proper number of clusters to use is not entirely clear, some guidance exists from the use of the gap statistic, the actual market distribution over a given k number of clusters, and the total within sum-of-squares. Using these methods, the proper number of clusters appears to be within two and seven clusters, depending on the size of the network. Using the 2-parameter logistic-fit, with FP sell-up estimation generally provides the largest increase in revenue over the previous market or system-based aggregation level. Compared to the input FRAT5c, clustering creates revenue gains of 0.75% and 0.44% for AL1 in Network D6 Unrestricted and Network T4, respectively. While the clustering method appears to work well, some may argue that the number of markets actually generating sell-up estimates and being clustered is too low, leaving numerous markets with the input FRAT5c. This is primarily due to the requirements of the logistic-fit, and increasing the demand multiplier to create more clusterable markets actually slightly decreases revenue and causes over-aggressive FRAT5 curves.

Combining clustering with the success of the piecewise input FRAT5 does not result in a high-performing method. Much of the failure is due to the inability to deal with outliers and lower FRAT5 values. Unlike the logistic-fitter, using a mean value over a given time period does not allow a FRAT5 curve to reach its maximum value, thus creating low estimates of sell-up, especially in later time frames when they matter most.

Finally, the introduction of Rational Choice Forecasting does provide a workable, simple, data-based forecasting method. While its revenue levels are not as high as those of Hybrid Forecasting, it does not need an external estimate of sell-up, as sell-up is accounted for by the inclusion of numerous passenger type categories. This method also removes the assumptions that passengers arrive in an inverse willingness-to-pay order and that all bookings in the lowest open class are price-oriented. While RC Forecasting is very sensible, it creates less aggressive forecasts than Hybrid Forecasting that result in more spiral down and thus slightly lower revenue. Even when using a forecast multiplier, RC Forecasting does not exceed the total revenue of Hybrid Forecasting. Using RC Forecasting with Fare Adjustment, however, increases revenue in all networks tested by creating more aggressive forecasts. Much of the success of Hybrid Forecasting stems from its ability to close lower classes early in the booking process. Because of this, more capacity exists in later time frames, allowing for not only more sell-up, but more recapture and spill-in passengers from other airlines.

CHAPTER 7

INSIGHTS INTO CLUSTERING AND ITS APPLICATION

7.1 Regression Analysis

While the previous chapters introduce the sell-up clustering method and describe its performance in various scenarios, this chapter aims to delve further into what market characteristics drive a market into belonging to a particular cluster. Being able to understand why a market has a certain estimate of sell-up, beyond that of just using its historical booking observations, is crucial and may provide more insight into market sell-up predictability. Because the x1 and x2 logistic-fit parameters are the primary indicators of the FRAT5 curve in clustering, this chapter seeks to predict, or explain, a market's x1 and x2 parameters through the use of independent market variables with various forms of regression models.

7.1.1. Setup

The markets used for this experiment all come from Network T4, where non-LCC markets have a "more restricted" fare structure and LCC markets are all completely unrestricted, making sell-up observations more abundant for all markets. However, recall that when run at a demand multiplier of 1.0, which produces normal load factors of about 79 percent, only 247 of the 572 markets received logistic-fit parameters (in the FP case), leaving the majority of markets to use the input FRAT5c. For this experiment, it is essential to have as many markets as possible with logistic-fit parameters, and as shown in Section 6.4.2, increasing the demand multiplier fixes this problem at the expense of slightly overestimating sell-up. A demand multiplier of 1.50 is used to obtain the most markets with non-zero sell-up parameters (only 43 still do not have an estimate of sell-up) via FP sell-up estimation, creating a data set of 529 usable markets.

Of these 529 markets in the complete data set, a randomly selected 75 percent of them (397 markets) are set aside to be the training set, while the remaining 25 percent (132 markets) serve as the test set. The regression model's performance, while developed from the training set, is measured by the average absolute residuals of the model's application to the test data set (test error), and serves as the most important indicator of how well the model performs on other data. In addition, R-squared values give an insight on how well the model explains the variance in the training data, with the p-values serving as primary indicators of the significance of a particular independent variable.

This experiment is divided into two sub-experiments, one trying to predict the value of x1, and the other to explain the value of x2. For each experiment, the following independent variables exist for every market (observation).

Whether or not the Market is an LCC (binary)
Average Yield
Percentage of Business Passengers
Fare Ratio
Constraining Load Factor
Whether or not the Airline has a Route Advantage (binary)

For a market to be classified as an LCC, it must be served by AL3 in Network T4. The average yield and percent of business passengers are determined over all flights per day that serve the market. While the percent business demand is known exactly in PODS by looking at the passenger attributes, it is not unreasonable that an airline in real life would be able to estimate what portion of passengers are classified as business for a particular market.

The fare ratio is defined as the highest fare divided by the lowest fare for the market, and simply shows the maximum range of sell-up that one may observe in a market. Keep in mind that all willingness-to-pay attributes are distributed randomly throughout the passenger population in PODS. Therefore, a market with a low fare ratio does not allow those with a higher maximum willingness-to-pay from actually paying up to that level, thus constraining what may be observed in the market. Additionally, if a market has a high fare ratio, more observations of higher sell-up should exist, thus theoretically resulting in a more aggressive FRAT5 curve.

The constraining load factor is another interesting market attribute that may be able influence an estimate of sell-up. For a higher average load factor, the sell-up estimate should be higher due to more closures of low fare classes, meaning there is more sell-up. Therefore, those markets that have higher average load factors should hypothetically have more aggressive FRAT5 curves. To determine the constraining load factor in Network T, keep in mind that for every market, there are up to three paths of two connecting legs throughout the day (west spoke to hub, hub to east spoke). The constraining load factor is defined as the maximum of the two sets of averages (west spoke to hub average, hub to east spoke average).

Last, a route advantage for AL1 in a particular market is given to all markets that begin or end with AL1's hub (Minneapolis-St. Paul). If a competitor airline wishes to carry passengers to, or out of AL1's hub, it must first fly through its own hub. Therefore, because AL1 offers non-stop service for the particular market involving MSP, and the competition may only offer connecting service, the market is classified as a route advantage.

7.1.2. Data Relationships

To gain a better insight into each of the regression experiments, it is essential to understand the meaning of the x1 and x2 parameters, as well as their relationship with the independent variables. The logistic-fit parameters define different aspects of the shape of the FRAT5 curve.

