The Report Committee for Libin Koshy Varghese
Certifies that this is the approved version of the following report:

Study of a Share Based Passenger Mix Model

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Erhan Kutanoglu

David P Morton

Study of a Share Based Passenger Mix Model

by

Libin Koshy Varghese, B.Tech

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

May 2012

Dedicated to my beautiful wife, Sherin

Acknowledgements

I am grateful to Professor Erhan Kutanoglu for his guidance and support. I would like to thank him for providing me the resources needed in completion of this work. I am thankful to Rajesh Vijayagopal, Sergey Shebalov, Renzo Vaccari and the development team at Sabre Holdings for all their help and support in completion of this work.

I am thankful to Professor David P. Morton for his invaluable time in review of this report. I am also thankful to all my teachers who have inspired me over the years through their teaching and research.

I would like thank my friends Harshad, Neeraj, Raghu, Ahmed and Jayasree for their help and support.

Finally, I would like to express my thanks to my wife, Sherin, and my parents, Mary and Koshy Varghese for their constant encouragement and endless support. I am thankful to my brothers, Mobin and Sibin, for all the interesting discussions and keeping me motivated.

Abstract

Study of a Share Based Passenger Mix Model

Libin Koshy Varghese, M.S.E
The University of Texas at Austin, 2012

Supervisor: Erhan Kutanoglu

A passenger mix model (PMM) is used by airlines to find out how many passengers will fly on a fleet schedule. There are numerous ways of modeling passenger mix models and this report studies a share based passenger mix model, proposed by Sabre, and tests its efficacy against a deterministic linear program (DLP) passenger mix model. A DLP passenger mix model cannot recapture spilled passengers and requires iterations of the same model to recapture passengers. In order to eliminate the iterative nature of the DLP model Sabre proposed a new model which eliminates iterations for recapturing passengers. This report studies the proposed share based passenger mix model and compares it with the DLP model in terms of traffic allocation and speed of solution. It is found that the share based model allocates traffic in the same manner as the tried and tested DLP model.

Table of Contents

List of Tables

List of Figures

1. Introduction

A major problem that an airline has to solve is to find out how many passengers will fly on a given origin destination pair. As far as the passenger is concerned there are many itineraries which he can choose for getting from a city to another. The airline has to accurately predict how many passengers from the total demand they can capture and how many they will lose to the competition. If they do lose passengers to the competition can they change their flight schedule to capture more passengers? These questions can be answered by solving a *passenger mix model (PMM)*. The passenger mix model has the ticket revenue gained from the passengers on all flights as the objective function. A passenger chooses an itinerary based on the attractiveness of the itinerary to his needs. If a passenger, for example, is looking for a noon flight and there are five flights during the day, he is going to choose the flight which is closest to noon. A passenger mix model splits the total market demand based on consumer choice. Since a flight has a limited number of seats it can carry only as many passengers as its capacity therefore the model makes sure that only that many passengers from the total demand are allowed to fly on that leg. The demand and market share estimations are discussed in Section 3.2. The passengers who cannot be accommodated on the flight are *spilled* and they can either fly on another flight operated by the airline or on a competitor flight. If the spilled passenger chooses another flight of the same airline to fly then that passenger is said to be *recaptured* or else he is a *lost* passenger. The method of modeling spill and recapture in a passenger mix model varies and this report explores one way of doing that.

A *fleet schedule* is a flight schedule of an airline which has the optimal allocation of aircrafts which is obtained by solving a fleet assignment model (FAM). This fleet schedule is then used to solve the PMM to know how much traffic is to be assigned on a leg from all the itineraries that fly on it. Figure 1 shows the inputs and outputs of a PMM. Various passenger mix models have been proposed in the literature including Glover et al. (1982), Dror et al. (1988), Phillips et al. (1991), Farkas (1995), Kniker (1998) and Barnhart et al. (2002) which are reviewed in Section 2.

The current passenger mix model being used in Sabre is a *passenger mix bid price model*. This model is also known as a Deterministic Linear Programming (DLP) model. The DLP model does not take into account recapture of spilled passengers hence a modified version is formulated in such a way that it requires multiple iterations to recapture the spilled passengers. The motivation behind this report is to study and test a model for solving the passenger mix problem which eliminates the need for iterations of the DLP model and has to be comparable to it in terms of allocated traffic. Since the DLP model is a tried and tested model it is safe to assume that traffic allocated is accurate. Hence if the traffic allocation of the new model is comparable to the DLP model we can safely assume it is working correctly. The FAM being used has a PMM integrated in it and it is not iterative. The DLP model is iterative and hence there is an inconsistency when comparing traffic in between these two models. Therefore, we want a PMM which recaptures passengers without using iterations. The reason for studying a model which eliminates iterations is to have uniformity in traffic allocations between both the FAM and the PMM. The model studied in this report is called the *share based passenger mix model* which was recently proposed by Sabre. The report develops and tests the share based PMM and compares it to the DLP model and discusses its efficacy. This report does not propose a new model but discusses how the share based model is modified so as to meet the practical requirements such as size and solution speed.

The remaining report is organized as follows. Section 2 contains a literature review on previous work done in passenger mix models. Section 3 contains a brief description of the terminology and the notation used in the PMM models. Section 4 describes the DLP model and how spill and recapture are calculated in this model. Section 5 talks about the share based PMM, its advantages and disadvantages and how by using time windows the problem could have been reduced in size. Section 6 has the final share based model that was implemented at Sabre. The

DLP model and the share based PMM were tested on a scenario which has about 3000 markets and 2 million passengers and this is described in Section 7.

