# Modeling Airline Passenger Choice: Passenger Preference for Schedule in the Passenger Origin-Destination Simulator (PODS)

by

Emmanuel Carrier

M.A., Economics
Université Paris-I Panthéon-Sorbonne, 1997

B.A., Business
Hautes Etudes Commerciales (HEC), 1996

SUBMITTED TO THE DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN TRANSPORTATION AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2003



Signature of Author.......................................................................................................
Department of Civil and Environmental Engineering
May ___, 2003

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

Accepted by .......................................................................................................
Cynthia Barnhart
Chairman, Transportation Education Committee

Accepted by .......................................................................................................
Oral Buyukozturk
Chairman, Departmental Committee on Graduate Studies

**Modeling Airline Passenger Choice: Passenger Preference for Schedule in the Passenger Origin-Destination Simulator (PODS)**

by

Emmanuel Carrier

Submitted to the Department of Civil and Environmental Engineering on May, 14, 2003 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Transportation

# ABSTRACT

This thesis examines how to model the choice of individual travelers among various possible travel alternatives in the airline industry. A review of the models used to represent that choice situation in the Passenger Origin-Destination Simulator (PODS) was undertaken for two reasons. First, the development of computational capabilities has lead to advancements in consumer choice theory that enabled the implementation of more flexible models like mixed logit models. Second, the increasing competition of low-cost new entrant airlines has put great pressure on pricing practices of traditional network carriers. This increasing competition has also compelled these carriers to focus on their strengths, for example, schedule coverage. In this thesis, after a comparison between the PODS Passenger Choice Model and the literature on consumer choice theory, we will then focus on how to model passenger preference for schedule.

The review of the literature on air traveler choice reveals that most authors have used discrete choice models, like standard logit or nested logit models, to represent the choice of individual passengers among travel alternatives. However, the logit model has two limitations in the air traveler choice problem: it can accommodate neither random taste variation in some elements of the passenger utility function nor the complex substitution patterns across travel alternatives modeled in PODS. However, we show that the highly flexible mixed logit model brings a solution to these limitations and the choice process modeled in PODS can be approximated by a set of mixed logit models.

In the second part of the thesis, we focus on how passenger preference for schedule is modeled in PODS. In the current model, a constant replanning disutility is added to the cost of all paths that are not convenient to the passenger. However, the current approach does not differentiate among paths based on their level of schedule inconvenience and this leads to distortions in the valuation of the revenue advantage of the carrier offering the best schedule. We propose in this thesis an alternative approach called the variable replanning disutility model. In this model, the replanning disutility added to the cost of paths depends on the time location of the path and its level of schedule inconvenience. PODS simulation results show that the variable replanning disutility model leads to a more realistic valuation of the revenue advantage associated with a better schedule coverage.

Thesis Advisor: Dr Peter Paul Belobaba

Title: Principal Research Scientist, Department of Aeronautics and Astronautics

# ACKNOWLEDGEMENTS

First of all, I would like to thank my academic advisor, Dr. Peter Belobaba, for his support and advice during this long journey. His knowledge of the airline industry is immense and I have been able to learn so much about air transportation and revenue management since I joined MIT and started working with him.

I would like also to thank Craig Hopperstad for his modeling ideas as well as his programming talents. His passion for research in airline revenue management as well as his dedication to the PODS simulator is unparalleled.

I would like also to express my gratitude to Professor Moshe Ben-Akiva. Through his classes and our discussions, he helped me make progress in the understanding of discrete choice models and the references he provided have been crucial for the development of this research.

This research was funded by the PODS consortium, which gathers some of the best revenue management expertise in the industry. I would like to thank all the sponsor airlines of this research, Continental, Northwest, Delta, Lufthansa, KLM, SAS and Swiss for their financial support and all the participants of the PODS summits for their comments and their ability to survive 15 straight hours of PODS presentations.

In addition, it was great to work with an amazing team of ICAT and PODS students. In particular, I would like to thank Seonah Lee, whose last job at MIT was to give me my first PODS lesson, to Alex Lee for his huge knowledge of PODS simulations, to my fellow PODS students Adeem Usman and Diana Dorinson. In addition, thanks to my fellow mileage junkies Thomas Gorin and Andrew Cusano for introducing me to the world of airline frequent flyer programs. Working with you has made it so much fun.

This thesis would not have been possible without the support of my family. I thank my father for giving me a passion for the transportation industry and my mother for teaching me that everything in life is possible. I also thank my brother Bruno and my sister Florence for their support. However, this work would not have been possible without the love, care and extensive culinary expertise of you, Ngoc.

Finally, this thesis is dedicated to the loving memory of my grandmothers, Paulette Ullman and Ginette Lion.

# TABLE OF CONTENTS

# LIST OF FIGURES

CHAPTER 5

# LIST OF TABLES

# Chapter 1 Introduction

## 1.1. *Setting, Purpose and Motivation*

In liberalized air transportation markets, the consumers of air transportation services have typically the choice between several fare class products on one or more available flight itineraries offered by the airlines serving the desired markets. Depending on their characteristics and preferences, air travelers will choose a particular airline, flight and fare class to fulfill their travel needs. Demand for air travel at the path/class level is then the result of the tradeoffs individual air travelers make when they choose among different airlines, flights and fare classes.

Gaining insight about air traveler preferences and understanding the determinants of demand for air travel at the path/class level is important to support key airline planning decisions like flight scheduling, pricing, fare class restrictions design and seat allocation among path/classes (revenue management). In that context, MIT and a consortium of seven leading airlines developed in the mid-nineties the Passenger Origin-Destination Simulator (PODS) to examine the impact of revenue management methods, especially seat allocation decisions at the airline network level. However, unlike most revenue management simulators used previously, demand for each particular path/class in PODS is not exogenous; it is the result of millions of individual choices at the air traveler level between available airlines, flight schedules and fare classes. As a result, what makes PODS unique among airline revenue management simulators is its passenger choice model that reproduces how hypothetical air travelers choose among various airline, flight schedule and fare class products available in an air travel market.

Over the years, PODS and its Passenger Choice Model have proved to be a valuable tool to measure the impact of various revenue management methods designed to maximize airline revenues at the network level. With the increase in computational power during the nineties, PODS has been used to simulate passenger choices in increasingly complex airline networks including transatlantic alliance networks and has produced stable and consistent results in the study of various revenue management issues: origin-destination revenue management methods, alliance revenue management, forecasting, passenger behavior issues like sell-up and recapture.

A review of the PODS Passenger Choice Model today seems necessary for two main reasons. The first reason is driven by a structural change in the airline industry with the increasing competition from low-cost new entrant airlines. These new carriers along with the downturn in airline business travel since 2001 have put a great pressure on the fare and restriction structure used by the traditional network carriers. As a result, PODS member airlines have shown great interest in using PODS to study the revenue impact of changes in their pricing and restriction policies. However, the PODS Passenger Choice Model has been primarily designed and calibrated to study the impact of revenue management decisions like seat allocation decisions in the context of the competition between several traditional network carriers offering similar fares and restriction structures. In order to test the impact of a new entrant airline or different fare/restriction structures, a review of some elements of the PODS passenger choice model would be useful and a recalibration of the model could well be necessary.