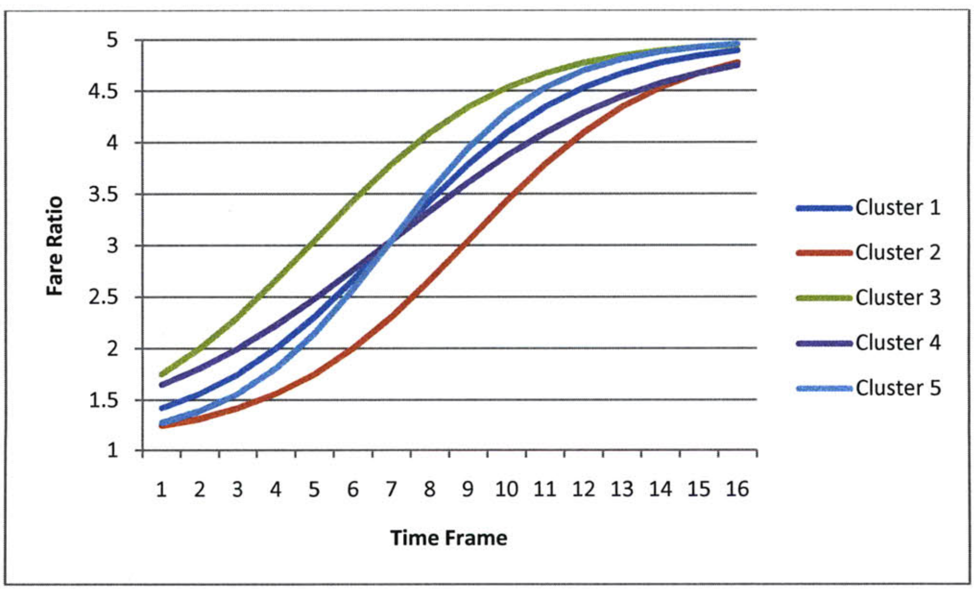


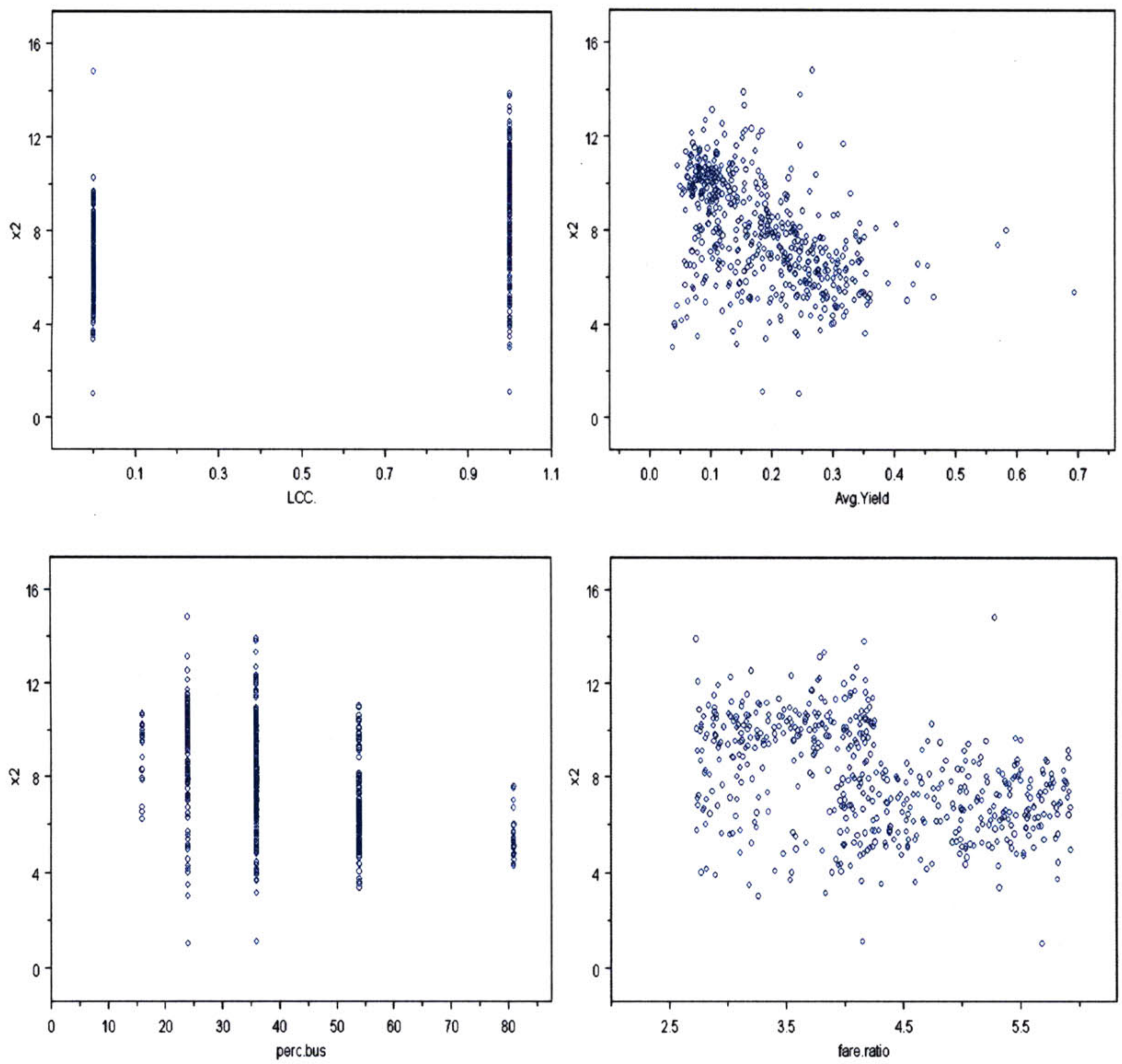
Figure 102: FRAT5 Example Curves (FP 5 Clusters)

Cluster	*1*	*2*	*3*	*4*	*5*
x1 value	0.4	0.4	0.4	0.3	0.5
x2 value	7.0	9.0	5.0	7.0	7.0

Table 23: Logistic-fit Parameters for FRAT5 Example Curves

First, holding x1 constant and changing the value of x2, it is evident that x2 is responsible for the curve's height and aggressiveness. Also, there exists an inverse relationship—as x2 decreases, the height of the curve increases. Next, holding the x2 value constant, it is easy to see that the x1 value is responsible for the steepness of the curve. A higher value of x1 implies a steeper FRAT5 curve, all else equal. Because x2 is more responsible for the overall aggressiveness and maximum point of a FRAT5 curve, it is the first and main area of focus for the regression analysis, and likely provides stronger relationships with the independent variables associated with each market.

The easiest method to determine if a relationship exists between a dependent and independent variable in a regression is to simply look at the scatter plots. These relationships also give an initial insight about the signs of the variable coefficients in the regression.



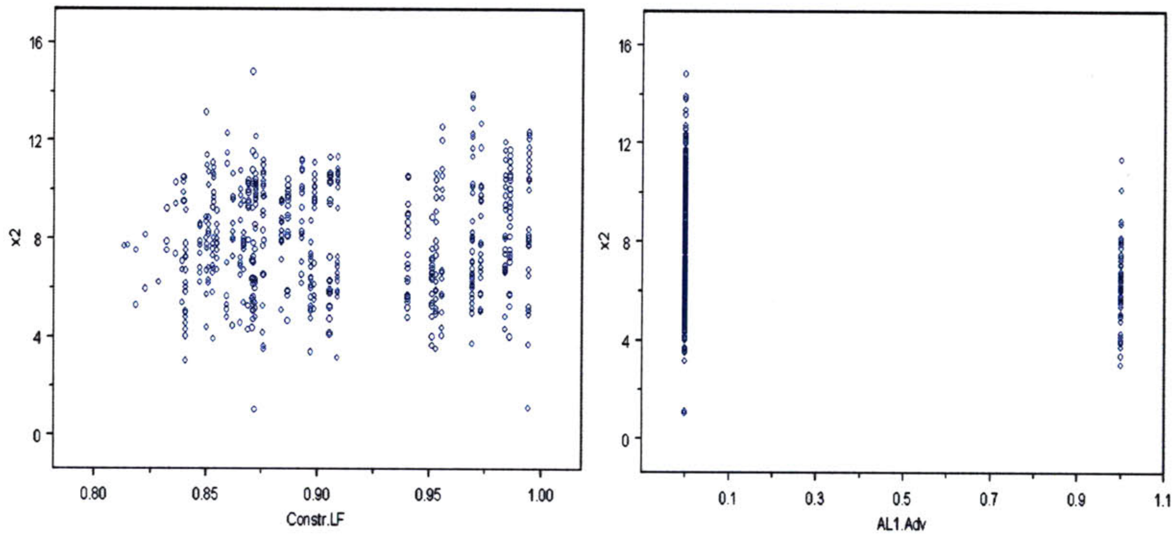
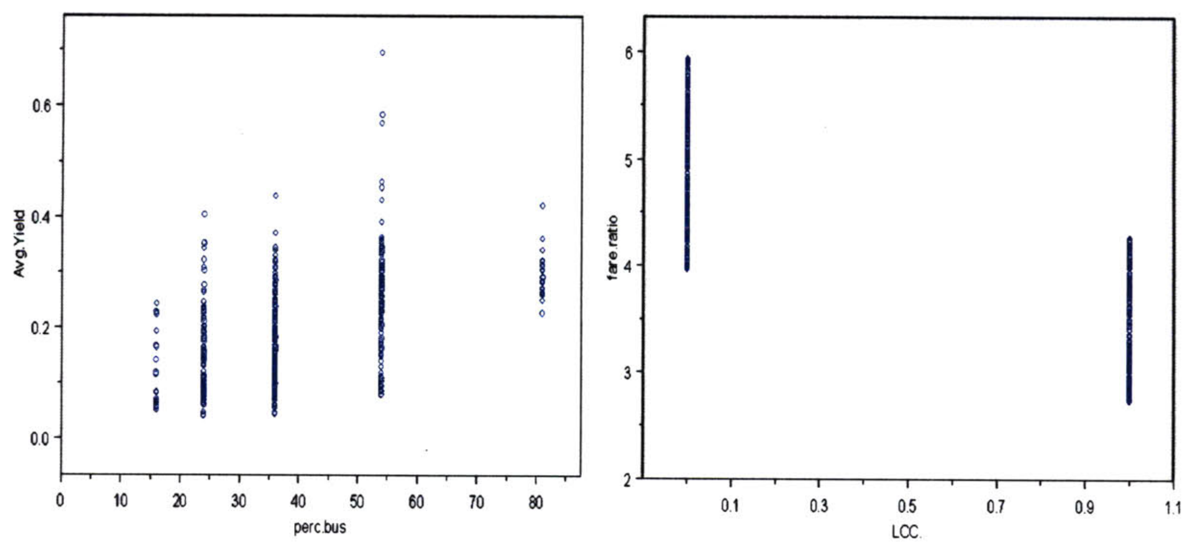


Figure 103: x2 vs. Independent Variable Scatter Plots

Using the scatter plots, it appears that the only variable with a positive correlation with x2 is the LCC variable. This reasonably suggests that LCCs have a tendency to have lower FRAT5 curves. All other relationships appear to either have a negative correlation with x2 (meaning an increase in curve height) or a lack of a significant correlation (such constraining load factor).

In addition to plotting the dependent response variables versus the independent variables, it is also important to plot the independent variables versus themselves in search for potential regression problems that violate some assumptions of OLS.



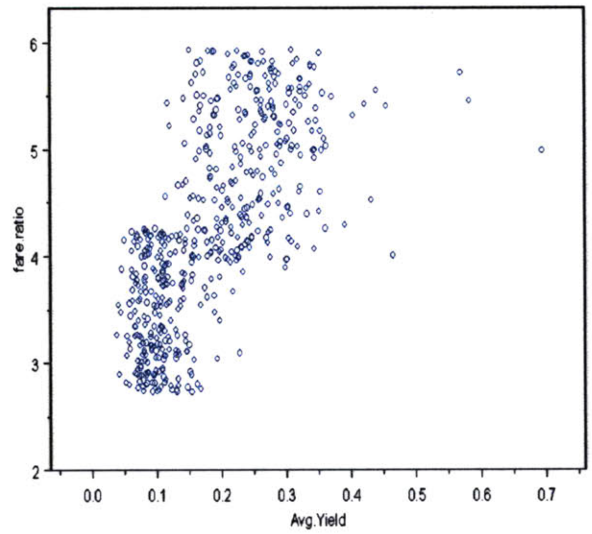


Figure 104: Independent Variable Relationships

While it appears that several of these independent variables may be strongly related to each other (especially fare ratio and average yield), bringing to light the potential of multicollinearity problems, these are addressed with various regression models in Section 7.1.3.