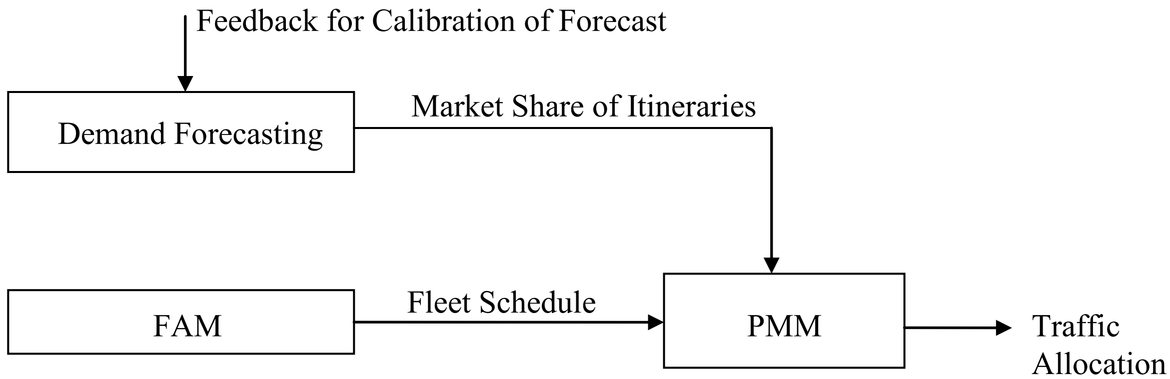

Figure 1: Inputs and Outputs of the PMM

2. Literature Review

This section deals with the literature review on the previous work done in passenger mix models. Glover et al. (1982) proposed a DLP model for identifying the optimum passenger fare-class mix for an airline. The paper describes a network-related model which divides the passengers based on fare class into an itinerary. Spill and recapture of passengers is not taken into account in this model though it is mentioned that it can be minimized by fixing the maximum seats to be allotted to a fare class. The demand difference can be studied by varying the various fare prices and then solving again to see how many passengers are gained or lost. Dror et al. (1988) propose a version of a multicommodity network flow model for allocation of passengers to an entire fleet of aircraft types over a time period of operation. This model is also a DLP model; spill and recapture of passengers are not taken into account. The existing DLP model used at Sabre is a general passenger mix model which is described in various references like Williamson (1992), van Ryzin and Talluri (2002) and Klabjan (2005). They all agree that the major drawback of the DLP model is that it does not incorporate passenger recaptures. Phillips et al. (1991) proposes an iterative method of allocating passengers by taking into account leg-dependent effects. The paper factors spill and recapture of passengers in an extension to the proposed model. Farkas (1995) discusses a combined passenger flow and FAM, but the PMM in this combined model does not incorporate passenger recapture. Farkas assumes that the spilled passengers are lost. Kniker (1998) and Barnhart et al. (2002) build on the work of Farkas (1995) and propose a combined PMM and FAM called as Itinerary-based Fleet Assignment Model (IFAM). This PMM is similar to the model explained and tested in this report. They use recapture rates to specify which airline a spilled passenger is going to choose. The objective function of the model is a minimization of spill plus carrying cost. They take leg dependencies into account and the demand is static in nature. Sandhu and Klabjan (2006) propose a fleeting model that combines both passenger flow and cargo traffic into the FAM. They use the DLP model to incorporate passenger flow into the FAM. Jacobs, Smith and Johnson (2008) uses

Benders' decomposition to integrate FAM and an origin-and-destination revenue management model. The model is called an O&D FAM and this model tackles demand uncertainty unlike the model in this report which assumes demand to be deterministic. Dumas and Soumis (2008) also propose a PMM that has stochastic demand which is not what this report assumes.

Compared to the DLP model proposed by Glover et al. (1982), Dror et al. (1988) and Williamson (1992), the share based model in this report models spill and recapture. Hence, given a fleeted schedule we can see what the true traffic is. The existing model used by Sabre is an iterative model much like the model proposed by Phillips et al. (1991). As mentioned earlier, the goal of the report is to come up with a non-iterative model. Farkas (1995) and Sandhu and Klabjan (2006) also do not calculate passenger recapture which this report does.

3. Terminology and Notation

3.1 Terminology

The terms used in the description of the models are described here. A *market* is defined as an origin – and – destination (*O&D*) pair of airports. A PMM in general has about 4000 markets. A *flight leg* is a nonstop flight between two airports. An *itinerary* is a flight schedule in which the first leg originates from the origin airport in a market and the last legs ends at the destination airport. Generally there exists only a maximum of three legs in an itinerary but there can be exceptions. There can exist, in a market, any number of itineraries. In a PMM the total number of itineraries in the problem can go up to 3 million in some cases. The *traffic* on a leg can be classified as *local* traffic and *flow* traffic. Local traffic involves those passengers who are traveling between the two airports of the leg and the starting airport is their origin and the destination airport is their final destination. Flow traffic is those passengers whose final destination is not the destination airport of the flight leg. In Figure 2, LAS-SFO is a market and there are two itineraries in that market:

1. A nonstop flight from LAS to SFO.
2. A one stop flight from LAS to SFO via LAX where the first flight leg is from LAS to LAX and the second flight leg is from LAX to SFO

In this particular example there are only two itineraries in the LAS – SFO market, but in reality LAS – LAX is also a market and there can be a nonstop itinerary *3* between them. Assuming there is just one flight operating on LAS – LAX, the traffic on this flight will have both local, from itinerary *3*, and flow traffic, from itinerary *2*.

Each market is analyzed by dividing it into time intervals. Generally, a PMM problem is solved for a week which is divided into 167 hourly time intervals. The passengers are divided into *fare classes* based on high fare (business) and low fare (leisure). Before the model is solved the demand is known for each time interval of every market and this is split among itineraries based on their *attractiveness* in that market to the passengers. The details of how attractiveness is actually calculated are explained in Section 3.2. Each flight leg has a *capacity* of passengers that it can accommodate. When the demand of all the itineraries on a flight leg exceeds its capacity, the excess passengers are said to be spilled. If there is another flight leg, in the same market, that has capacity to accommodate the spilled passengers, then the passengers are said to be recaptured. The recaptured passengers are divided among itineraries that have remaining capacity according to their re-attractiveness. *Re-attractiveness* is calculated based on similar attributes on which attractiveness is calculated, the only difference being that only the itineraries that have capacity to accommodate the spilled passengers are considered. An example is used to explain spill and recapture in Section 4.2.

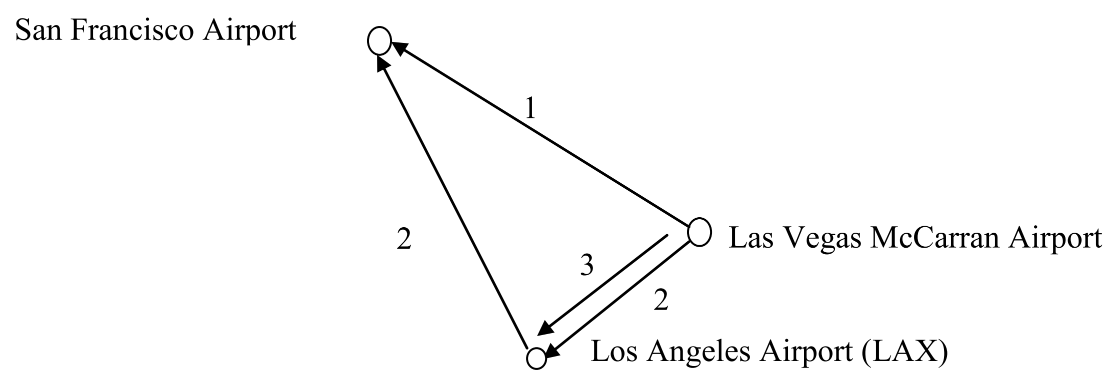

Figure 2: Airline Route Map

3.2 Determining Market Share

The total market demand is calculated by various methods of demand forecasting. The demand in the models used in this report is assumed deterministic. A multinomial logit model (Ben Akiva and Lerman, 1985) is used to estimate what percentage of this total demand each fare class of an itinerary in that market will receive. The general mathematical form of the multinomial logit model is

$$P_i = \frac{e^{V_i}}{\sum_j e^{V_j}}$$

where P_i is the probability of choosing itinerary i and V_i is the utility of itinerary i. The utility is defined as

$$V_i = \sum_j \beta_j X_{ij}$$

where j indexes the attributes in the multinomial logit model, X_{ij} is the explanatory variable of itinerary i for the j^{th} attribute and β_j is the weight given to the attributes. The term *attractiveness of an itinerary* referred to in this report is the same as the utility of an itinerary discussed above.