In addition, due to the development of low-cost competition in the United States and Europe in the recent years, network carriers need to focus on their strengths including network coverage and frequency of service. These industry trends have lead some consortium members to show interest into investigating the impact of schedule asymmetry on PODS simulation results. However, the PODS Passenger Choice Model has been conceived to simulate the competition between airlines offering similar schedules but using different revenue management strategies. As a result, to assess the impact of offering a superior schedule on airline revenues requires reviewing how passenger preference for schedule is modeled in PODS.

The second reason is driven by progress in the theory of consumer choice. The development of computational capabilities in the nineties has also enabled significant progress in consumer choice theory, in particular in the field of discrete choice models. Essentially, these advancements have been mostly centered on the use of simulation methods, which is the researchers' response to the inability of computers to perform complex integration. Simulation allows the estimation of otherwise intractable models: almost any model specification can be estimated by some form of simulation. As a result, the researcher is freed from previous constraints on model specification and is not limited to a few model specification alternatives that have favorable mathematical properties but also some severe limitations. Simulation allows a more creative, precise and realistic representation of the hugely varied choice situations that arise in the world, including air traveler choice among airlines, flight schedule and fare class. A new class of discrete choice models has emerged that offer a lot more flexibility than standard models used in the past. As a result, we can compare the PODS Passenger Choice Model and its hypotheses to these new developments in discrete choice models.

The objective of this thesis is then to review how air traveler preferences are modeled in PODS and relate the options used in the PODS Passenger Choice Model to the new competitive environment in the industry and the advancements in consumer choice theory. In particular, this thesis will compare the PODS approach to the literature and theory on air traveler choice focusing on discrete choice models including the new classes of flexible discrete choice models like mixed logit models. In addition, due to the interest expressed by some consortium airline members and the potential for improvement identified during the passenger choice model review, this thesis will include a case study on how to model passenger preference for schedule in PODS.

## *1.2. Outline of the Thesis*

Chapter 2 reviews the literature on consumer choice theory focusing in particular on discrete choice models used to describe the choice of a consumer among a discrete number of alternatives. This chapter includes a description of the most widely used discrete choice model, the logit model, its advantages and its limitations and some more complex and flexible models developed to overcome these limitations like for instance nested logit models or mixed logit models.

Chapter 3 provides a detailed description of the PODS Passenger Choice Model including how it was developed based on the Boeing Decision Window Model. In addition, this chapter compares the approach used in the PODS Passenger Choice Model with models used in the transportation literature primarily discrete choice models.

After analyzing the PODS Passenger Choice Model in its entirety and relating it to the consumer choice literature, the two subsequent chapters focus on one particular element of the choice model: passenger preference for flight schedule. Chapter 4 reviews the literature on schedule choice in intercity travel and compares it to the approach used in PODS. Based on that analysis, alternative approaches are designed.

In Chapter 5, PODS is used to simulate the impact of alternative approaches to model passenger preference for flight schedule. Detailed analysis of the simulation results is included in this chapter.

Finally, Chapter 6 concludes this thesis with a summary of the findings from the literature review, the comparison between the PODS approach and the methods used in the literature and the flight schedule case study. At the very end of this chapter, we address some of the issues for future research directions involving the PODS Passenger Choice Model and discrete choice models in air transportation.

# Chapter 2 Discrete Choice Models

## 2.1. Introduction

Discrete choice models describe decision-makers' choice among various alternatives. To fit within a discrete choice model framework, the set of alternatives called the choice set must exhibit three properties: the alternatives must be mutually exclusive, collectively exhaustive and the number of alternatives must be finite. Indeed, the first and second criteria are not restrictive as the researcher can always ensure that the alternatives are mutually exclusive and collectively exhaustive by an appropriate definition. In contrast, the third condition is actually restrictive as it is the defining characteristic of discrete choice models and distinguishes their realm of application in consumer choice theory from that of regression models, where the dependent variable is continuous, which means that there are an infinite number of possible outcomes.

In addition, discrete choice models usually assume that the decision-maker has a rational behavior and will choose the alternative that maximizes its utility. However, the utility that each alternative brings to the decision-maker is not known with certainty but is divided into two parts: an observed element known to the researcher and a random element, which remains unknown. As a result, the researcher is not able to predict precisely the choice of the decision-maker (the alternative that maximizes its utility) but rather estimates the probability that each alternative might be chosen. This probability depends on the observed part of the utility known to the researcher and the assumed distribution of the error terms (random element). As a result, discrete choice models are known as random utility maximizing models.

The purpose of the PODS Passenger Choice Model is to reproduce the choice of individual travelers among various travel alternatives defined along three dimensions: the choice of an airline, a flight schedule and a fare class. In the real world, air travelers have a choice between various travel alternatives when planning a trip. The number of available alternatives varies market by market based on the number of airlines serving that market, the number of flights offered by each airline and the number of fare-class products they actually market. However, this number is always finite for a given time period. Each individual traveler has to make a choice among a finite number of possible travel alternatives. The choice set might vary across decision-makers based on their preferences or on their access to information but the number of alternatives is always finite.

As a result, the choice of a travel alternative by an individual air traveler fits within the discrete choice model framework. The study of the most widely used discrete choice models, with their respective strengths and weaknesses is then necessary to better understand the design of the PODS Passenger Choice Model and to compare it to the approaches usually used in the literature.

## *2.2. The Logit Model*

The logit model is by far the easiest and most widely used discrete choice model. The description of the logit model in this section is based on Ben-Akiva and Lerman (1985). The popularity of logit is due to its mathematical simplicity: the formula for the choice probabilities takes a closed form and is readily interpretable. To derive the logit model, let us introduce the following notation. A decision-maker labeled n faces a choice among J alternatives. The utility that the decision-maker obtains from alternative j is decomposed into a part labeled $V_{nj}$ that is known by the researcher and an unknown part $\varepsilon_{nj}$ that is supposed to be random:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \ \forall \ j$$

The logit model is obtained by assuming that each random element $\varepsilon_{nj}$ is distributed independently and identically extreme value. The probability that the decision-maker n chooses alternative $i \in J$ is then:

$$\begin{aligned} & \text{Prob } (V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}) \ \forall \ j \\ \Leftrightarrow \quad & \text{Prob } (V_{ni} - V_{nj} > \varepsilon_{nj} - \varepsilon_{ni}) \ \forall \ j \end{aligned}$$

Given the extreme value distribution of the error term, the choice probability of alternative i becomes:

$$\Pr(i) = \frac{\exp(Vi)}{\sum_j \exp(Vj)}$$

If $V_j = a\ X_j \ \forall \ j$ with $X_j$ observed by the researcher the formula then becomes:

$$\Pr(i) = \frac{\exp(aXi)}{\sum_j \exp(aXj)}$$

The value of the parameters a can be estimated using maximum likelihood techniques.