7.1.3. Regression Models and Results

The following analysis includes several regression models, starting with a basic linear regression. Before beginning, we note the six assumptions of the classical regression model. First, the shape of the relationship between the response variable, x2, and the independent variables must be linear, which is shown by the simple regression equation. Next, there must be no multicollinearity, meaning that the columns of the matrix including all independent variables must be independent. This already poses a problem as shown by the scatter plots indicating some relationships between independent variables. Third, the expected value of the error term must equal zero. Fourth, the variance of the error term must be constant. In other words, when plotting the residuals versus the fitted values of the dependent variable, there must be a constant variance about the x-axis. Next, the error terms must be independent. If these first five assumptions are met, then this indicates that this is the best linear unbiased estimator. Finally, requiring the error term to be normally distributed ensures that hypothesis testing is possible, allowing the use of the t-statistic and p-values to determine the significance of a particular coefficient. Any violation of these assumptions creates potential problems and inaccurate models. To counter these issues as they arise, several different regression models are used in the following analysis.

Full OLS Model

The first regression method used in this study is the simple OLS regression of x2 on all independent variables. This method provides a good starting point and will give insight into what OLS assumptions are violated, creating the need to correct them and/or use other regression models.

$$x2 = \beta_0 + \beta_1 \cdot LCC + \beta_2 \cdot Avg\ Yield + \beta_3 \cdot Percent\ Business + \beta_4 \cdot Fare\ Ratio + \beta_5 \cdot Constraining\ LF + \beta_6 \cdot AL1\ Advantage + \varepsilon$$

	Coefficient	Std. Error	t value	Pr(>|t|)
(Intercept)	10.7678	1.9607	5.4919	0.0000
LCC	1.9692	0.3773	5.2190	0.0000
Avg. Yield	-3.9736	1.8244	-2.1780	0.0300
Percent Bus.	-0.0331	0.0070	-4.7211	0.0000
Fare Ratio	0.1886	0.1738	1.0851	0.2786
Constr. LF	-2.6390	2.0312	-1.2992	0.1946
AL1 Adv.	-2.4712	0.3313	-7.4583	0.0000

R-squared	0.4165

Table 24: Linear Regression (Full Model) Output

Looking at the results, it is evident that not all of the coefficients make sense. In addition, not all p-values show that the coefficients are significant. For example, a positive sign on the fare ratio coefficient does not necessarily make sense (this means a higher fare ratio implies a lower FRAT5 curve), which is accounted for by a large standard error, small t-value and its p-value being insignificant at the 0.05 level of confidence. Besides the fare ratio variable, as well as the constraining load factor variable lacking in significance, the rest of the signs of coefficients coincide with the initial thoughts about their relationship with the x2 parameter.

Also, note that the R-squared value is only 0.4165, even with every variable in the model (adding variables only cause the R-squared value to increase, regardless of how "good" they are). This means that the regression only explains about 42 percent of the variation in the model, whereas a perfect fit achieves an R-squared of 1.00.

Stepwise Regression

To account for the insignificant variables in the full OLS model, stepwise regression is often used to pick the "best subset" of variables to include in the model, while keeping all p-values significant. The stepwise algorithm works in two directions—moving forwards, it begins with zero variables in the model and adds the best variable one at a time while maintaining significance, or, by moving backwards, begins with a model including all variables and removes the least significant variable until all included variables' coefficients are significant. The algorithm used in this analysis produced the same results for both directions.

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	9.4655	0.4319	21.9156	0.0000
LCC	1.5327	0.2757	5.5586	0.0000
Avg. Yield	-4.4626	1.6144	-2.7642	0.0060
Percent Bus.	-0.0326	0.0069	-4.7536	0.0000
AL1 Adv.	-2.5174	0.3254	-7.7352	0.0000

R-squared	0.4118

Table 25: Stepwise Regression Output

Despite removing two variables from the model, the R-squared value only decreased from 0.4165 to 0.4118, meaning that the two variables (fare ratio and constraining load factor) provide little insight on the prediction of x2. The stepwise regression model appears to make sense, implying that LCCs will have lower FRAT5 curves, and markets with higher yield, more business passengers, and where AL1 has an advantage, will all have higher FRAT5 curves, all else equal. While it is clear that multicollinearity exists to a certain extent in this model due to the inclusion of both "average yield" and "percent business" variables, it is also necessary to see if heteroscedasticity is a problem.

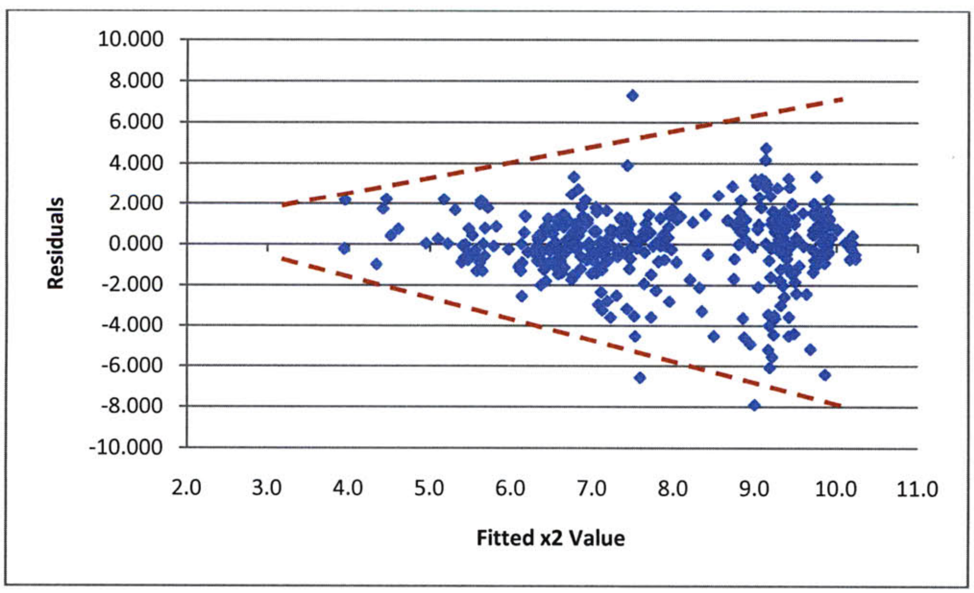


Figure 105: OLS (Best Subset) Residuals vs. Fitted Values

If the residuals were homoscedastic, they would show constant variance over the range of the fitted values. However, there is clear evidence in Figure 105 that this is not the case, indicating an increase in variance as the fitted value of x2 increases. In order to account for this, weighted least squares (WLS) and robust regression are two methods that are tested later in this section. First, ridge regression and partial least squares (PLS) provide two methods to account for the potential problems of multicollinearity.

Ridge Regression

If multicollinearity exists and independent variable terms are correlated, the least squares coefficient estimates become highly sensitive to random errors in the observed values of x2, the dependent variable, which will produce a large variance. For example, a very large positive coefficient on one variable may be cancelled by a very negative coefficient on a related variable. Therefore, the goal of ridge regression is to impose a size constraint on the coefficients, as indicated by the ridge parameter k, which will reduce the variance of estimates and often result in a smaller MSE.

$$\hat{\beta}^{ridge} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2$$

$$\text{subject to} \sum_{j=1}^{p} \beta_j^2 \leq 1/k$$

		Ridge		
	OLS	k=0.05	k=5.0	k=50
(Intercept)	10.7678	10.7686	10.8397	11.1059
LCC	1.9692	1.9682	1.8736	1.3909
Avg. Yield	-3.9736	-3.9742	-4.0297	-4.0819
Percent Bus.	-0.0331	-0.0331	-0.0330	-0.0315
Fare Ratio	0.1886	0.1882	0.1511	-0.0376
Constr. LF	-2.6390	-2.6374	-2.4888	-1.7181
AL1 Adv.	-2.4712	-2.4708	-2.4364	-2.1531

Table 26: Ridge Regression Output

Using ridge regression to shrink the coefficients appears to only have a noticeable effect for large values of k. It is interesting to note that the "fare ratio" and "constraining load factor" variables shrink the most, which were also the two that were cut in the stepwise regression.

Partial Least Squares (PLS)

Like ridge regression, partial least squares is another coefficient shrinking technique. PLS creates components, which are linear combinations of the original independent variables that have a large covariance with x2, and then regresses x2 on the newly created components. This is an effective technique because it uses information about the variances of the dependent and independent variables, as well as the correlations between them.

	Component				
	1	2	3	4	5
LCC	-3.875	-8.115	-1.490	-3.970	-0.599
Avg. Yield	0.967	1.131	-0.180	0.567	-0.889
Percent Bus.	303.069	-1.314	0.024	-0.018	-0.001
Fare Ratio	5.570	17.413	-2.162	-1.971	-0.241
Constr. LF	-0.135	-0.251	-0.006	-0.090	-0.749
AL1 Adv.	-0.163	-0.472	4.841	-2.991	-0.326

Table 27: PLS Component Relationship to Independent Variables

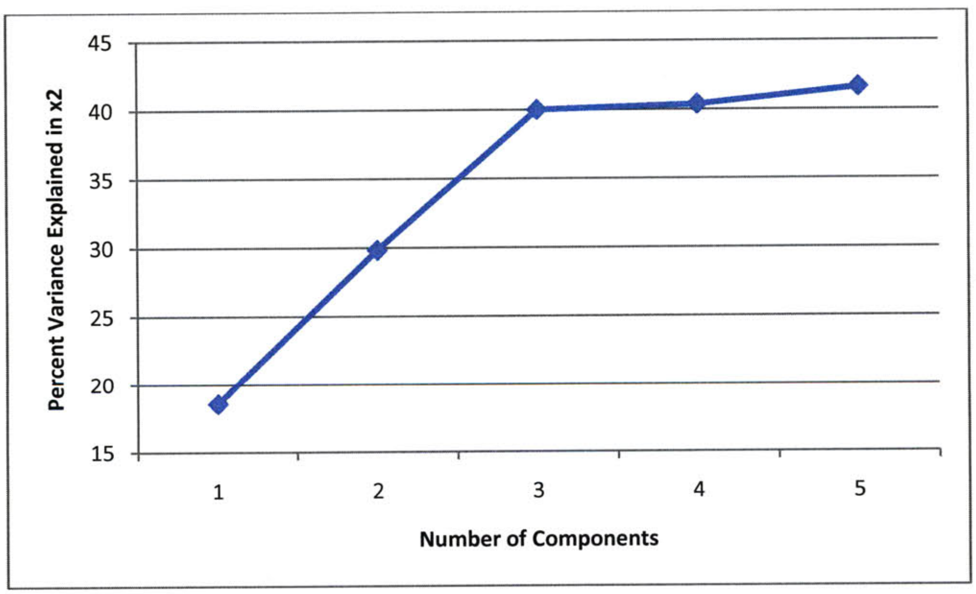


Figure 106: Variance Explained by Components in PLS Regression

Table 27 shows the linear combination of components that approximate the original independent variables for PLS in the five component case. However, the addition of components does reach a saturation point in regards to the amount of variance explained in the model. Figure 106 shows the majority of the variance is explained by three components, with minimal benefit gained beyond that. In the last step of

PLS, based on the regression of x2 on the components, coefficients may be estimated for each of the original independent variables, as reported in Table 28 for both the three and five component cases.