3.2.1 Attributes

The various attributes that are used to calculate attractiveness are

1. Service Type: This captures the passenger preference of itineraries based on their type of service. Examples of the various types of services are non-stop, through, online connections, interline connections, inter airport connections, etc. A through connection is a multi-stop itinerary where the passengers do not need to disembark the plane unless they are local passengers. An online connection is a connection itinerary in which all the flights in that itinerary are operated by a single airline. An interline connection is a connection itinerary in which all the flights in that itinerary are not operated by a single airline.
2. Elapsed Time: When there are multiple connecting flights in a market, the itineraries are penalized based on how long they are compared to the shortest connection flight in the itinerary
3. Aircraft Type: This captures the passengers' preference for one aircraft type over another. Examples of various types are wide-body jet, narrow-body jet, different engine types, etc. A wide-body aircraft can seat seven to ten passengers abreast and are generally preferred for transcontinental and transatlantic flights. A narrow-body aircraft can seat two to six passengers abreast and are generally preferred for regional flights.
4. Time: This captures the time preference of the passenger. A penalty is levied based on how far the itinerary is from the desired time the passenger wants to take the trip. This

attribute is dynamic which means that it changes with the time the passenger wants to travel. Consider that we have two similar itineraries but at different times say 8am and 4pm, for a passenger who wants to travel at 8am the first itinerary will be more attractive. This attribute brings the time preference into the utility function.

3.3 Notation

An origin and destination pair is connected by a leg l which has a capacity of C_l and all the legs we are interested in belong to the set *L*. An itinerary *i*, belonging to the set *I*, can consist of multiple legs *l*. An itinerary belongs to a market m which is defined by its starting and destination airport. All markets belong to the set *M* and each market m has a demand of D_m. The demand is divided into time intervals τ which belongs to the set *T*. With this resolution the demand can be expressed as D_m^τ also. The number of passengers allocated to an itinerary i is called traffic t_i. Different passengers belonging to the same itinerary can belong to different fare classes *j*, which belongs to the set *F*. The fare of an itinerary i for the fare class j is f_{ij}. This fare can be consolidated into a single fare f_i for an itinerary by weighting the rate with the ratio of demand of a fare class to the total demand for that itinerary. Using these resolutions the traffic can also be expressed in term of fare classes and different time intervals as t_{ij}^τ. In a market *m* it is possible that there are passengers who do not want to fly on any itinerary. These passengers are the no-fly traffic u_m and with the time interval resolution it can be expressed as u_m^τ . As there are multiple itineraries on a market m, they are ranked relatively to each other by using a multinomial logit model as described in Section 3.2. Therefore each fare class j in itinerary i in the time interval τ has an attractiveness A_{ij}^τ associated with it. This attractiveness is the same as the utility of an itinerary V_i mentioned in Section 3.2. A_{ij}^τ can be consolidated into the attractiveness of an itinerary A_i by weighting it with the probability of passengers willing to travel in the various time intervals and then weighting it by the ratio of demand for that itinerary in a fare class to the total demand for that itinerary.

3.3.1 Sets

L: Set of all legs flights indexed by *l*.

M: Set of all markets indexed by *m*.

I: Set of all itineraries.

I_m: Set of all itineraries in market *m*.

$I_{m(l)}$: Set of all itineraries in market *m* containing leg *l*.

T: Set of all time intervals indexed by τ

F: Set of all fare classes indexed by *j*

3.3.2 Data

C_l: Maximum number of passengers that can fly on leg *l*.

f_i: Fare for itinerary *i*.

f_{ij}: Fare for itinerary *i* for fare class *j*.

f_o: Penalty for not flying.

D_m: Passenger Demand in market *m*.

D_m^τ: Passenger Demand in market *m* in time interval τ.

A_i: Attractiveness of itinerary *i*.

A_{ij}^τ: Attractiveness of itinerary *i* of fare class *j* in time interval τ.

U_m: Attractiveness of the option of not flying in market *m*.

U_m^τ: Attractiveness of the option of not flying in market *m* in time interval τ.

3.3.3 Decision Variables

t_i: Traffic on itinerary *i*.

t_{ij}^τ: Traffic on itinerary *i* of passenger in fare class *j* travelling in time interval τ.

u_m: Traffic that chooses not to fly in market *m*.

u_m^τ : Traffic that chooses not to fly in market *m* in time interval τ.

4. Deterministic Linear Programming Model

The current passenger mix model being used in Sabre is a DLP model where the demand is deterministic. The market share of each itinerary in the market is calculated using the multinomial logit model described earlier. One assumption of this model is that in a market the whole demand will choose to fly if seats are available. In one run of the model, passengers are assigned to the itineraries based on their market share. If there is not sufficient capacity on the legs to accommodate the allocated passengers they are spilled from the itineraries. A new iteration is necessary to recapture the spilled passengers.