Using the extreme value distribution for the error terms is nearly the same as assuming that they are normally distributed. The extreme value distribution gives slightly fatter tails than a normal, which means that it allows for slightly more aberrant behavior than a normal distribution. But the key assumption of the logit model is not the shape of the distribution but the independence of the error terms. This means that the unobserved utility of one alternative is unrelated to that of another alternative, which can be fairly restrictive. Stated equivalently, this means that the researcher has specified the systematic part of the utility $V_{nj}$ precisely enough that the remaining unobserved portion is just essentially white noise. This can be considered as the ideal goal of any researcher: specify the utility so well that a logit model is appropriate. Seen in this way, the logit model is ideal rather than restrictive. If the researcher thinks that the unobserved portion of the decision-maker utility is correlated across alternatives, he has basically three options: use a different model that allows for such a correlation, re-specify the systematic utility so that errors are now uncorrelated or use the logit model as an approximation. The last option might however lead to some errors, especially if the researcher plans to investigate substitution patterns.

The logit model has two main advantages: its mathematical simplicity and a very large flexibility in the definition of the choice set. As already mentioned, the choice probabilities take a closed form and are very easy to calculate. In addition, the choice set can vary from an individual to the next individual and only a subset of the alternatives can be included in a decision maker particular choice set. Indeed, the standard logit estimation procedure by likelihood maximization remains valid if only a subset of alternatives is included in the choice set, if all alternatives have the same chance of being chosen into each decision-maker choice set (uniform conditioning).

However, the logit model has also three main weaknesses: it cannot accommodate random taste variation in the population, it implies a very specific substitution pattern and it is not appropriate to deal with panel data. Let us examine first the issue of random taste variation. Random taste variation occurs when there is heterogeneity in the population response to an alternative attribute.

For instance, the impact of a Saturday night stay restriction associated with a discount fare may vary from traveler to traveler and this variation might be unobserved by the researcher. Some travelers, especially those with family commitments, might consider that having to stay over the weekend at their destination is a very serious disadvantage and has a very negative impact on their utility: they will give a very high value to the Saturday night stay coefficient. However, for some other travelers like young unmarried students, staying at destination over the weekend might not be such a hassle and could even be seen as an opportunity. These passengers will give a very low value to the Saturday night stay coefficient. As a result, the coefficient of the Saturday night stay in the utility function of the discount fare alternative is not fixed but follows an unknown distribution: this variation in the population response is called random taste variation.

If tastes vary with unobserved parts of the utility, then the logit model is not appropriate as the error terms will necessarily be correlated across alternatives. A logit model is then a misspecification. As an approximation it might be able to capture the average taste fairly well since the logit formula is typically fairly robust to misspecifications. However, even if the logit model were to provide an acceptable approximation of the average taste, it does not give information on the distribution of tastes around the average. This distribution can be important in many situations and to incorporate random taste variation appropriately, a mixed logit model will be preferred.

In addition to the random taste variation issue, the logit model implies a very specific substitution pattern among alternatives. Due to the assumption of independence between error terms, the ratio of choice probabilities of two alternatives remains constant (the independence of irrelevant alternatives property or IIA property) and there is a proportional substitution between alternatives. Any increase in the choice probability of one alternative leads to a decrease in choice probabilities of all other alternatives by the same percent. This very specific substitution pattern can be clearly unrealistic in some situations as illustrated by the famous blue bus/red bus paradox.

Suppose that a commuter has the choice between using his private car or riding the bus to go to work and that each alternative has a 50% choice probability. Now suppose that another bus service is introduced that is equal to the existing buses in all its attributes except for the color of the bus. We now have red and blue buses as well as driving a private car as the all the available alternatives. Under the logit model, the choice probability of each alternative is 33.33%. However, this is unrealistic because the commuter will most likely consider the two bus modes as similar and treat them as a single alternative: in this case the probability of the car alternative will remain 50% and each of the bus alternatives will have a 25% choice probability. Proportional substitution between alternatives in this case seems completely unrealistic and the logit specification is not an appropriate approach to model such a choice situation.

However, as already mentioned, the IIA property of the logit model has a major advantage: it allows to estimate model parameters consistently only on a subset of alternatives (if each decision-maker choice set is chosen randomly). This can be a tremendous benefit when the number of alternatives is so high that estimation would be otherwise too computer-intensive. It allows also great flexibility as the choice set can vary across decision-makers.

Whether the IIA property seems realistic or not can be tested. Indeed, if the IIA property holds, the coefficient estimates obtained on a subset of alternatives are not significantly different from those obtained on the full set of alternatives. A test of that hypothesis constitutes a test of the IIA property and several procedures have been defined like for instance the McFadden-Hausman test (McFadden and Hausman, 1984). In addition, as the logit model is often a special case of more complex models, the IIA property can generally easily be tested.

The third limitation of the logit model is with panel data. Data that represent repeated choice over time by the same decision-maker are called panel data. If the unobserved factors that affect the choice of decision-makers are independent over the repeated choices, then logit can be used with panel data. However, in most cases, errors can be assumed to be correlated over time. In these situations, either the model needs to be re-specified to bring the sources of correlation into the observed part of the utility or another model like mixed logit should be used.

The air traveler choice problem i.e. the choice by an individual air traveler of an airline, a flight schedule and a fare class might involve all three main limitations of the logit model. For instance, we can reasonably assume that there is some significant heterogeneity in the response of the air traveler population to some parameters like schedule convenience or the disutilities associated with low-fare restrictions. Indeed, air travelers flying for business purposes are known to place a high emphasis on schedule convenience and flexibility and people flying for leisure purposes on price. Even within the population of business and leisure travelers, they should be significant differences on how passengers value these elements of their utility function.

In addition, in the case of the air traveler problem, we do not expect the IIA property to hold. Indeed, we expect for instance a higher degree of substitution among lower restricted discounted fare class products than between discounted fare classes and fully flexible full fare products. Similarly a higher degree of substitution can be expected between two flights offered by the same airline and two flights offered by two different airlines. As a result, a model able to accommodate more flexible substitution patterns than the logit model may be preferred.

Finally, a large proportion of air traffic is actually flown by a relatively small population of regular frequent fliers. Indeed, most airlines have developed very complex and extensive frequent flyer programs. Membership to these frequent flyer programs is open to all air travelers but their benefits are non-linear and they are especially targeted to seduce that regular frequent flyer population. These regular users of air travel services typically make repeated choices of airlines, flight schedule and fare class and these repeated choices can be assumed to be fairly correlated over time based on the decision-maker preferences and characteristics. As a result, a model able to take into account some correlation between repeated choices over time might be useful to our analysis of the air traveler choice problem.