	Coefficients	
	3 Component	5 Component
(Intercept)	9.2199	10.8799
LCC	1.8660	1.9828
Avg. Yield	-0.2367	-3.8707
Percent Bus.	-0.0406	-0.0333
Fare Ratio	-0.0665	0.1864
Constr. LF	-0.1124	-2.7743
AL1 Adv.	-2.4749	-2.4643
R-squared	0.4002	0.4165

Table 28: PLS Model Output

The PLS regression creates similar results to previous regression models, all indicating the comparable relationships between x2 and the independent variables (except the fare ratio coefficient in the 5 component case).

Weighted Least Squares (WLS)

In order to correct for heteroscedasticity, WLS assigns weights to determine the contribution of each observation to the final parameter estimates. The weights are inversely proportional to the variance at each level of the explanatory variable. A higher variance implies less weight, resulting in a residual versus fitted values plot with a more random distribution of residuals, showing an equal variance over all fitted values.

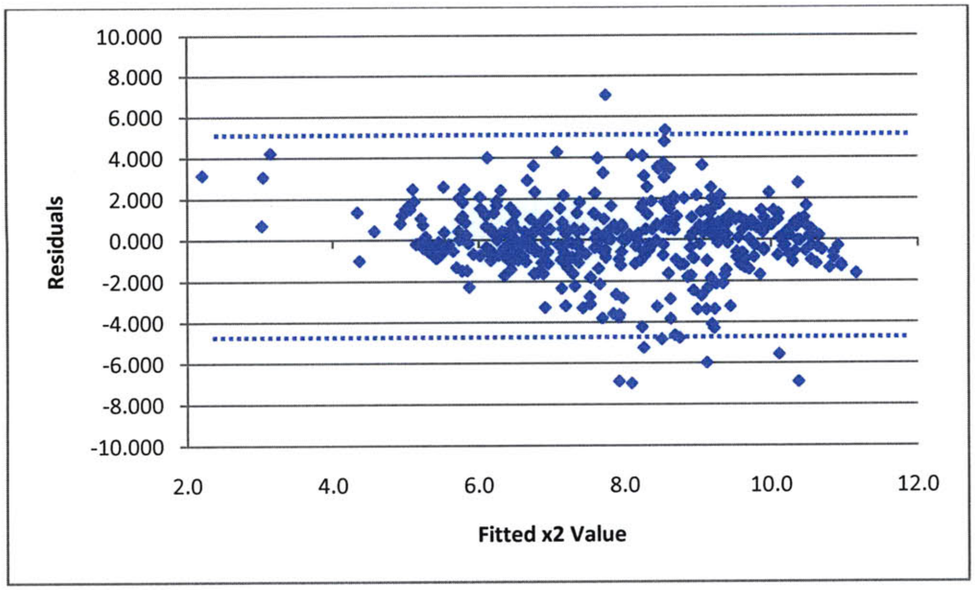


Figure 107: WLS Residuals vs. Fitted Values

WLS Full

	Coefficients	
	OLS	**Robust**
(Intercept)	10.7678	17.95
LCC	1.9692	0.9963
Avg. Yield	-3.9736	-8.4778
Percent Bus.	-0.0331	-0.0369
Fare Ratio	0.1886	-0.0454
Constr. LF	-2.6390	-7.8389
AL1 Adv.	-2.4712	-2.6847

WLS Stepwise Best Subset

	Coefficients	
	OLS subset	**Robust subset**
(Intercept)	9.4655	17.7209
LCC	1.5327	1.0557
Avg. Yield	-4.4626	-8.5916
Percent Bus.	-0.0326	-0.0366
Fare Ratio		
Constr. LF		-7.8148
AL1 Adv.	-2.5174	-2.7343

Table 29: WLS Regression Output

The WLS regression also produces sensible coefficients for the variables included in the model. In the full variable case, in comparison to the first OLS regression, the WLS coefficients all appear to be in line with the original hypothesis, correcting the coefficient for fare ratio from a positive to a negative value. Applying stepwise regression to WLS after the reweighting process creates a new subset of variables, similar to the OLS best subset, but also including "constraining load factor" as another variable.

Robust Regression

Robust regression works by a process called iteratively reweighted least squares. In the first iteration, all data receive equal weight, and in subsequent iterations, those data points that are farther from the previous model, are given less weight. For example, if an outlier is skewing the results, that observation would receive less weight in the next iteration. This process repeats until there is convergence within a specified limit. This is essentially just an iterative version of WLS, where the weights are not necessarily all inversely proportional to the variance of the variable.

Robust Full

	Coefficients	
	OLS	Robust
(Intercept)	10.7678	10.3085
LCC	1.9692	2.6526
Avg. Yield	-3.9736	-2.4093
Percent Bus.	-0.0331	-0.0372
Fare Ratio	0.1886	0.2328
Constr. LF	-2.6390	-2.6092
AL1 Adv.	-2.4712	-2.4442

Robust Stepwise Best Subset

	Coefficients	
	OLS subset	Robust subset
(Intercept)	9.4655	9.2068
LCC	1.5327	2.2497
Avg. Yield	-4.4626	-2.6268
Percent Bus.	-0.0326	-0.0368
Fare Ratio		
Constr. LF		
AL1 Adv.	-2.5174	-2.5675

Table 30: Robust Regression Output

The output of robust regression shows interesting results. For the full model, compared to the OLS estimates, more weight is particularly given to whether or not the market is an LCC and the market fare ratio, which is contrary to previous models that either shrunk the "fare ratio" variable or removed it all together. In both the full variable model and the best subset model, less weight was allocated to the average yield, which has been a primary component in every regression model shown in the analysis.

7.1.4. Regression Summary

x2 Regression

A summary of the coefficients for all regression models used to predict x2 are shown below in Table 31. Each model tested serves a purpose—stepwise regression creates a best subset of variables, removing some element of multicollinearity. Ridge regression and PLS also attempt to remove multicollinearity by shrinking coefficients of those variables that would otherwise result in higher variance of the estimates. Both WLS and Robust Regression attempt to reweight observations that are more beneficial to creating homoscedastic residuals and normal error terms.

There are several common relationships among the independent variables throughout the different methods tested that are worthy of note. It is apparent that by a market being an LCC market with an unrestricted fare structure, it is predicted to have a lower FRAT5 curve. However, several factors influence a market to have a higher FRAT5 curve—higher average yield, more business passengers, a higher load factor, or a route advantage. When evaluating the significance of the variables, several regression techniques remove "fare ratio" and "constraining load factor" from the model. Statistical insignificance for a given variable is often shown by its coefficients changing signs in various models. This is definitely the case for "fare ratio," and as a result, the value of those models that include fare ratio could be discounted.

	Lin Reg. (all)	Lin Reg. (subset)	Ridge k=0.05	Ridge k=5.0	Ridge k=50	PLS (3 comp)	PLS (5 comp)	Robust Reg. (all)	Robust Reg. (subset)	WLS (all)	WLS (subset)
(Intercept)	10.768	9.466	10.769	10.840	11.106	9.220	10.880	10.309	9.207	17.950	17.721
LCC	1.969	1.533	1.968	1.874	1.391	1.866	1.983	2.653	2.250	0.996	1.056
Avg. Yield	-3.974	-4.463	-3.974	-4.030	-4.082	-0.237	-3.871	-2.409	-2.627	-8.478	-8.592
Percent Bus.	-0.033	-0.033	-0.033	-0.033	-0.032	-0.041	-0.033	-0.037	-0.037	-0.037	-0.037
Fare Ratio	0.189		0.188	0.151	-0.038	-0.067	0.186	0.233		-0.045	
Constr. LF	-2.639		-2.637	-2.489	-1.718	-0.112	-2.774	-2.609		-7.839	-7.815
AL1 Adv.	-2.471	-2.517	-2.471	-2.436	-2.153	-2.475	-2.464	-2.444	-2.568	-2.685	-2.734

Table 31: x2 Regression Models Summary

The effectiveness of the regression models hinges on its comparison to a baseline model when applied to the test data set. For these cases, the baseline is simply choosing the average value for x2 for every predicted value in the test set. The measure of the test error is the average absolute residual, or the absolute difference between the predicted value of x2 and the actual value of x2 in the test set.