4.1 Mathematical Formulation

The DLP model maximizes total revenue from the traffic on all itineraries in all markets subject to leg capacity and attractiveness constraints.

$$\text{maximize} \sum_{m \in M} \sum_{i \in I_m} f_i \, t_i \tag{1}$$

$$\text{subject to} \sum_{m \in M} \sum_{i \in I_m(l)} t_i \leq C_l \qquad l \in L \tag{2}$$

$$0 \leq t_i \leq D_m \frac{A_i}{\sum_{j \in I_m} A_j} \qquad \forall i \in I_m, m \in M \tag{3}$$

The objective (1) maximizes the revenue from all itineraries in all markets without fare and time differentiation. Constraint (2) is a capacity constraint which ensures that the traffic of all itineraries on a leg cannot be greater than the capacity of the leg. Constraint (3) restricts the traffic on an itinerary to its market share which is the product of the utility ratio with the total market demand. The final traffic allotted by solving this linear program for an itinerary is then split into various fare classes based on the ratio of fare class demand to the total demand as shown in equation (4). The demand is the output of the demand forecasting module which is an input to the PMM as shown in Figure 1.

$$t_{ij} = t_i \frac{Demand\ of\ fare\ class\ j\ of\ itinerary\ i}{Total\ Demand\ of\ itinerary\ i} \qquad \forall j \in F \qquad (4)$$

4.2 Spill and Recapture Calculation in DLP Model

Upon solving the model above, any passenger who cannot be accommodated on a leg is spilled. To see how recapture occurs in the DLP model we can consider a simple example. Let us look at a market which has only three itineraries. Consider the flight legs in these itineraries having a given capacity and attractiveness in a market which has a total passenger demand of 10. The itineraries are graphically shown in Figure 3. Attractiveness of an itinerary is based on the attributes it has which is listed in the second column of Table 1. Also assume that an itinerary is operated by a single flight.

Itineraries in Market	Attractiveness of Itinerary	Capacity of legs	Demand Share	Traffic Allocated
Itinerary 1 • Non Stop Flight • 9am Departure	60%	4	10*60% = 6	4
Itinerary 2 • 1 Stop Flight • 12pm Departure	30%	4	10*30% = 3	3
Itinerary 3 • 2 Stop Flight • 8am Departure	10%	4	10*10% = 1	1

Table 1: Optimal Traffic Allocation after the 1st Iteration

Itineraries in Market	Re-attractiveness	Capacity	Demand Share	Traffic Allocated
Itinerary 2 • 1 Stop Flight • 12pm Departure	50%	1	2*50% = 1	1
Itinerary 3 • 2 Stop Flight • 8am Departure	50%	3	2*50% = 1	1

Table 2: Optimal Traffic Allocation after the 2nd Iteration

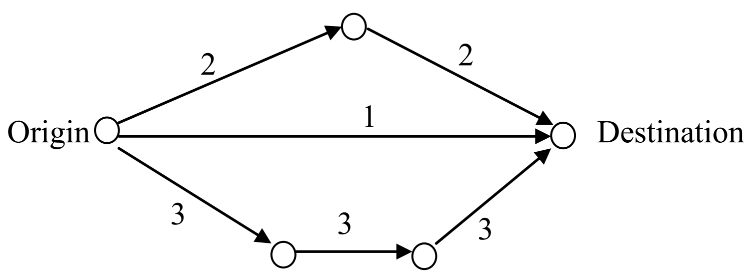

Figure 3: Itineraries in market

On solving the DLP model, constraint (2) will force the traffic on Flight 1 to be 4 passengers. In this market two passengers are lost due to spill even though Flights 2 and 3 have spare seats. So if we were to stop the optimization here we would have lost those two passengers permanently. Hence we iterate again but this time we trim the itineraries based on remaining capacity. Any itinerary that has reached its capacity is removed from set *I*. Table 1 shows the traffic allocated to the itineraries after iteration 1. The total market demand D_m is reduced by the traffic allocated on the trimmed itineraries. In the second iteration of the DLP model, Flight 1 is trimmed because it has reached capacity. The passenger demand in the market is also reduced to

2 because 8 passengers were allocated in the first iteration. The capacity of Flight 2 and 3 are also reduced based on how many passengers were allocated in the first iteration. Re-attraction rates are calculated for each itinerary in the modified set I and the model is solved again. Re-attraction rates are calculated in a similar method as the attractiveness of an itinerary which is discussed in Section 3.2. Table 2 shows the allocated traffic after the second iteration. Therefore at the end of the second iteration, in this small example, all the passengers in the market size of 10 are allocated and hence there are no further iterations.

4.3 Disadvantages

In the above example it took only two iterations for all the passengers to be allocated but that is not generally always the case. In problems where there are millions of passengers a couple of iterations are generally not enough to allocate all the passengers. This is assumed to cause long run times when the number of itineraries in the problem is significantly large. Hence in most cases any passenger not recaptured after four iterations is considered lost. Also the assumption that all the passengers will fly, given infinite capacity, may not be true as some passengers may choose not to fly at all if they do not get their desired itinerary. The last disadvantage is the inconsistency when comparing traffic between FAM and PMM as mentioned in Section 1. The aim of the models we describe in the next section is to come up with a problem formulation which eliminates the iterations needed to recapture spilled passengers and also includes variables which can accommodate the passengers who choose not to fly in a market.

5. Share Based Passenger Mix Model with Fare Classes and Time Intervals

This model is called the share based PMM because the traffic on an itinerary is not limited by its market share like the DLP model. Instead a demand constraint forces the itineraries to divide the whole market demand based on their relative attractiveness to each other. A new variable is introduced to take into account passengers who will choose not to fly in a market. To get a better understanding of traffic and how it is allocated, the decision variable in this model is changed to include the traffic on an itinerary at all time intervals and all fare classes. In the DLP model the traffic is classified as t_i but in the share based model the traffic variable is t_{ij}^{τ}. For example in this model we can examine the traffic on an itinerary and see how many passengers who wanted to fly first class (j) at 8am (τ) on Monday choose that particular itinerary (i).

5.1 Mathematical Formulation

The objective function maximizes the revenue for all fare classes, markets and itineraries. Constraint (6) is the capacity constraint where all the traffic on every fare class and every time interval of all itineraries on the leg should not exceed than the capacity of the leg. The new constraint, constraint (7), is a demand constraint that ensures that all the demand in a market in a time interval has to be met by the itineraries in that market and the no-fly traffic. Constraint (8) represents the market share constraints which assign traffic to itineraries based on their relative attractiveness to each other with respect to the time interval they are in.

$$\text{maximize} \quad \sum_{m\in M}\sum_{i\in I_m}\sum_{j\in F}\sum_{\tau\in T} f_{ij}t_{ij}^{\tau} \tag{5}$$

$$\text{subject to} \quad \sum_{m\in M}\sum_{i\in I_m(l)}\sum_{j\in F}\sum_{\tau\in T} t_{ij}^{\tau} \leq \; C_l \;\; \forall l \in L \tag{6}$$

$$t_m^{0\tau} + \sum_{i\in I_m}\sum_{j\in F} t_{ij}^{\tau} \; = \; D_m^{\tau} \quad \forall m \in M;\; \forall \tau \in T \tag{7}$$

$$0 \leq \; U_m^{\tau} t_{ij}^{\tau} \; \leq \; A_{ij}^{\tau} u_m^{\tau} \quad \forall i \in I;\; \forall \tau \in T;\; \forall j \in F \tag{8}$$

Special attention should be brought to constraint (8) as this is different from constraint (3) in the DLP model. The drawback of constraint (3) is that it restricts the traffic t_i to be only as big as its market share. Hence even if the leg has capacity, as we see in the example in Section 4.2, the constraint does not allow it to exceed its market share. But since constraint (8) takes the relative attractiveness into consideration it allows the traffic of different itineraries to be proportional to their attractiveness and not limit it to their actual market share.