In the case of the air traveler choice problem, the assumptions of the logit model are actually fairly restrictive. Another model specification that is able to account for random taste variation, complex substitution patterns and correlation between repeated choices over time is probably more appropriate. The next two sections will then describe two alternatives to the standard logit specification: the Generalized Extreme Value (GEV) family of models that allows integrating more complex substitutions patterns and the mixed logit model, which is fully general and highly flexible and provides a solution to all three limitations of the logit model.

## 2.3. The GEV Family of Discrete Choice Models

Generalized extreme value (GEV) models constitute a large class of models that exhibit a variety of substitution patterns. GEV models are consistent with utility maximization and their unifying attribute is that the unobserved portion of utility for all alternatives is jointly distributed as a generalized extreme value, which allows for some correlation patterns across alternatives. GEV models relax one of the three limitations of the logit model and have the advantage that the choice probabilities usually take a closed form such that they can be relatively easily estimated without resorting to simulation.

The most widely used model of the GEV family is called the nested logit model. The nested logit model is appropriate when alternatives can be grouped into nests and exhibit the following substitution patterns: the ratio of the choice probabilities of any two alternatives in the same nest is independent of the attributes or existence of other alternatives. IIA holds within the nest. However, for two alternatives in different nests, the ratio of probabilities can depend on the attributes of other alternatives. IIA does not hold across nests. The error terms are correlated for two alternatives in the same nest but remain independent for alternatives in different nests.

A consistent way to picture the substitution patterns is with a tree diagram. In such a tree, each branch denotes a subset of alternatives within which IIA holds and every leaf on each branch denotes an alternative. There is proportional substitution across twigs within a branch but not across branches.

The tree in Figure 2.1. illustrates the situation in which there is a proportional substitution pattern among various flights offered by the same airline but not across flights from different airlines:

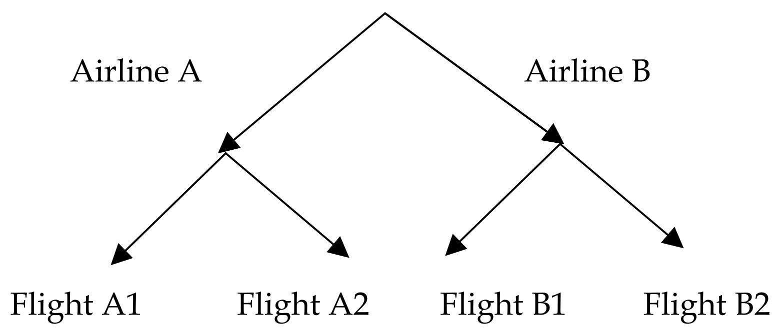


**Figure 2.1.** : The nested logit tree structure

In this case, if airline A were to introduce a third flight in the market, demand for flights A1 and A2 would decrease by the same proportion but that proportion would be different from the decrease in passenger demand for flights B1 and B2.

If we suppose that $U_{nj} = V_{nj} + \varepsilon_{nj}\ \forall\ j$ where $V_{nj}$ is observed by the researcher and $\varepsilon_{nj}$ is an unobserved random variable, the nested logit model is obtained by assuming that the vector of errors has a certain type of generalized extreme value distribution. Then, the choice probabilities take a closed form and the model can be estimated using maximum likelihood techniques. The standard logit model is of course a special case of the nested logit model in which the generalized extreme value of the error terms collapses into an iid extreme value distribution. Indeed, it is possible to test the nested logit specification against the logit specification and verify if IIA might hold across nests (McFadden and Hausman, 1984).

If the nested logit model has the ability to accommodate some non-proportional substitution patterns, it can only apply if the choice situation can fit within this particular tree structure. This means that alternatives can be grouped into nests and that the choice problem must be divided into several dimensions with a specific hierarchy between these dimensions. In the example above, the choice of passengers is bi-dimensional with first the selection of an airline and then the selection of a particular flight schedule. We will discuss in the next chapter whether such a hierarchy is appropriate in the case of the air traveler choice problem.

In the standard nested logit, each alternative belongs to only one nest. This limitation is sometimes restrictive and several kinds of GEV models have been specified with overlapping nests to accommodate more complex substitution patterns. However, the GEV family of models does not provide the researcher with a complete freedom in exploring all kinds of substitution patterns. In addition, GEV models are not a solution to the other two limitations of the logit model i.e. random taste variation and panel data. The next section describes the mixed logit model, which resolves all three limitations of the logit model but, unlike logit and GEV models, requires the use of simulation methods to estimate the choice probabilities.

## 2.4. The Mixed Logit Model

Mixed logit is a highly flexible model that can approximate any random utility model. It resolves all three limitations of standard logit models and allows for random taste variation, any substitution pattern and correlation in unobserved factors over time. The mixed logit model is defined on the basis of the functional form of its choice probabilities. Any behavioral specification whose derived choice probabilities take this form is called a mixed logit. The description of the mixed logit model in this section is based on Train (2000).

Mixed logit choice probabilities are the integral of standard logit probabilities over a density of parameters.

$$P_{ni} = \int L_{ni}(\beta) f(\beta) \partial\beta$$

$$\text{where } Lni(\beta) = \frac{\exp(Vni(\beta))}{\sum_j \exp(Vnj(\beta))}$$

and $f(\beta)$ is a density function. $V_{ni}(\beta)$ is a portion of utility that depends on parameters β. If utility is linear in β, then $V_{ni}(\beta) = \beta' x_{ni}$. Then the mixed logit probability takes its usual form:

$$Pni(\beta) = \int \frac{\exp(\beta' xni)}{\sum_j \exp(\beta' xnj)} f(\beta) \partial\beta$$

It is a weighted average of logit choice probabilities evaluated at different values of the parameters β with the weights given by the density $f(\beta)$. In the statistics literature, the weighted average of several functions is called a mixed function and the distribution that provides the weights the mixing distribution.

Standard logit is a special case of mixed logit model, where the mixing distribution is degenerated to fixed parameters. This mixing distribution can also be discrete. A discrete function can be a useful specification if there are distinct segments in the population, each of which has its own behavioral pattern. In most cases, it is however a continuous function. It can be specified to be a normal or lognormal distribution. By specifying the explanatory variables and mixing distribution appropriately, the researcher can represent any type of random utility maximizing behavior as well as many forms of non-utility maximizing behavior.

An issue of terminology arises in mixed logit models. There are two sets of parameters that enter a mixed logit formula. There are the parameters used in the logit formula and there are the parameters that describe the mixing distribution, typically its mean and variance. Usually, the researcher is interested in estimating the second ones. As a result, we will focus here on estimation techniques to get estimates of the mixing distribution parameters.

Using a mixed logit specification to represent random taste variation is then straightforward. The utility specification is the same as for standard logit except that the parameters are supposed to vary across decision-makers rather than being fixed (the parameters are random variables). For each decision-maker, the researcher observed the value of the explanatory variable but neither their coefficient, nor the unobserved part of the utility function. The researcher has then to specify a distribution for each coefficient of the systematic utility and estimate the parameters of this distribution. Several specifications are possible: normal but also lognormal when the coefficient is known to have the same sign for all decision-makers like for instance a price or cost coefficient.