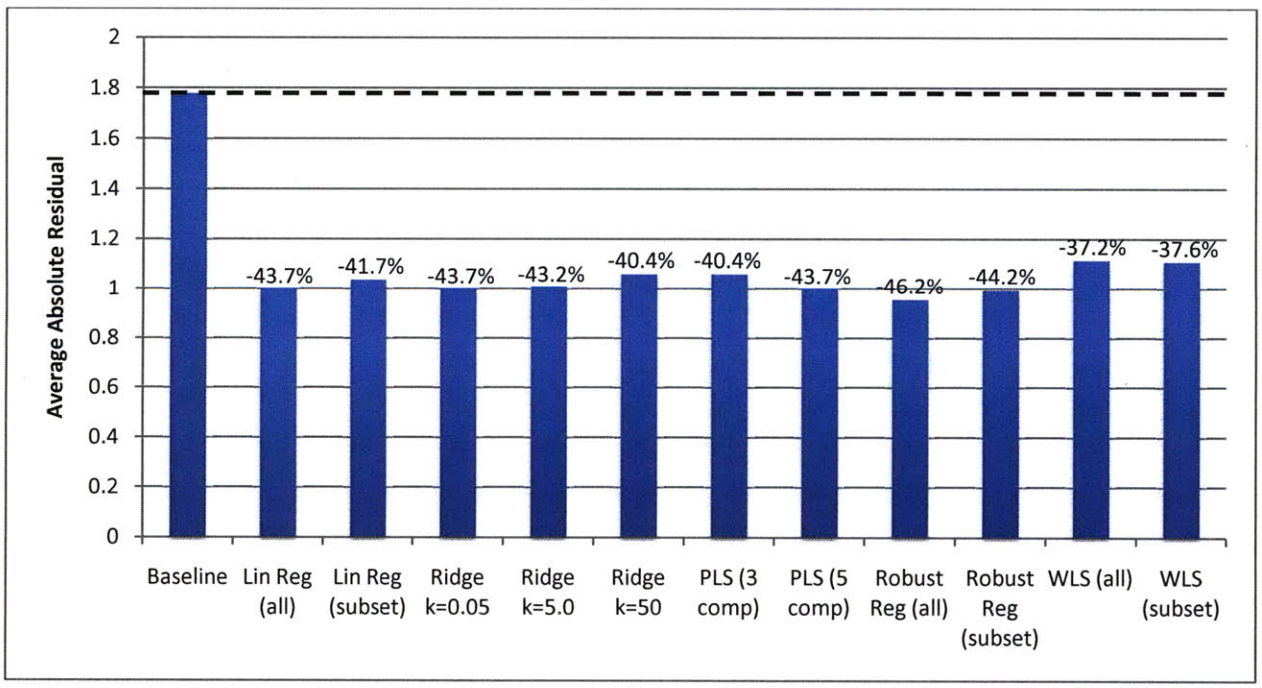


Figure 108: x2 Test Error per Regression Model

The best performing regression model is robust regression, creating a 46.2 percent decrease in test error compared to the baseline case for the full model (all variables). However, as discussed previously, this model gives the variable "fare ratio" a positive coefficient, going against what was hypothesized. The remaining models all lay within a 37 to 44 percent range of improvement over the baseline, with robust regression (best subset) at 44.2 percent.

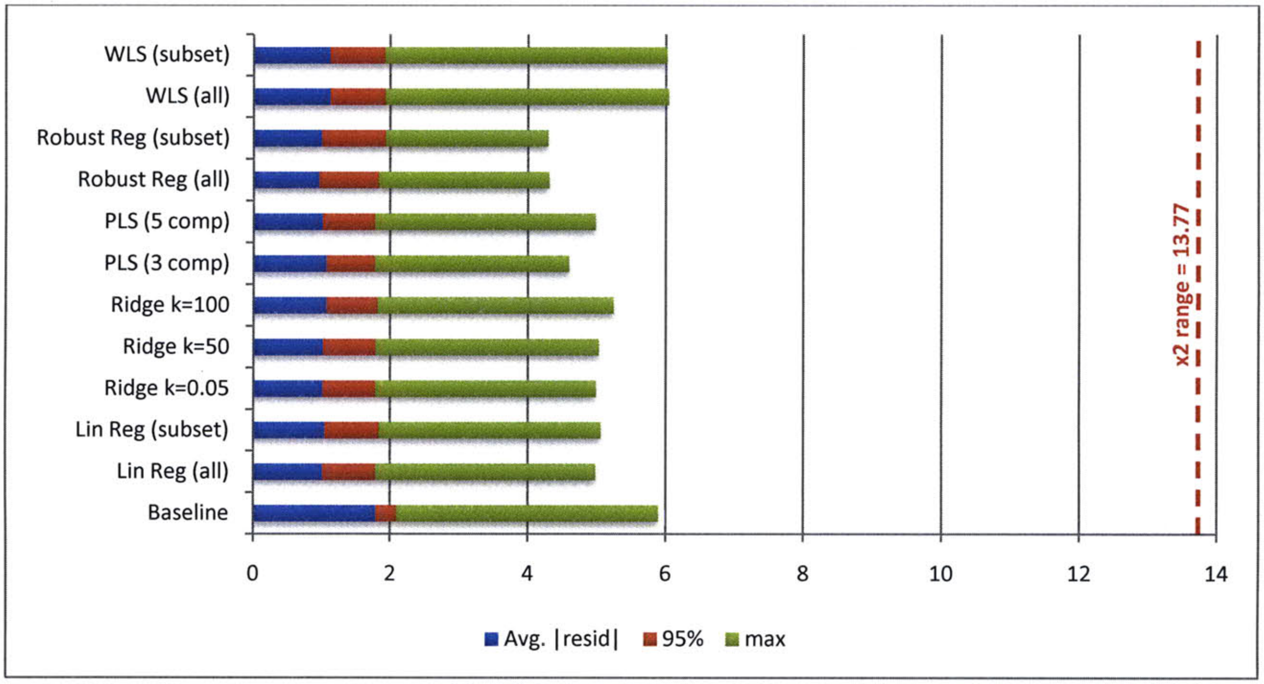


Figure 109: x2 Test Error Average, 95% C.I., and Maximum Values per Regression Model

To better evaluate the effectiveness of each regression method, Figure 109 shows the range of x2 residual values against the range of the entire x2 variable. The worst that the baseline method can do by choosing the average x2 value is about half of x2's range. Comparing this, the 95 percent confidence interval of residuals, and the average residual to the other methods, most of the regression models perform better than the baseline. Weighted Least Squares, while creating a better average residual, actually performs worse in regards to the maximum residual. It is apparent that robust regression again performs better than other regression models, creating the smallest maximum residual. However, PLS and Ridge Regression both have smaller 95 percent confidence intervals compared to robust regression, and are comparable in average absolute residuals.

x1 Regression

While some predictive power exists for x2 regression models, the x1 parameter does not cooperate as nicely. It is difficult to associate the steepness of a FRAT5 curve with different market variables. For example, the overall height of the curve implies the overall maximum sell-up behavior of those passengers in the market, but a low level curve and a high level curve may both be steep, making the market characteristics unclear in their predictive power. However, using the same regression models as x2, the regression output of coefficients is shown below in Table 32. All of the models showed an R-squared value of only 0.11, meaning that x1 is extremely variable on its own and difficult to predict.

	Lin Reg. (all)	**Lin Reg. (subset)**	**Ridge k=0.05**	**Ridge k=5.0**	**Ridge k=50**	**PLS (3 comp)**	**PLS (5 comp)**	**Robust Reg. (all)**	**Robust Reg. (subset)**	**WLS (all)**	**WLS (subset)**
(Intercept)	0.6997	0.7174	0.6996	0.6934	0.6532	0.5700	0.6646	0.6701	0.7067	0.5499	0.5541
LCC	0.0137		0.0137	0.0138	0.0139	0.0026	0.0077	0.0038			0.0110
Avg. Yield	0.0009		0.0009	-0.0030	-0.0232	0.0008	-0.0517	-0.0527			-0.0087
Percent Bus	-0.0004	-0.0005	-0.0004	-0.0004	-0.0004	-0.0004	-0.0003	-0.0005	-0.0006	-0.0005	-0.0004
Fare Ratio	-0.0199	-0.0253	-0.0199	-0.0193	-0.0157	-0.0227	-0.0188	-0.0153	-0.0204	-0.0184	-0.0137
Constr. LF	-0.1689	-0.1527	-0.1688	-0.1646	-0.1362	-0.0058	-0.1251	-0.1470	-0.1672		-0.0333
AL1 Adv.	-0.0707	-0.0695	-0.0707	-0.0700	-0.0636	-0.0716	-0.0740	-0.0610	-0.0574	-0.0682	-0.0706

Table 32: x1 Regression Models Summary

Despite the lack of predictive power of the regression models, some common coefficient results indicate that increases in the percent of business passengers, a market's fare ratio, its constraining load factor, and if the market has an advantage, the market will have a flatter FRAT5 curve. In addition, these models indicate that LCC markets should have steeper FRAT5 curves, but the LCC variable and average yield variable are often cut from the model in best subset selections.