5.2 Advantage over DLP model

The main advantage of this model over the DLP model is that it does not need any additional iteration for recapturing the spilled passengers because constraint (7) has to satisfy all the market demand and the spilled passengers are the ones that choose not to fly or have no seats since all the capacity is used.

5.3 Disadvantage with regard to DLP model

A major disadvantage is that the number of decision variables has increased significantly. Consider an airline which has 300,000 itineraries and a one-week problem with four fare classes. The DLP model has only 300,000 variables compared to the share based model which has

$$300{,}000\ itineraries\ X\ 4\ fare\ classes\ X\ 168\ time\ intervals = 201{,}600{,}000\ variables.$$

Constraint (7) has a total of $168\ X\ Number\ of\ markets$ constraints, while in the DLP model that constraint does not exist. Hence the problem size becomes much bigger. As the problem size increases the time taken to solve the problem also increases substantially as compared to the DLP model.

5.4 Reducing Problem Size by Including only Itineraries in a Time Window

Since the share based model size is too big, methods are investigated to reduce the size. One way is to reduce the number of variables on the demand constraint. In constraint (7), the demand at a time τ has to be satisfied by all the itineraries and the no fly variable in the market irrespective of its operating time. It is feasible to think that a spilled passenger who wants to fly on Monday 8 am will not choose an itinerary on Friday 8 am but will instead choose not to fly.

An approach would be to consider only itineraries within a specific window of τ. An example of time window is shown in Figure 4 where the time window is 2 hours. For example for the time interval of *0900* the only itineraries that will appear on constraint (7) are the itineraries in the 2 hour time window from *0800 – 1000*. In constraint (6), all the time intervals for traffic t_{ij}^{τ} will not be included but only the ones in the window. This approach drastically reduces the number of variables in the problem.

A major challenge with this approach concerns the calculation of the attractiveness A_{ij}^{τ} of an itinerary. The attractiveness of an itinerary is calculated by a demand forecasting module which serves as an input to the PMM as shown in Figure 1. The demand forecasting module gives attractiveness for an itinerary in a time interval in terms of all the itineraries in that market. If we introduce time windows we will have to calculate attractiveness for itineraries with respect to the itineraries that are in that time interval only. Consider, for example, the itinerary at *0900* which is the elongated arrow in Figure 4. The attractiveness of this itinerary in the *0800 – 1000* time window is its attractiveness with respect to the remaining itineraries that are in that particular time window and not with all the remaining itineraries. Changing the demand forecasting module would require changing the calibration methods and an entire database hence it was decided to be not feasible to use time windows and this method of reducing variables was dropped.

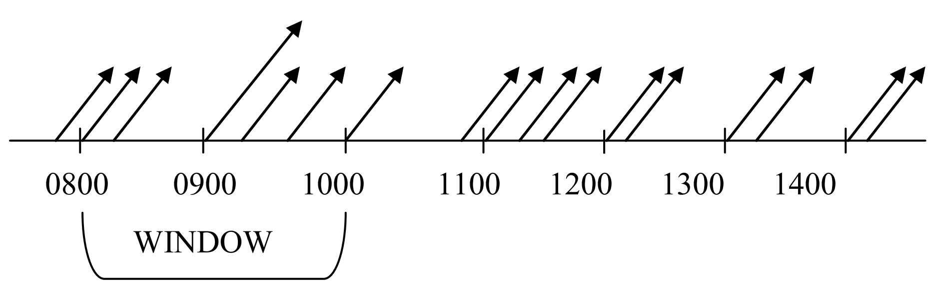

Figure 4: Airport Timeline with a Time Window

6. Share Based Passenger Mix Model without Time Intervals and Consolidated Fare Class

Since there is no major way to cut down the size of the problem through time windows, we consolidate decision variables in fare and time interval. This reduces the granularity that we can gain out of the earlier share based model with fare classes and time intervals but reduction in problem size outweighs the loss of fidelity. This model is the same structure as the earlier share based model with fare classes and time intervals but with improvements to cut down on the number of variables and constraints. The main differences are that the fare classes have been consolidated into one and the time intervals have been eliminated.

6.1 Mathematical Formulation

The model maximizes the revenue due to traffic allocation less the penalty on the traffic allocated to passengers not flying across all markets. Constraint (10) is a capacity constraint which ensures that the traffic of all itineraries on a leg cannot exceed the capacity of the leg. Constraint (11) is the demand constraint that ensures that all the demand in a market has to be met by the itineraries in that market and the no-fly traffic. Constraint (12) assigns traffic to itineraries based on their relative attractiveness to each other.

$$\text{maximize} \quad \sum_{m \in M} \sum_{i \in I_m} f_i\, t_i - \sum_{m \in M} f_o u_m \tag{9}$$

$$\text{subject to} \quad \sum_{m \in M} \sum_{i \in I_m(l)} t_i \leq C_l \qquad \forall l \in L \tag{10}$$

$$t_m^0 + \sum_{i \in I_m} t_i = D_m \qquad \forall m \in M \tag{11}$$

$$0 \leq U_m\, t_i \leq A_i u_m \qquad \forall i \in I \tag{12}$$

The changes made to the original share based model are mentioned below. This is the final model that was tested against the DLP model to see if it allocated traffic properly and if the speed of solving the problem was comparable. The results of the test follow in Section 7.

During the formulation of the problem the traffic based on fare class decision variable is eliminated and instead a single variable traffic for all the fare classes is used. The final traffic allotted by solving this linear program for an itinerary is then split into various fare classes as is done in the DLP model.

It is clear that by eliminating the time intervals we reduce the number of decision variables and also the number of constraints. The multinomial logit model is used, as in the DLP model, to estimate the market share of each itinerary based on the time interval the itinerary lies in and its attractiveness with respect to the other itineraries.

7. Comparison of the DLP model and the Share based PMM

We use a sample scenario consisting of 2993 markets, 360,000 itineraries and about 2.12 million passengers to run a test to see how the share based PMM model, discussed in Section 6, compares to the original DLP model. The average fare in the markets range from $30 to $3000 and market demand ranges from 2 to 30,000 passengers. Analysis was done to study and explain the changes in passenger revenue and spill for this test scenario. The penalty for not flying was taken as $1 and further study needs to be done for other penalty rates and its effect on the traffic allocation.