For instance, if we go back to the example of the Saturday night stay requirement developed in Section 2.2., as mentioned earlier, in a logit model, we need to assume that the coefficient of the Saturday night stay restriction is fixed. However, under a mixed logit specification, the coefficient of this restriction is not necessarily fixed. It can vary from one decision-maker to the next. For instance, we can assume that it follows a normal distribution and estimate its mean and variance. As a result, a mixed logit specification allows the researcher to investigate heterogeneity in response of the population to some part of the utility function.

A mixed logit model can also be used without a random coefficient interpretation but to simply represent error components that create correlation among the utilities of different alternatives. The error component is then composed of two terms, one that is distributed iid extreme value across alternatives and another one that can be correlated over alternatives. Various correlation patterns, hence substitution patterns, can be obtained by an appropriate choice of the variables that enter the error component. For example, an analog of nested logit is obtained by specifying a dummy variable for each nest that equals 1 if the alternative belongs to the nest and zero otherwise. The variance of the dummy coefficient will capture the magnitude of the correlation of alternatives that belong to the nest.

In fact, any random utility model can be approximated by a mixed logit specification with the appropriate choice of variables and mixing distribution. A mixed logit specification just requires that part of the error component is distributed iid extreme value. Adding an iid extreme value term to the utility of all alternatives might change the decision-maker behavior. However, by scaling up the utilities appropriately, the researcher can assure that this will never occur. As a result, adding an extreme value term to the true utility, which makes the model into a mixed logit does not change it in any meaningful way when the scale of the utility is sufficiently large. A mixed logit can approximate any random utility model by simply scaling up utility sufficiently. However, in most cases, this scaling-up might not be necessary if some part of the true utility can be assumed to be iid extreme value. In this case, the researcher's task is just to find variables and a mixing distribution that capture the other parts of utility, i.e. the parts that are correlated.

Once the researcher has specified the model, the estimation procedure is composed of two steps. First, the choice probabilities are approximated by simulation. The choice probability simulation proceeds as follows: draws of the parameters are taken from the mixing distribution. Then, for each draw, the choice probability is calculated using the classical logit formula. The simulated choice probability is the average of the choice probabilities calculated for each draw of the parameters. These simulated estimates are unbiased estimators of the true choice probabilities. Their variance diminishes as the number of draws used in the simulation increases.

In a second step, these choice probabilities estimates are used to estimate the mixing distribution parameters through for instance a maximum likelihood procedure. Under some conditions, these simulated maximum likelihood estimators will be unbiased and consistent estimates of the unknown true parameters.

## *2.5. Some Applications in the Air Transportation Literature*

Prossaloglou and Koppelman (1999) use a logit model to investigate the choice of air carrier, flight and fare-class. They consider air travelers as rational decision-makers that choose the alternative with the highest utility. The authors justify the existence of an error term in the trip utility function to account for the lack of complete information, possible measurement errors and the inability to properly observe and account for all factors affecting choice behavior. The factors that are examined as explanatory variables include fare class restrictions, fares, carrier market presence, quality of service, participation in carrier frequent flyer programs and flight schedules. Separate models were estimated for business and leisure passengers.

Estimation of these logit models is based on stated-preference data. Data collection was based on a two-tier survey: initial data concerning passenger characteristics (past trips, trip purpose, permanent address, frequent flyer membership) was collected through a mail survey. A random sample of mail survey respondents was then chosen for a phone-based survey designed to simulate individual travelers' search for information about air travel options and the selection among available alternatives like during a booking process. Based on the answers to the mail survey, each interviewee was presented with the scenario of either a business trip or a vacation trip in one of the two following markets: ORD-DEN (7 morning flights available on three different airlines) or DFW-DEN (9 morning flights available on 4 different airlines). Business travelers had a three-day advance notice and had the choice between three fare classes: first class, unrestricted coach and restricted coach. Leisure travelers had three-week advance notice and also the choice between three fare classes all with restrictions. Each traveler had to ask the agent over the phone on the available alternatives (carrier, flight, fare class) and finally make a booking decision based on the information provided and their own preferences like during a booking process with a regular travel agent.

The results of the model suggest significant differences in travel behavior between leisure and business travelers. As expected, business travelers are more time-sensitive and less price-sensitive than leisure travelers. Indeed, business travelers are willing to pay $60 per hour of reduced schedule delay compared to only $17 for leisure travelers. They also place more emphasis on frequent flyer programs. They are willing to pay a $21 premium to travel on an airline, which frequent flyer program they already belong to and $52 more to fly with the airline of their most preferred frequent flyer program. For leisure passengers, those values are only $7 and $18 respectively.

In addition, there has been a number of studies of airport choice in multi-airport regions that are based on discrete choice models, especially logit and nested logit models. For instance, Kanafani (1983) uses a multinomial logit model to study the choice the choice of airports by air travelers flying between the Los Angeles metropolitan area and the San Francisco Bay area. The explanatory variables in his model include for instance the frequency of service at each airport and the level of the fares.

Regarding the application of mixed logit models to the air transportation field, Mehndiratta (1996) studies in his doctoral dissertation the impact of time of day preferences on the scheduling of business trips in the domestic US focusing mainly on trips involving air transportation. His assumption is that the current models used in inter-city travel analysis do not take into account the spatial and temporal variations in the value of time and that these variations have a large influence on the choice of travel alternatives at the individual level. He attempts to incorporate these variations in a discrete choice model and to study their impact on the selection of travel alternatives.

Mehndiratta divided a regular 24-hour schedule into three periods: work, leisure and sleep time. He proposed and formulated a theory to accommodate variations in the value of time among these three periods of the day. He then proceeded to implement his theoretical approach. As he wanted to test whether there might be some heterogeneity in the population response, he used a mixed logit model specification to study the impact of disruption of work, leisure and sleep time on the choice of intercity travel alternatives. The coefficients for the value of disruption of leisure and sleep time were specified to be random. Based on stated preference data, he estimated the mean and standard deviation of the distribution of these random coefficients. As the standard deviation of these coefficients turned out to be statistically significant, the assumption of heterogeneity in response of the population to disruption of various periods of their regular schedule was validated.

In addition, another conclusion of this study is that business travelers tend to give a higher value to sleeping time than to work and leisure time. In addition, sleep and leisure time spent at home or around home is more valuable than leisure time and sleep time at the business destination. As a result, travelers will avoid staying overnight at their destination unless staying home and leaving very early in the morning would disrupt too largely their normal sleep schedule.

## *2.6. Summary*

In this chapter, we have described the most widely used discrete choice models with their respective strengths and weaknesses. In addition, we have shown how these models have been used to study a variety of choice situations that air travelers may face including the choice of an airline, a flight schedule and a fare class. In the next section, we will describe how these choices are modeled in PODS and how the PODS Passenger Choice Model relates to these theoretical models used in consumer choice theory in general and in air transportation choice in particular.