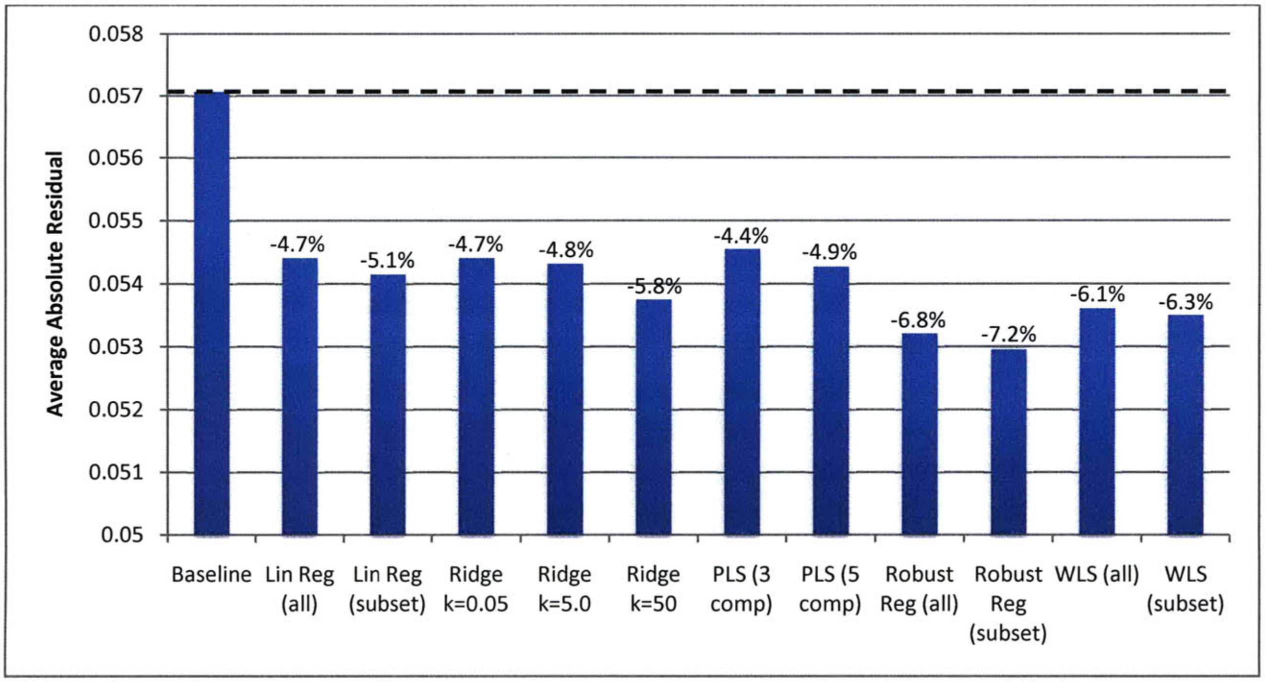


Figure 110: x1 Test Error per Regression Model

Compared to the baseline of choosing the average x1 value for every market in the test set, these regression models only provide a slight improvement (recall that the improvement over the baseline for the x2 variable is up to about 44 percent). However, robust regression again appears to be the best performing method, achieving a 7.2 percent improvement over the baseline. Relative to the overall range of x1, it is clear that the predictive power of these regression models is quite poor, showing very small improvement over the baseline.

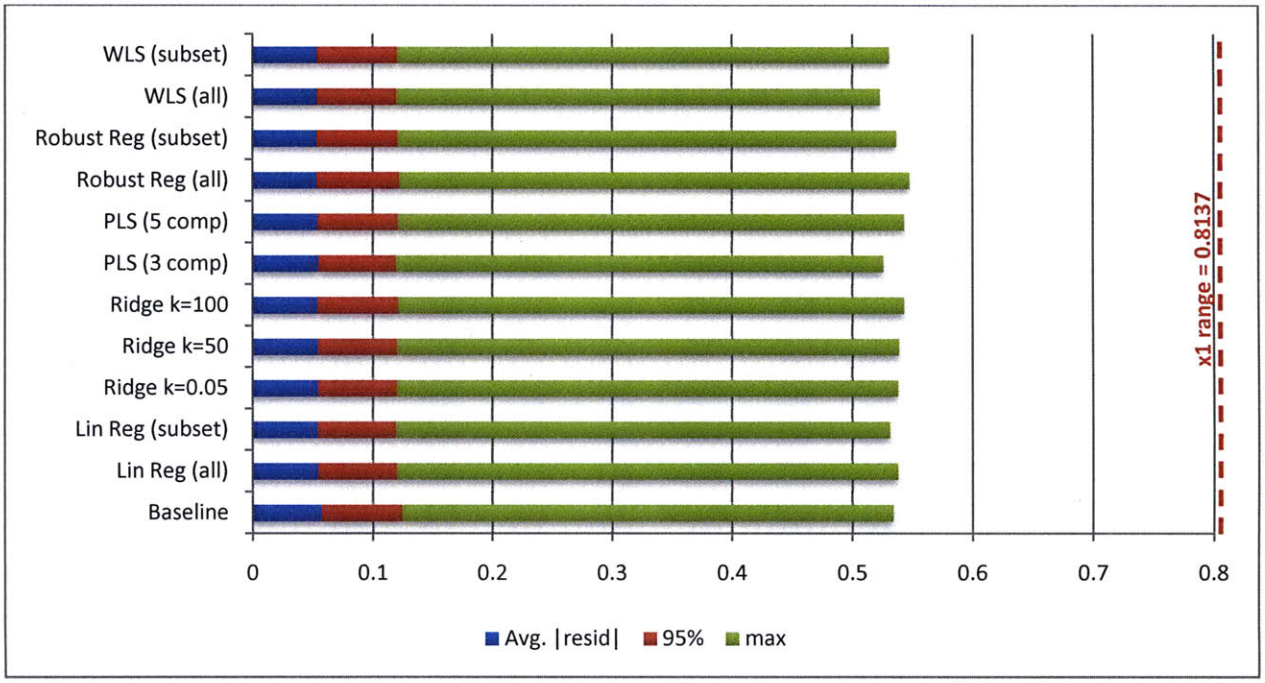


Figure 111: x1 Test Error Average, 95% C.I., and Maximum Values per Regression Model

All of the regression models have maximum residuals all around the baseline maximum residual, with some exceeding it. Similar to the x2 regression, PLS has a slightly smaller 95 percent confidence interval. Weighted Least Squares actually performs better in the x1 regression, rivaling the average residual of robust regression, and outperforming all others in the maximum residual.

Model Interpretation

These results all indicate that it is more possible and more effective to create a regression model that predicts the x2 logistic-fit parameter, which controls the overall height of the FRAT5 curve. Using the best performing and most logically sound model, robust regression (best subset), one may gain the following insights. Using the average x1 value of 0.445, and holding all else equal, it is easy to visualize the effect of each variable included in the model.

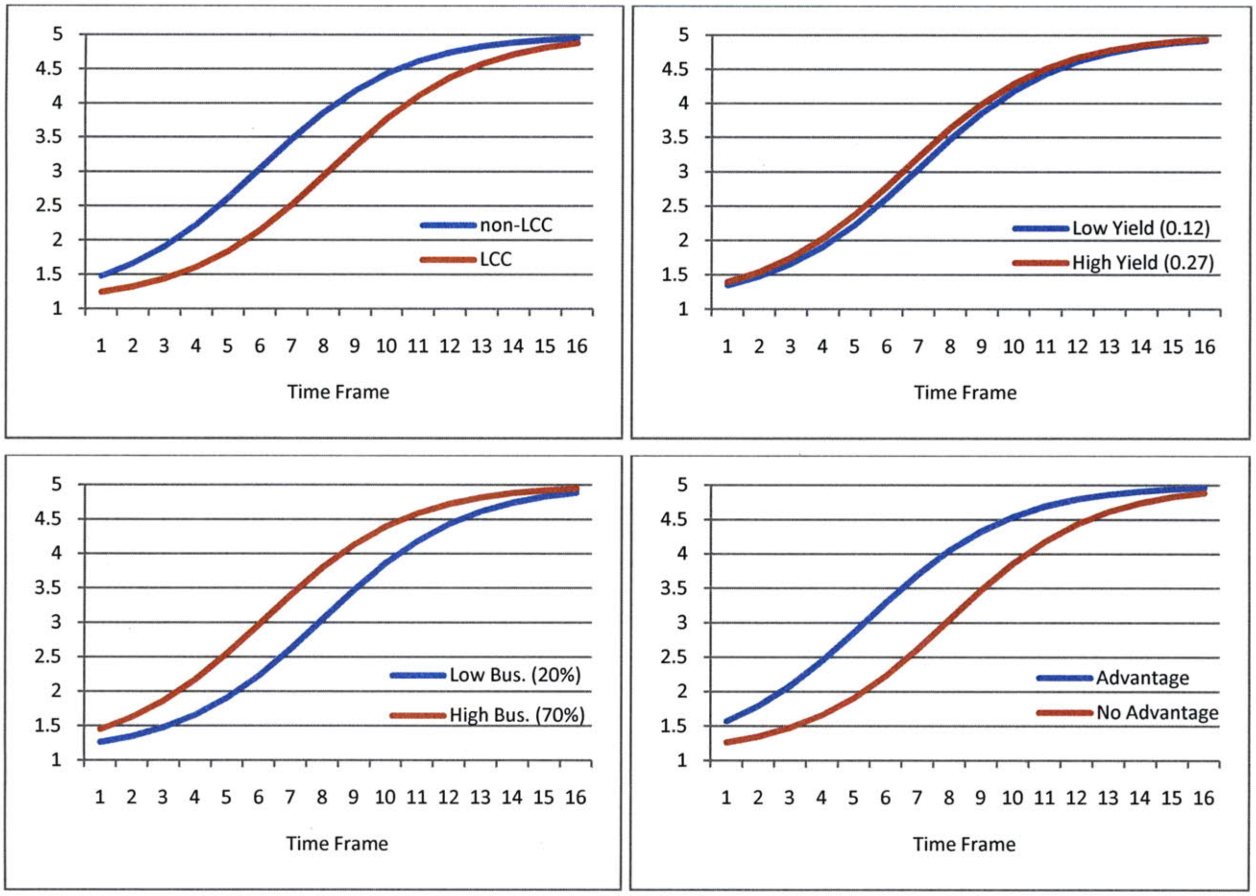


Figure 112: Robust Regression Model Impact on FRAT5 Curves

A market's non-LCC versus LCC distinction creates a large gap in FRAT5 curves, with a non-LCC market having a much more aggressive curve than LCC markets throughout the entire booking period. However, average yield plays a small role in the height of a FRAT5 curve—market with a high yield of \$0.27/mile creates a FRAT5 curve just above that of a low yield market (\$0.12/mile). In addition, if a market has a high percentage of business passengers, it is likely to have a significantly more aggressive FRAT5 curve. This is due to business passengers having a higher willingness-to-pay, making them more likely to sell-up. Last, a market with a route advantage over the airlines will have a more aggressive FRAT5 curve, indicating that those passengers will be more willing to sell-up due to the disutility of connecting flights with other carriers.