7.1 Passenger Revenue

Table 3 shows the basic results from the analysis. The total allocated traffic in the DLP model is larger than the share based model, but the share based model achieves higher total revenue by $83,149. This increase in revenue is because the share based PMM fills more high fare itineraries than the DLP model. Figure 5 shows the revenue difference based on markets ordered based on descending average fare. The graph shows a large positive difference for the high average fare markets. This shows that the share based PMM fills more high fare itineraries, hence increasing the revenue by $83,149. These findings are specific to this instance and further studies should be done on other instances.

	DLP model	**Share Based**	**Difference**
Number of Markets	2993		
Total Demand for all Markets	2,141,707.35		
Total Allocated Traffic	2,128,853.96	2,110,802.309	-18,051.65
Total Revenue	$758,812,819.52	$758,895,968.76	$83,149.24

Table 3: Traffic and Revenue Comparison

7.2 SPILL

The share based PMM spills more passengers than the DLP model. This means that it accepts fewer passengers on its itineraries than the DLP model. The main reason for this is the existence of the variable u_m in the demand constraint (7). The DLP model forces all the passengers in the market to be allocated to the itineraries as long as there is capacity but recall that the share based PMM has the no fly traffic variable, u_m, because some of the passengers would choose not to fly based on the itinerary choices they see in the market. Looking at Figure 6 we can get an idea of the difference in spill between both models. While the numbers may seem large in some markets we need to look at spill as a percentage of demand to understand it, as shown in Figure 7. The graph shows that compared to the DLP model, the new model spills not more than 3%. The large negative spill difference in some markets in Figure 7 is due to the share based model accommodating more passengers than the DLP model in those markets. Looking at a histogram in Figure 8 it is clear that the major part of the markets show a spill difference between 0 – 0.5%. Hence the share based PMM models spill is comparable to the DLP model. Hence constraint (8) is successful in eliminating the need for iterations by allocating traffic on itineraries relative to their attractiveness unlike constraint (3).

7.3 SPEED OF THE MODELS

In the case of the DLP model the run time includes reading the input, constructing data structures, running iterations and displaying the output. The run time of the share based PMM includes the same as the DLP model except for the fact that it does only a single iteration. Even though the share based model does not go through iterations it is not noticeably faster than the DLP model. The total run time required was 360s for the DLP model and 370s for the share based PMM.

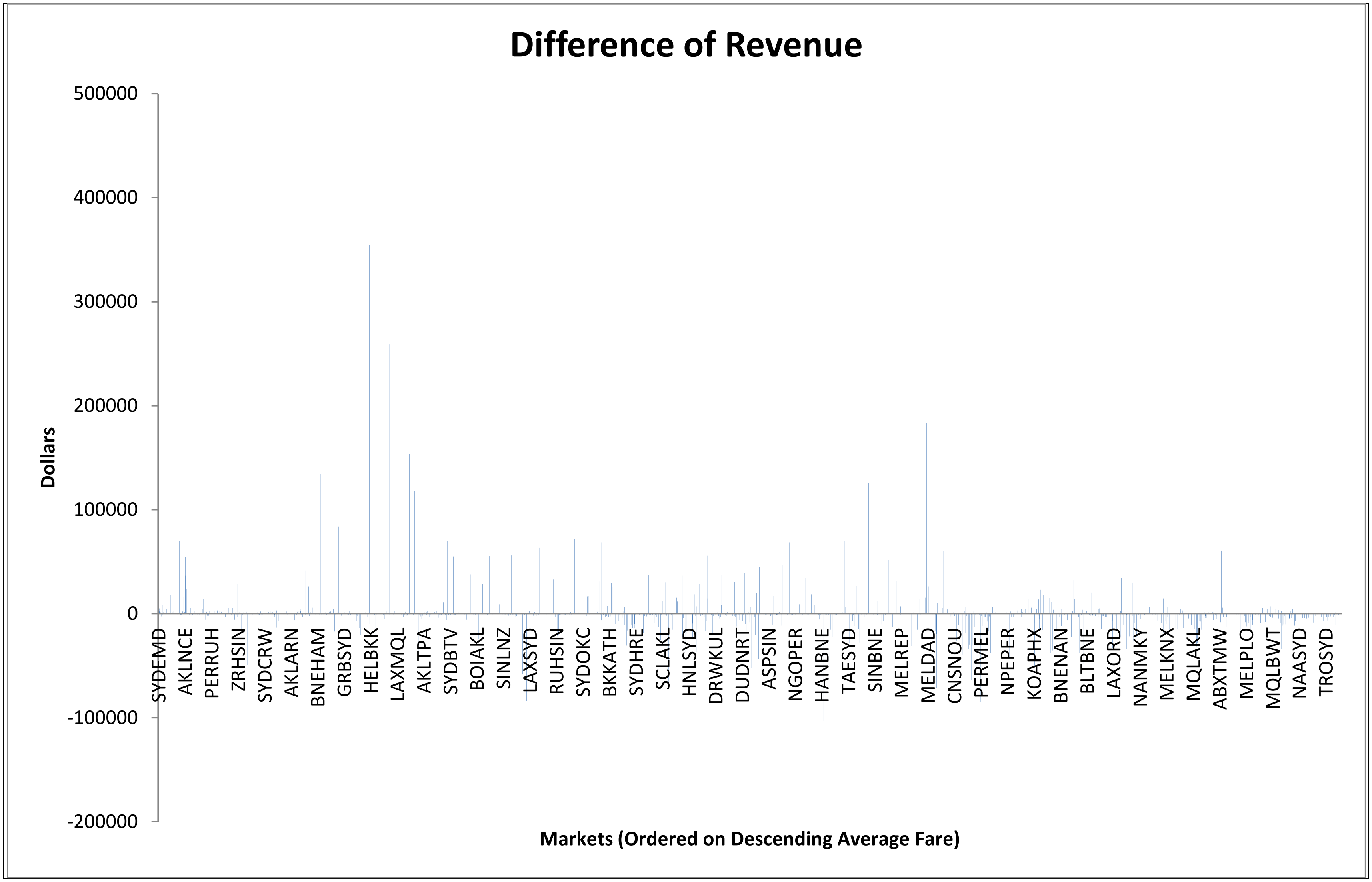

Figure 5: Difference of revenue between the models when markets are ordered based on descending average fare, where a positive difference means that the share based model had larger revenue in that market than the DLP model