# Chapter 3 The PODS Passenger Choice Model

## 3.1. Introduction

In the previous chapter, we have described the discrete choice models usually found in the literature on consumer choice theory and applied in the air transportation field. This chapter will concentrate on the Passenger Origin-Destination Simulator and in particular on its passenger choice model that reproduces the choice process of an individual air traveler among various possible travel alternatives. After a general overview of the simulator, we will focus on a description of all the elements that affect air traveler choice, how they are modeled in PODS and how the PODS approach relates to the theory introduced in the previous chapter.

## 3.2. Overview of the PODS Simulator

PODS is a computer simulation tool designed to investigate airline revenue management techniques. It was originally developed by Hopperstad, Berge and Filipowski at the Boeing Company and is an extension of the Boeing Decision Window Model used to study the impact of flight schedules on airline market share. For a description of the Boeing Decision Window Model, the reader is referred to Chapter 4. PODS has served as a revenue management experimental tool for the PODS consortium, a partnership between MIT and seven major international airlines.

To this end, PODS simulates millions of choices by individual air travelers flying over a network of origin-destination markets served by several airlines. More precisely, it simulates the interactions between passengers and airlines during the booking period for a single day of departure. The booking period extends over 16 successive time frames, the first time frame beginning 63 days before departure and the last ending on the departure day. After the simulation is over, it is possible to analyze the results of each airline. The simulation outputs are the results of these individual choices and include airline traffic, revenues and loads. With these outputs, researchers are able to analyze the relative performance of various revenue management strategies.

To this end, PODS runs an iterative process, performing multiple "trials" for the same departure day. This allows the airlines to progressively build the historical database they need to operate the forecasting component of their revenue management systems: initial numbers in the database are progressively replaced by passenger demand that results from the choices of thousands of individual air travelers. To be more precise, under the current default settings, each PODS simulation consists of 5 independent trials, each composed of 600 successive (and thus correlated) simulations of a departure day (also called a sample). The initial 200 samples of each trial are discarded to eliminate the initial condition effects, and the results from the 5 trials are averaged to give stable and statistically significant results. Under these settings, in the most widely used network environment (PODS Network D), simulation results are then the outcome of the simulation of 2,000 departure days or samples with about 16,000 booking requests per sample. This means that each PODS run involves in Network D the simulation of the choice process of 32,000,000 air trips.

The PODS architecture consists of four components, which are linked as shown in Figure 3.1.

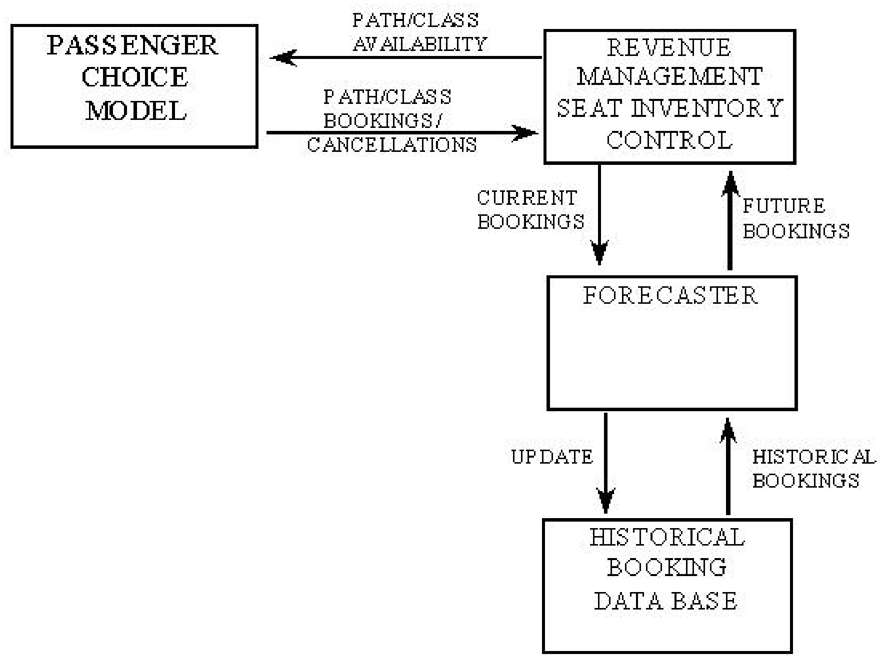


**Figure 3.1.** The PODS Architecture (courtesy of Hopperstad).

The first component is the historical database. It is created by keeping a record of booking histories starting from the first booking of the simulation. These booking histories are then used to generate forecasts for future flights. Forecasting demand is the task of the second component of the PODS simulator, i.e. the forecaster. Indeed, to set the booking limits for each fare class of future flight departures, the revenue management optimizer requires as inputs a demand forecast by fare class or path class based on the characteristics of the revenue management method used. In the following section, we describe the forecasting process for fare class demand forecasts. The process is similar for path class demand forecasts.

Demand forecasts are based on historical bookings for the same flight and fare class. There are currently four fare classes in PODS labeled Y, B, M and Q with Y being the most expensive fare. Demand forecasting consists of two steps: first, the forecaster performs detruncation of observed historical bookings. Indeed, if a fare class was closed before departure on a past flight, the bookings recorded in the historical database are a constrained observation of the actual demand for that flight/fare class. As a result, to estimate the total past demand for that flight/fare class, it is necessary to estimate what the demand would have been if there was no capacity constraint i.e. if capacity for that flight/fare class was infinite. Several detruncation methods can be used in PODS to get estimates of the unconstrained demand for a flight/fare class based on the observation of actual bookings constrained by capacity. In a second step, these estimates of the unconstrained demand by fare class of past flight departures are used to forecast unconstrained demand for future flights by fare class. Several techniques are available in PODS to perform forecasts based on these historical data. The forecasts are then transferred to the revenue management optimizer at the beginning of each time frame: they are updated 16 times during the booking process for each future flight departure to take advantage of the latest information recorded in the historical booking database. For a more complete description of forecasting and detruncation methods in PODS, the reader is referred to Zickus (1998).

Based on the remaining capacity of the aircraft (total aircraft capacity minus current bookings) and the demand forecasts, the optimizer sets booking limits for each fare class on future flights. Booking limits are nested so that all of the remaining aircraft capacity is always available for booking requests in the highest fare class. Booking requests for each fare class will be accepted up to the booking limit set by the optimizer. Each time a booking is made, booking limits are decreased by one unit for that fare class and all fare classes above due to the nested structure of fare class booking limits. As already mentioned, this booking also gets recorded in the historical database, which is constantly kept up to date. At the beginning of each time frame, booking limits are re-optimized based on updated forecasts delivered by the forecaster. For a description of the various revenue management methods, the reader should refer to Darot (2001).