7.2 Chapter Summary

Through the use of regression models, more insight is gained about why a particular market belongs to a specific cluster. While clustering enables markets to be separated based on their values for logistic-fit parameters of x1 and x2, it was not apparent why a particular market receives those parameters beyond that of its data-based FRAT5 sell-up estimates in the booking history. Each market has several different characteristics that may influence its sell-up categorization. Regression enables the analysis of the relationship between the logistic-fit parameters and those market characteristics. After extensive investigation with various regression models correcting and accounting for heteroscedasticity and multicollinearity issues within the data set, much information is gained about a FRAT5 curve's connection to its market.

The relationship between x2 and the market's independent variables is much stronger than that of x1, indicating that x2, accounting for the height and aggressiveness of the logistic-fit curve, is the primary driver of the estimate of sell-up. Of the variables tested, the principal forces behind a market's estimate of sell-up are whether or not the market is an LCC, its average yield, percentage of business passengers, and whether or not the market is a route advantage for the airline in question. While being an LCC market creates a lower FRAT5 curve, all else equal, an increase in yield, an increase in the number of business passengers, and having a route advantage all tend to boost the aggressiveness of the FRAT5 curve.

CHAPTER 8

CONCLUSIONS

8.1 Summary of Findings

The purpose of this thesis was to explore and examine sell-up estimation and its application to airline revenue management forecasting methods. The introduction of Hybrid Forecasting, as well as Fare Adjustment, makes the estimate of passenger sell-up behavior crucial to these airline revenue management models. Incorrect estimates of sell-up may either lead to the spiral down of fare class forecasts and revenues, or to the gross overprotection of seats. To evaluate the performance of these methods in a controllable competitive environment, the thesis used the Passenger Origin-Destination Simulator (PODS), which allowed the use of various fare structures, optimizers, and airline networks.

There are several different methods that can estimate passenger willingness-to-pay, all of which are related in this thesis by the FRAT5, or the fare ratio at which 50 percent of passengers will sell-up to a higher class than their original preference for a given time frame (or point prior to departure). These FRAT5 values, when plotted over the booking period for a single market, create a FRAT5 curve that essentially defines the overall sell-up characteristic of the market in question. Both methods of estimating sell-up in this thesis, Direct Observation and Forecast Prediction, are viable methods that use actual historical booking data available in airline RM databases to estimate sell-up. These estimates are used in the application of Q and Hybrid Forecasting, where forecasts are initially reduced to the lowest "Q" class, and then redistributed to higher fare classes based on the estimates of the probability of selling up from Q to a higher class. Because sell-up observations are often sparse for various markets over the booking period, cross-time frame fitters are applied to the data. This thesis examined both a logistic function fitter and a regression-based fitter, with the logistic fitter creating more reasonable and better performing FRAT5 curves.

The main question that this thesis addressed was the determination of the level of aggregation of these sell-up estimates. Attempting to estimate a FRAT5 curve for a single market often results in wildly inaccurate sell-up estimates because of the sparseness of data, that is, actual occurrences of sell-up in historical departures. Additionally, estimating one FRAT5 curve for the entire airline system creates a sell-up estimate that is too general for all markets, forcing some markets to have an overly aggressive FRAT5 curve. One resolution to this problem, and the focus of this thesis, is to cluster the markets based on each market's logistic cross-time frame fit parameters via the K-Means clustering algorithm. This method finds the middle ground between estimating sell-up on a per-market basis and estimating sell-up over the entire system, and increases the number of sell-up observations that determine each cluster's FRAT5 curve. Compared to the best revenues created by the system-based or market-based levels of sell-up aggregation, clustering created a revenue increase of 0.26 percent in the very basic Network D6 Unrestricted, and 0.19 percent in the more complex and competitive Network T4. Additionally, relative to the baseline input FRAT5c, clustering increased revenue by 0.75 percent in Network D6 Unrestricted and by 0.26 percent in Network T4.

In addition to clustering, various input FRAT5 methods were tested. Input FRAT5s have little credibility because they are arbitrary and non-data-based, but they are important because they often indicate shortcomings of some data-based methods. Initially testing completely flat input FRAT5s indicated the importance of the FRAT5 curve in later time frames, where seat availability is vital to gain more late booking, high willingness-to-pay business passengers. Using revenue values per time frame, based on various flat FRAT5s, a piecewise step function FRAT5 was created. This input piecewise FRAT5, while arbitrary, slightly outperformed even the best data-based clustering method in Network D6 Unrestricted and Network T4. This was due the nature of the logistic-fit curve, overestimating sell-up in the middle time frames compared to the piecewise input FRAT5 curve. In order to apply the benefit of clustering to the performance of the input piecewise FRAT5 curve, it was first essential to transform the input piecewise FRAT5 curve into a data-based FRAT5 curve. However, the sparseness of the data again seemed to negatively affect the formulation of the data-based piecewise curve. Additionally, applying the clustering algorithm to these curves did not create any revenue gains over the baseline FRAT5c.

The last chapter of this thesis concentrates on the use of regression analyses of the relationship between the logistic-fit parameters that define the shape and level of the FRAT5 curves with specific market attributes. It describes what characteristics drive particular markets into having certain sell-up estimates beyond that of its actual booking data. These characteristics included whether or not it was an LCC market with an unrestricted fare structure, its average yield, percentage of business passengers, fare ratio, constraining load factor, and whether or not it has a route advantage over the competition airlines (direct versus connecting.) The focus of this chapter was primarily to predict the x2 logistic-fit parameter that determines the height of the FRAT5 curve. After correcting for multicollinearity and heteroscedasticity, robust regression (with the best subset of variables determined through stepwise regression) yielded the best results with a test error 44 percent better than a random guess of the average. The regression results suggest that increases in the percentage of business passengers, as well as an airline having a route advantage in a market, generate significantly more aggressive FRAT5 curves. In addition, an increase in average yield produces a slightly more aggressive FRAT5 curve, all else equal. In contrast, if a market is classified as an LCC, it is predicted to have a lower FRAT5 curve.

While the sell-up estimation methods presented in this thesis are primarily for use in conjunction with Q and Hybrid Forecasting, this thesis also introduced the Rational Choice Forecasting method. Rational Choice Forecasting's sell-up estimates are embedded within the regression that creates its forecasts. The independent variables include all of the passenger types possible based on the total number of fare classes and the passengers' possible willingness-to-pay levels. This method, while simpler than the two part Hybrid Forecasting plus separate sell-up estimation, does not create forecasts that are nearly as aggressive. Rational Choice Forecasting, however, does improve on standard pick-up forecasting, in all fare environments. Much of the success of Hybrid Forecasting is due to its ability to keep lower fare classes more closed throughout earlier time frames, saving capacity for later time frames when not only sell-up revenue is crucial, but revenue gained from recapture and spill-in.

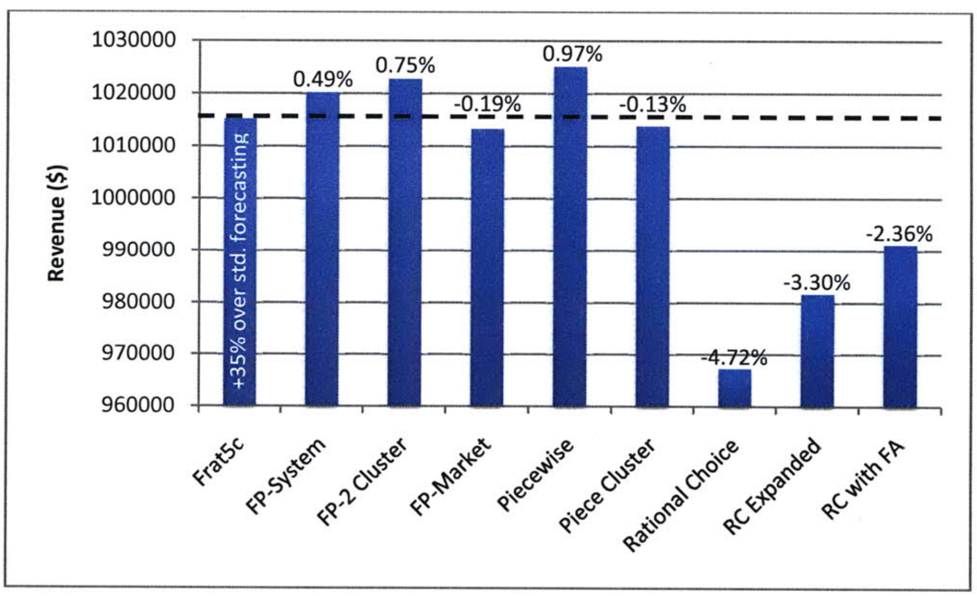


Figure 113: Summary Revenues in Network D6 Unrestricted

Even though the piecewise FRAT5 curve for Hybrid Forecasting created the largest revenue of any method tested in Network D6 Unrestricted, it is important to realize that this is an arbitrary curve, difficult for an airline to use in the real world. The best performing data-based methods were created through the use of clustering, ranging from two to seven clusters depending on the size of the simulation network. Rational Choice Forecasting as a whole performs worse than all Hybrid Forecasting methods, but it is much less complex and easier to implement.