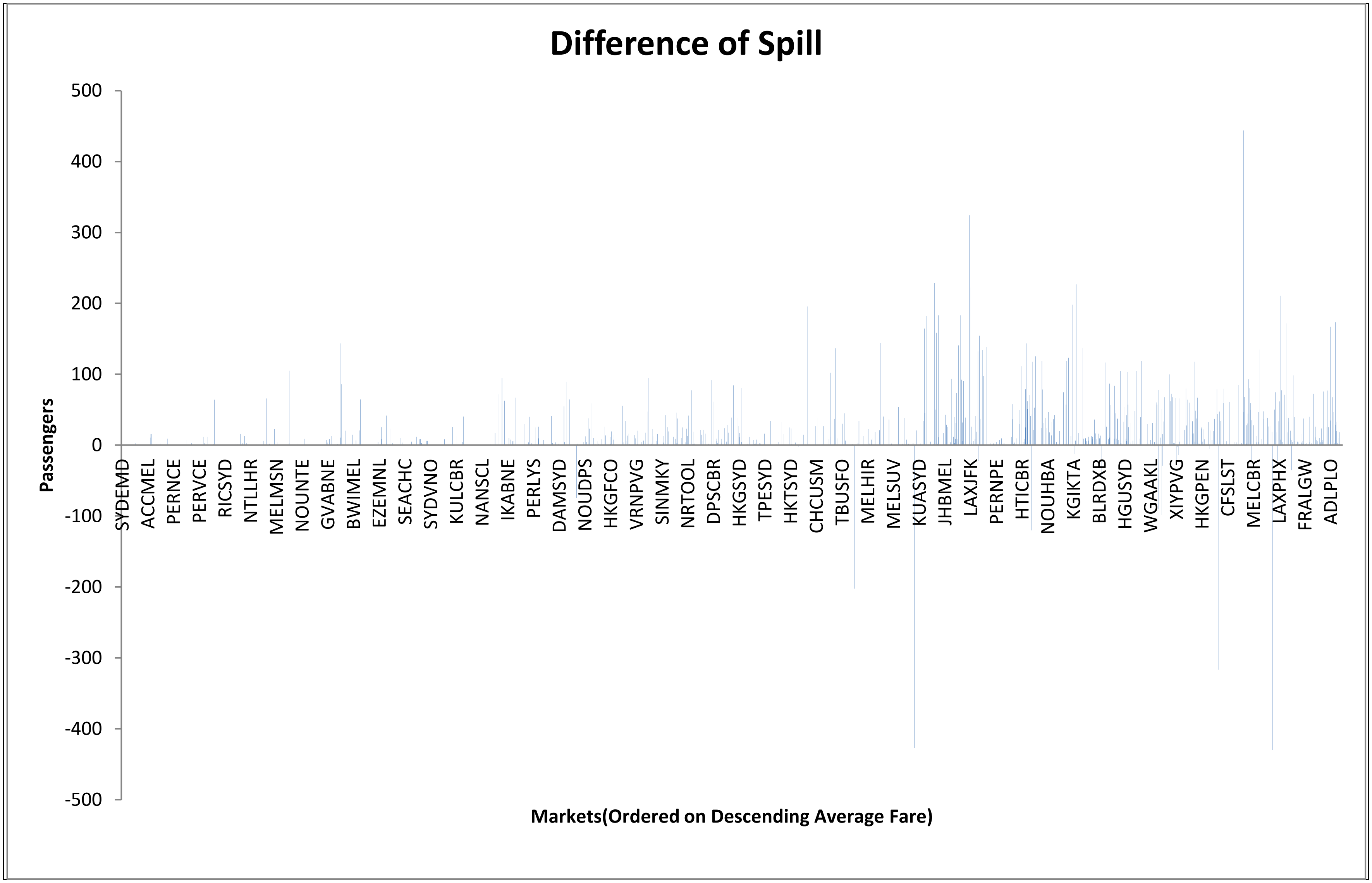

Figure 6: Difference of Spill between the two models, where a positive difference means that share based model had larger spill in that market than the DLP model