Finally, the last component of the simulator is the passenger choice model. This component generates the number of booking requests for each future flight departure. Then, based on passenger characteristics and fare class booking limits, each individual air traveler will choose among all the available travel alternatives that fulfills his travel need for getting from city A to city B. But before looking in more detail into the PODS Passenger Choice Model, the next section describes the most widely used network configurations.

## 3.3. PODS Network Configurations

We will use in this thesis two different network configurations, Network D and Network E. Network D represents the US domestic market and Network E a transatlantic international alliance market. In addition, in Network D, competing airlines offer very similar schedules whereas in network E, one airline has a significant schedule advantage over its competitors.

In Network D, two airlines compete in 482 US domestic origin-destination markets. Each airline offers 3 flights a day between its own hub and 40 spoke cities all over the US as well as to its competitor's hub. All flights use 100-seat aircraft. Network D is unidirectional as all passengers travel from the West to the East, except bi-directional flights in the hub-to hub market. Each airline offers three connecting banks at their respective mid-continent hub and their schedules are quite similar as the starting time of the connecting banks are the same for both airlines. Figure 3.2. below is a map of this US domestic network:

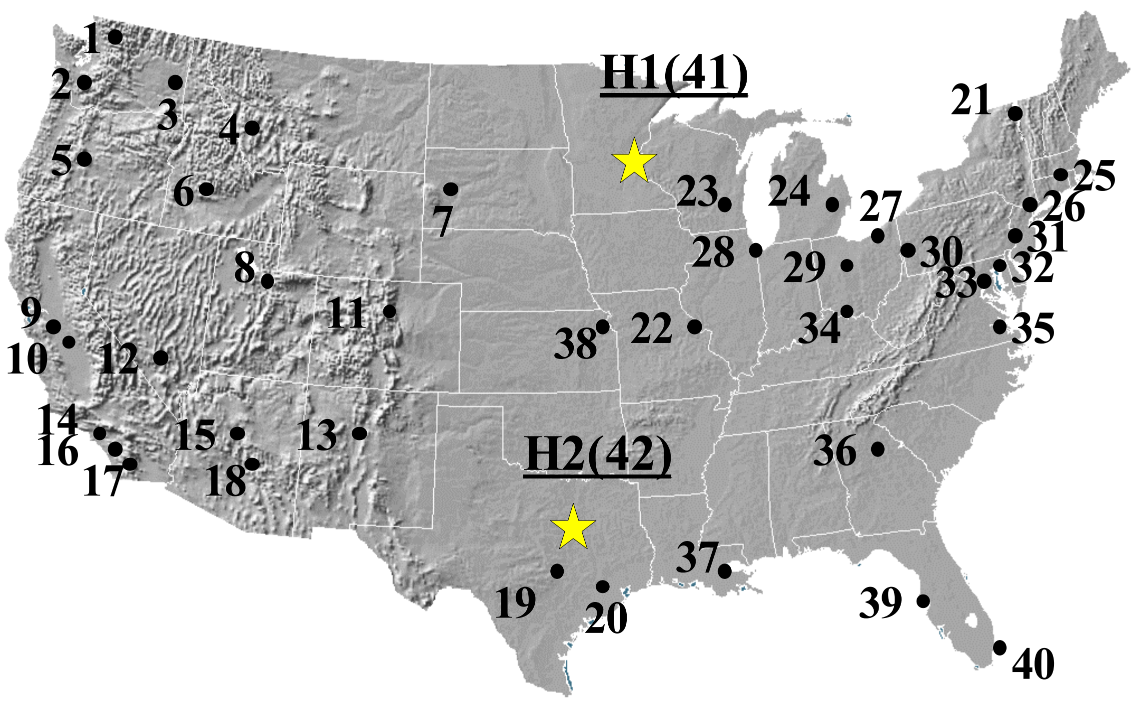


**Figure 3.2.** : Network D route map

In Network E, four airlines (two US and two European airlines) compete over their own domestic market as well as over the transatlantic market. Each airline offers flights from 10 origin cities to 10 destination cities in its own continent (from the West to the East of the US and from northern Europe to southern Europe). In addition, each airline offers transatlantic flights between its own hub and his alliance partner hub. As a result, each airline also offers transatlantic codeshare service to the 10 destination cities in its partner's continent. Airlines use aircraft of various sizes depending on the level of demand in each market with aircraft of larger size being used for transatlantic services. Figure 3.3. is a map of this transatlantic network:

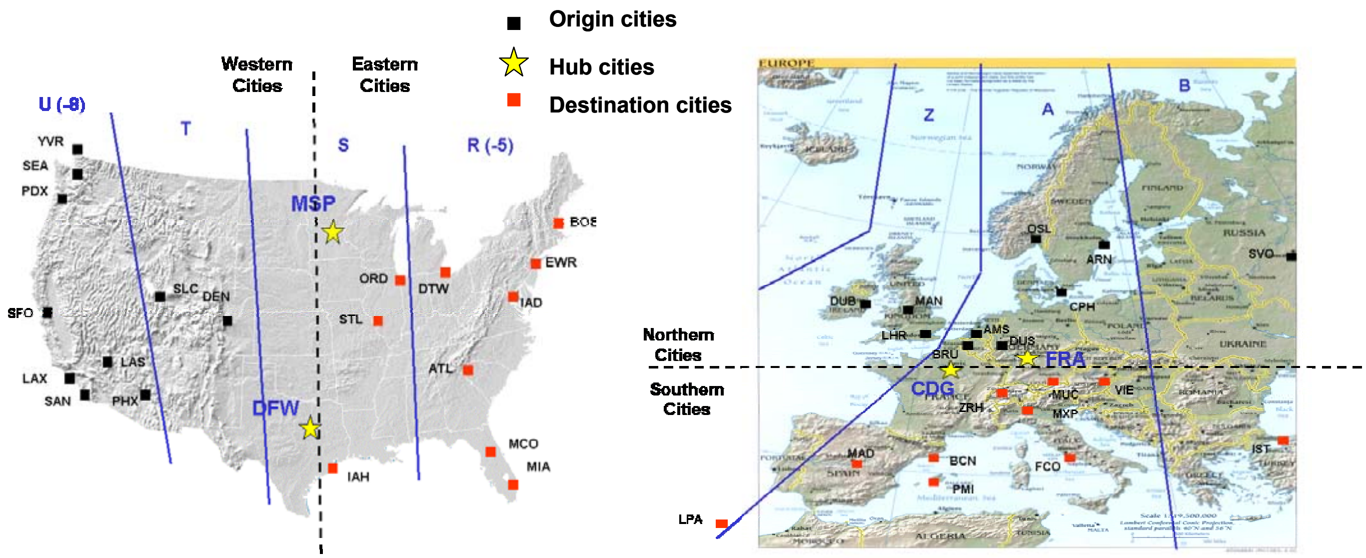


**Figure 3.3.** : Network E route map

Unlike in Network D, each airline offers quite different schedules in Network E. Like in Network D, each airline has three connecting banks with arrivals from all domestic origin cities and departures to all domestic destination cities. Two of these banks also include transatlantic flights. However, one of the US airline has a significant schedule advantage over its competitor thanks to a larger schedule coverage. European carriers have more similar schedules. Table 3.1. details the bank starting times for all four carriers:

| | US | | | | Europe | | | |
|---|---|---|---|---|---|---|---|---|
| | Starting time | Dom. | Int. Inbound | Int. Outbound | Starting time | Dom. | Int. Inbound | Int. Outbound |
| **Alliance A** (MSP-FRA) | 10 a.m. | X | X | | 7 a.m. | X | X | X |
| | 3 p.m. | X | X | X | 12 p.m. | X | X | X |
| | 8 p.m. | X | | X | 4 p.m. | X | | |
| **Alliance B** (DFW-CDG) | 12 p.m. | X | X | | 8 a.m. | X | X | X |
| | 3 p.m. | X | X | X | 12 p.m. | X | X | X |
| | 6 p.m. | X | | X | 4 p.m. | X | | |