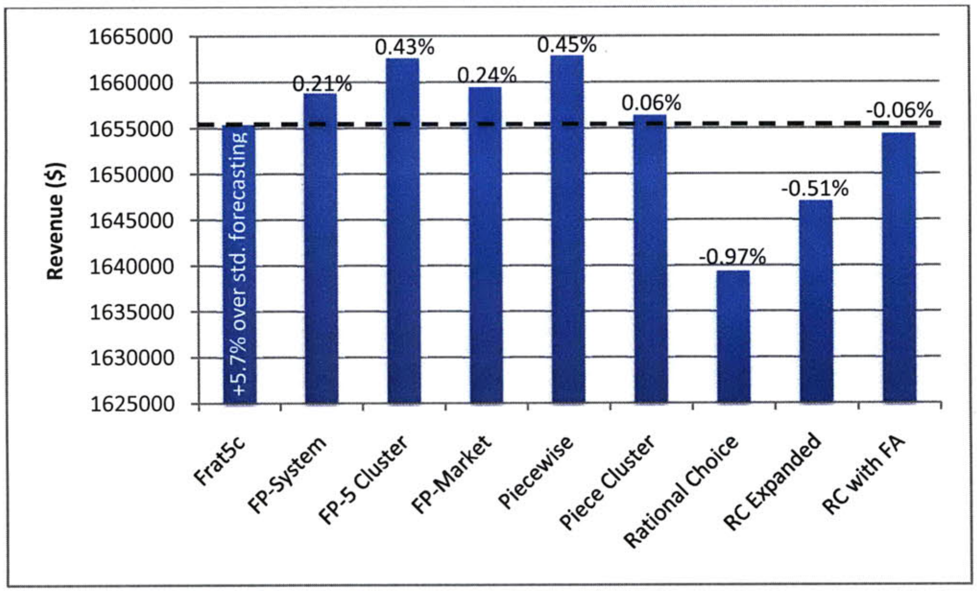


Figure 114: Summary Revenues in Network T4

In Network T4, the forecasting and sell-up estimation methods performed similarly to Network D6 Unrestricted. The best data-based clustering method nearly matches the performance of the best input piecewise FRAT5, creating a 0.43% revenue increase over the input FRAT5c. In addition, RC Forecasting performs much better relative to Hybrid Forecasting in Network T4. Rational Choice with fare adjustment nearly matches the revenue level of HF with an input FRAT5c. Overall, data-based sell-up estimation, when performed correctly with sufficient thought about data sparseness and aggregation, does provide significant benefit to forecasting in airline revenue management.

8.2 Future Research Directions

This thesis provided a comprehensive analysis of the estimation of passenger willingness-to-pay in regards to both Hybrid Forecasting and Rational Choice Forecasting in various fare structure environments. The foundation of these methods lies solely with the historical booking data for each airline. While these methods appear to benefit forecasting, there are other research directions worth pursuing to further improve them.

Many of the statistical problems with data-based sell-up estimation revolve around just dealing with the historical booking data. Perhaps an accurate estimation of willingness-to-pay should not be dependent on the bookings themselves, for these observations are obviously controlled by what fare classes are available. To better gather true unconstrained demand and eventually an estimate of sell-up, other sources of information exist, such as the Internet. A lot of today's Internet bookings are purchased by price-oriented passengers in search of the lowest fares. Using an Internet flight search engine to measure demand for particular markets through the number of searches for a specific market is another method worthy of consideration. Also, tracking a sample of the population through surveys or travel website memberships over time as prices evolve may lead to a better understanding of a passenger's true willingness-to-pay.

In addition to using alternate resources to gain more sell-up information, perhaps more extensive research can be devoted to continuing the regression analysis. Using existing market characteristics may lead to the actual prediction of a FRAT5 curve without the use of any booking data. Rather, the curve would be based solely on various market factors, such as its fare ratio and percentage of business passengers. Further pursuing data mining methods, in conjunction with regression analysis, would create numerous possibilities and potential insights to be gained in this research area.

In conclusion, this thesis motivates the need to further investigate sell-up estimation and its importance to airline revenue management. With an abundance of data available in airline revenue management systems, the opportunity exists to explore more sophisticated statistical estimation techniques for predicting passenger willingness-to-pay and choice behavior.

REFERENCES

Andersson, S.-E. (1998). Passenger Choice Analysis for Seat Capacity Control: A Pilot Project in Scandinavian Airlines. *International Transactions in Operational Research , 5* (6), 471-486.

Barnhart, C., Belobaba, P., & Odoni, A. (2003). Applications of Operations Research in the Air Transport Industry. *Transportation Science , 37* (4), 368-391.

Belobaba, P. (1987). *Air Travel Demand and Airline Seat Inventory Management.* PhD Thesis, MIT, Department of Aeronautics and Astronautics, Cambridge, MA.

Belobaba, P. (2002). Airline Network Revenue Management: Recent Developments and State of the Practice. In D. Jenkins (Ed.), *Handbook of Airline Economics* (pp. 141-156). New York: McGraw-Hill.

Belobaba, P. (1989). Application of a Probabilistic Decision Model to Airline Seat Inventory Control. *Transportation Science , 37* (2), 183-187.

Belobaba, P. (2007). Comparison of RM Methods for Reversing the Spiral Down. *AGIFORS YM Study Group Meeting.* Jeju Island, South Korea.

Belobaba, P. (2009). Fundamentals of Pricing and Revenue Management. In P. Belobaba, A. Odoni, & C. Barnhart (Eds.), *The Global Airline Industry* (pp. 73-112). West Sussex, United Kingdom: John Wiley & Sons, Ltd.

Belobaba, P. (1992). The Revenue Enhancement Potential of Airline Revenue Management Systems. *ASTAIR Proceedings of Advanced Software Technology in Air Transportation.* London.

Belobaba, P., & Hopperstad, C. (2004). Algorithms for Revenue Management in Unrestricted Fare Markets. *Meeting of the INFORMS Section on Revenue Management.* Cambridge, MA.

Belobaba, P., & Weatherford, L. R. (1996). Comparing Decision Rules that Incorporate Customer Diversion in Perishable Asset Revenue Management Situations. *Decision Sciences , 27* (2), 343-364.

Boyd, A. E., & Kallesen, R. (2004). The Science of Revenue Management when Passengers Purchase the Lowest Available Fare. *Journal of Revenue and Pricing Management , 3* (2), 171-177.

Carrier, E. (2008). *Modeling the Choice of an Airline Itinerary and Fare Product Using Booking and Seat Availability Data.* PhD Thesis, MIT, Department of Civil and Environmental Engineering, Cambridge, MA.

Cooper, W. L., Homem-de-Mello, T., & Kleywegt, A. J. (2006). Models of the Spiral-Down Effect in Revenue Management. *Operations Research , 54* (5), 968-987.

Fiig, T., Isler, K., Hopperstad, C., & Belobaba, P. (2010). Optimization of Mixed Fare Structures: Theory and Applications. *Journal of Revenue and Pricing Management , 9* (1-2), 152-170.

Gorin, T., & Belobaba, P. (2004). Revenue Management Performance in a Low-Fare Airline Environment: Insights from the Passenger Origin-Destination Simulator. *Journal of Revenue and Pricing Management, 3* (3), 215-236.

Guo, J. C. (2008). *Estimation of Sell-up Potential in Airline Revenue Management Systems.* Master's Thesis, MIT, Operations Research Center, Cambridge, MA.

Guo, J. C. (2008). Review: Methods for Estimating Sell-up Potential. *PODS Consortium.* Auckland, New Zealand.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning.* New York: Springer Science and Business Media.

Hopperstad, C. (2005). Passenger O/D Simulator: PODS v.10. Hopperstad Consulting Proprietary.

Kayser, M., Belobaba, P., & Hopperstad, C. (2008). *Summary of Rational Choice Forecasting and Rational Choice Sell-up Estimation.* Document, MIT, Cambridge, MA.

Littlewood, K. (1972). Forecasting and Control of Passenger Bookings. *AGIFORS Symposium Proceedings.* Nathanya, Israel.

McGill, J. I., & Van Ryzin, G. J. (1999). Revenue Management: Research Overview and Prospects. *Transportation Science, 33* (2), 233-256.

Ratliff, R., & Vinod, B. (2005). Airline Pricing and Revenue Management: A Future Outlook. *Journal of Revenue and Pricing Management, 4* (3), 302-307.

Reyes, M. H. (2006). *Hybrid Forecasting for Airline Revenue Management in Semi-Restricted Fare Structures.* Master's Thesis, MIT, Department of Civil and Environmental Engineering, Cambridge, MA.

Sa, J. (1987). *Reservations Forecasting in Airline Yield Management.* Master's Thesis, MIT, Department of Aeronautics and Astronautics, Cambridge, MA.

Swelbar, W. S. (2009, May 11). Current Airline Industry Challenges (Presentation to 16.75: Airline Management). Cambridge, MA.

Talluri, K. T., & Van Ryzin, G. J. (2004). *The Theory and Practice of Revenue Management.* New York: Springer Science and Business Media.

Talluri, K., & Van Ryzin, G. (2004). Revenue Management under a General Discrete Choice Model of Consumer Behavior. *Management Science, 50* (1), 15-33.

Tam, W., Belobaba, P., & Hopperstad, C. (2008). The PODS Primer (Version 1): Passenger Origin Destination Simulator (PODS) Summary of Processes and Models. MIT PODS Consortium.

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the Number of Clusters in a Data Set via the Gap Statistic. *Royal Statistical Society*, 411-423.

Weatherford, L. R., & Belobaba, P. P. (2002). Revenue Impacts of Fare Input and Demand Forecast Accuracy in Yield Management. *The Journal of the Operational Research Society*, *52* (8), 811-821.

Zeni, R. (2007). Can We Really Look to the Past to Forecast Demand? *Journal of Revenue and Pricing Management*, *6* (4), 312-314.