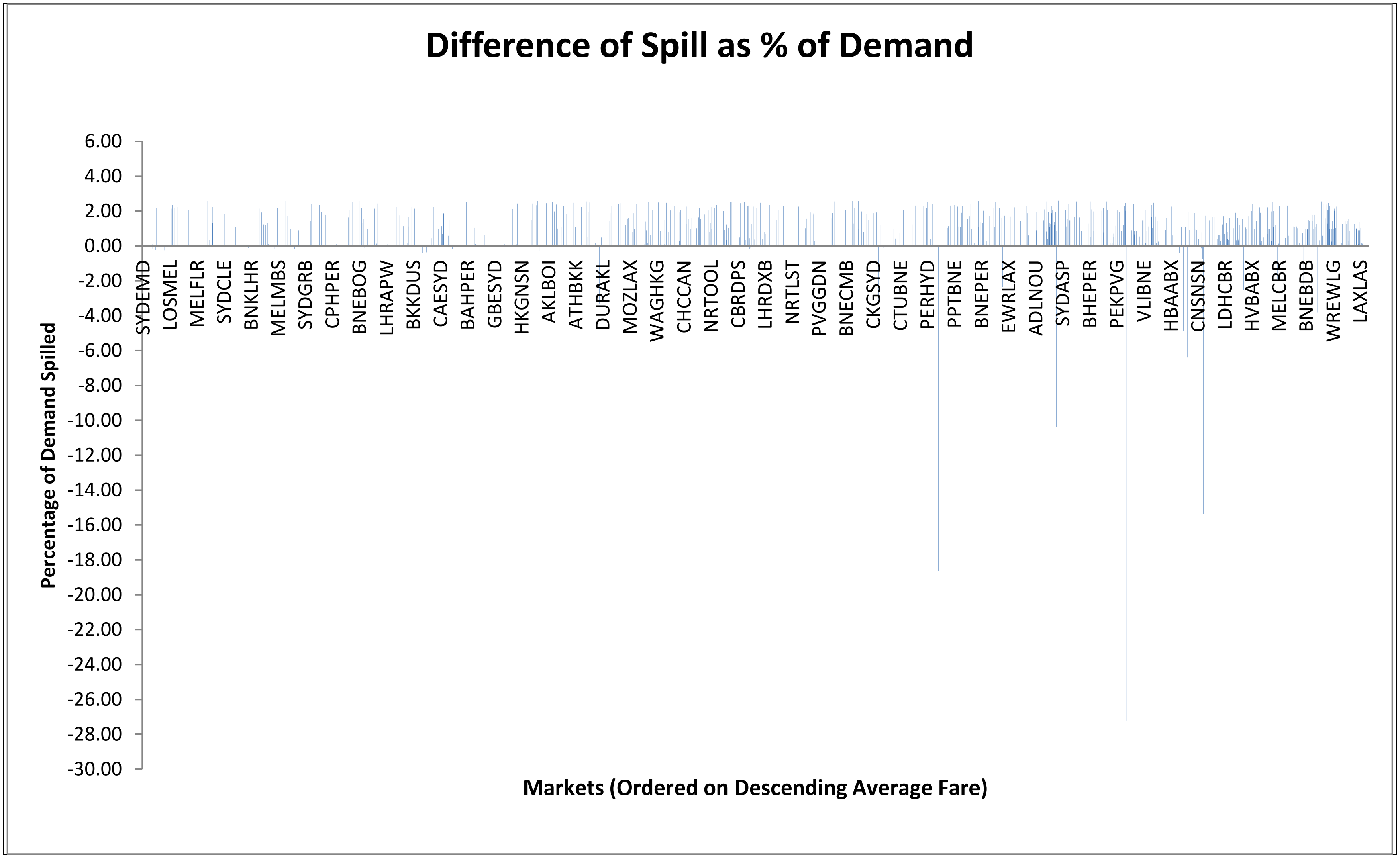

Figure 7: Difference of spill as a % of demand in a market between the two models, where a positive difference means that share based model had larger percentage of demand spilled in that market than the DLP model

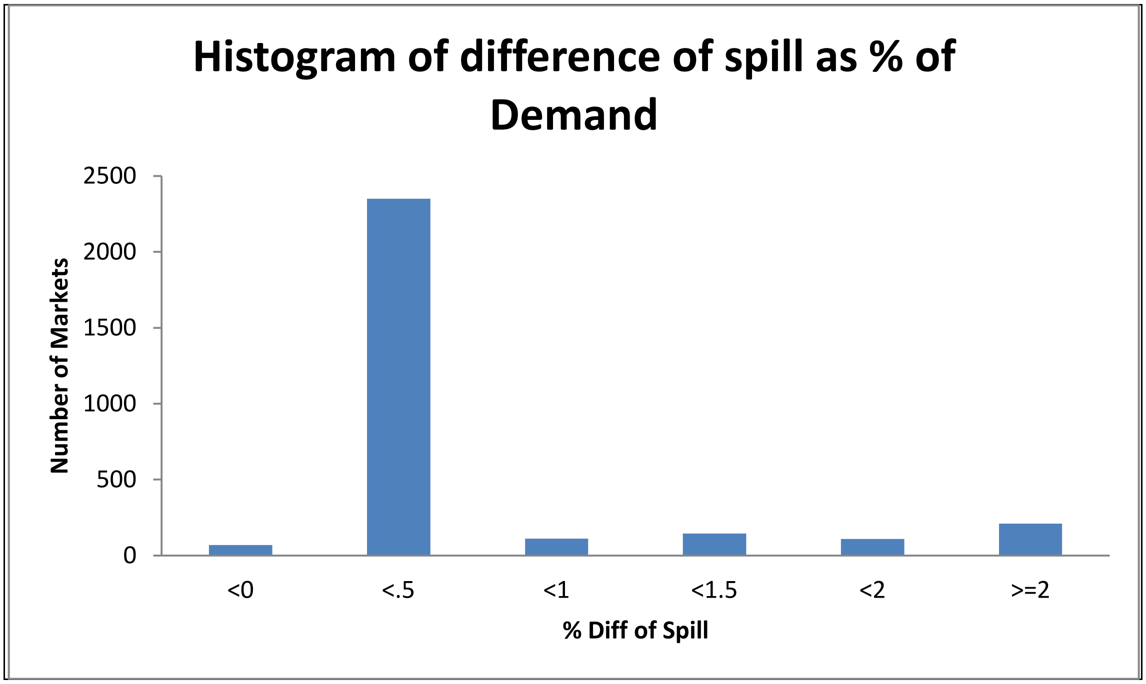

Figure 8: Histogram of difference of spill as % of demand

8. Concluding Remarks

In conclusion, the share based passenger mix model is comparable in terms of traffic allocation and speed to the DLP model. Even though the spill is greater in most cases by a small fraction, this does not result in a loss of revenue as the new model more aggressively fills the high-fare itineraries compared to the old model. The revenue earned is greater than the DLP model in this instance. By implementing this model for the passenger mix model, Sabre wishes to have a non-iterative PMM as in the FAM. Following this study the share based passenger mix model is currently under implementation at Sabre.

References

Barnhart, C., Kniker, T.S., Lohatepanont, M., 2002. Itinerary-based airline fleet assignment. Transportation Science 36, 199–217.

Ben Akiva, M., Lerman, S., 1985. Discrete choice analysis: Theory and application to travel demand. MIT Press, Cambridge, MA.

Dror, M., Trudea, P., Ladany, S.P., 1988. Network models for seat allocation on flights. Transportation Research 22b, 239–250.

Dumas, J., Soumis, F., 2008. Passenger flow model for airline networks. Transportation Science 42, 197–207

Farkas, A., 1995. The influence of network effects and yield management on airline fleet assignment decisions. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Glover, F., Glover, R., Lorenzo, J., McMillan, C., 1982. The passenger mix problem in the scheduled airlines. Interfaces 12, 73–79.

Jacobs, T. L., Smith, B. C., and Johnson, E. L., 2008. Incorporating network flow effects into the airline fleet assignment process. Transportation Science 42, 514-529.

Klabjan, D., 2005. Large-scale models in the airline industry. In Desaulniers, G., Desrosiers, J., and Solomon, M.M., editors, Column Generation, Kluwer Academic Publishers, Norwell, MA.

Kniker, T.S., 1998. Itinerary-based airline fleet assignment. Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Phillips, B.L., Boyd, D.W., Grossman, T.A., 1991. An algorithm for calculating consistent itinerary flows. Transportation Science 25, 225–239.

Sandhu, R., and Klabjan, D., 2006. Fleeting with passenger and cargo origin-destination booking control. Transportation Science 40, 517-528

Williamson, E.L., 1992. Airline network seat inventory control: Methodologies and revenue impacts. Ph.D. Dissertation, MIT Flight Transportation Laboratory, Cambridge, MA.

van Ryzin, G., and Talluri K., 2002. Revenue management. Hall, R., editor, Handbook of Transportation Science, 599-661. Kluwer Scientific Publishers, Norwell, MA, 599-661

Vita

Libin Koshy Varghese was born in Kollam, Kerala, India. Upon completion of his schooling in Doha, Qatar, he pursued Bachelor’s degree in Mechanical Engineering from Mar Athanasius College of Engineering, Kerala, India. He started his graduate studies at the University of Texas at Austin in fall 2010. He is currently interested in optimization of large scale systems.

Email: libin@utexas.edu

This report was typed by the author.