Table 3.1. : Connecting banks in Network E

## 3.4. *The PODS Passenger Choice Model*

This section describes the various elements of the PODS Passenger Choice Model. As already mentioned, the design of the PODS Passenger Choice Model is based on previous work by Boeing, especially the Boeing Decision Window Model. It can be divided into four different steps: first demand for air travel gets generated, a set of characteristics is then defined for each individual traveler to model his preferences, then the passenger choice set is defined based on the passenger characteristics and the state of the airline inventory and, finally, each passenger makes travel decisions by matching the attributes of all available travel alternatives and his own preferences. Figure 3.4. illustrates the structure of the passenger choice model:

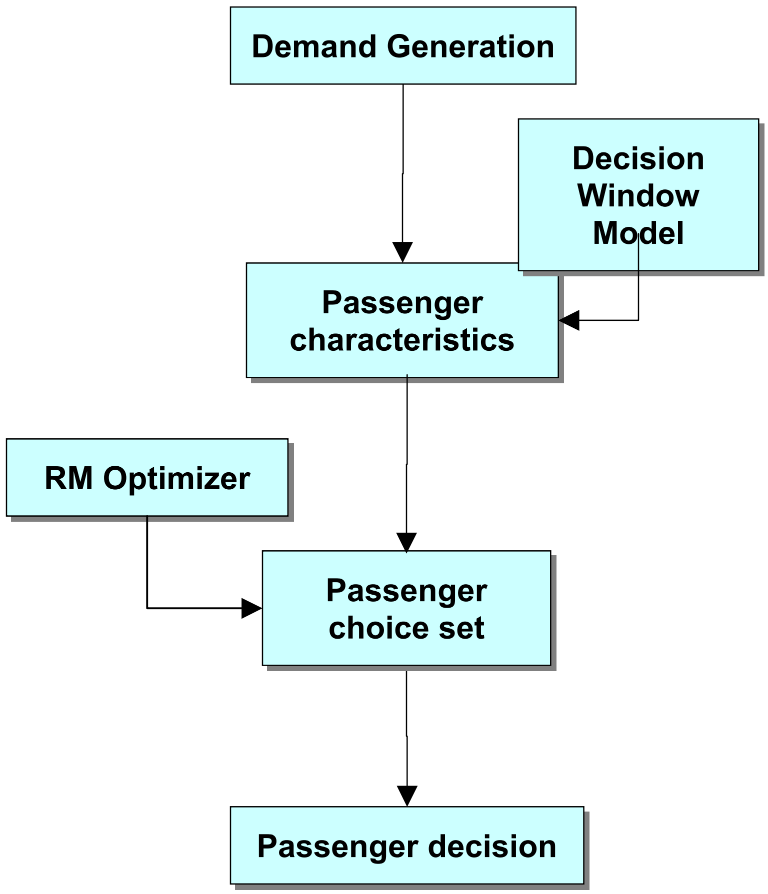


**Figure 3.4.** : The structure of the PODS Passenger Choice Model

### 3.4.1. Demand Generation

The first step of the Passenger Choice Model is to calculate the number of potential passengers who will make a booking request in a particular origin-destination market for example Los Angeles to New York. Air travelers are divided in PODS into two types based on trip purpose: leisure and business passengers are modeled separately and as we will see later on, they have different characteristics based on different behavioral assumptions. As a result, demand for air travel in each market is divided into demand for business and leisure travel.

Until recently (PODS Network D), the average demand for air travel was determined in each market through a gravity model based on the attractiveness of the origin and destination cities and the passenger mix was equally divided between business and leisure passengers. However, with the introduction of the international alliance network (PODS Network E), demand for air transportation is now based on real data provided by consortium member airlines. As a result, the split between business and leisure passengers now varies by market: in the international alliance network, some markets are largely dominated by leisure demand like for instance southern European destinations, whereas some other markets are more business-oriented.

Once the average demand for air transportation services has been determined for every O-D market in the PODS network, the simulator computes the demand for every single travel day (sample). The demand generation process does not incorporate some variability elements like seasonality or variations in the level of demand according to day of the week. However, even without taking these two elements into account, demand for air transportation will vary from one departure day (sample) to the next and some of this variation which cannot be easily explained or forecasted is modeled in the simulator.

To incorporate random deviations around the average demand, PODS follows the common industry practice of assessing a variability measure that depends on the magnitude of the mean. Two alternatives forms have been suggested to represent this stochastic variation referred to as k- and z-factors. PODS employs a combination of these two approaches to calculate the deviation from mean of the demand for air travel in a specific O-D market for a particular departure date.

Application of a k-factor supposes that the standard deviation $\sigma$ of a random variable is a constant times the mean $\mu$, $\sigma = k\mu$. From empirical analysis of airline demand data, researchers at Boeing have found that a k-value of 0.3 can be typically observed. If demand for air transportation is modeled as normally distributed, this means that 68% of the observations will be in the $\mu \pm \sigma$ range. Alternatively the z-factor approach suggests that the variation is proportional to the variance $\sigma^2$, or $\sigma^2 = z\mu$. More detail about variation in demand by departure day can be found in Wilson (1995).

Given the total number of booking requests for each departure date or sample, the allocation during the booking process must still be determined. The booking process in PODS starts 63 days before departure and is divided into 16 time frames, which are defined at the system-wide level. Time frames are initially as long as one week but their duration diminishes as the departure date becomes closer to reflect a more intense booking activity.

The allocation of the booking requests among the time frames is modeled through a booking arrival curve. These curves are different for business and leisure travelers as industry experience has shown that leisure travelers tend to book earlier than business travelers, a trend that is strengthened by the advance purchase requirements associated with lower fares (Y, B, M and Q have respectively 0, 7, 14 and 21-day advance purchase requirements in our PODS scenarios). In addition, they incorporate the impact of advance purchase requirements on the booking process with a stronger booking activity before the threshold booking dates of the lower fare-classes, especially for leisure travelers.

Figure 3.5. displays the booking curves used in all markets for business and leisure travelers:

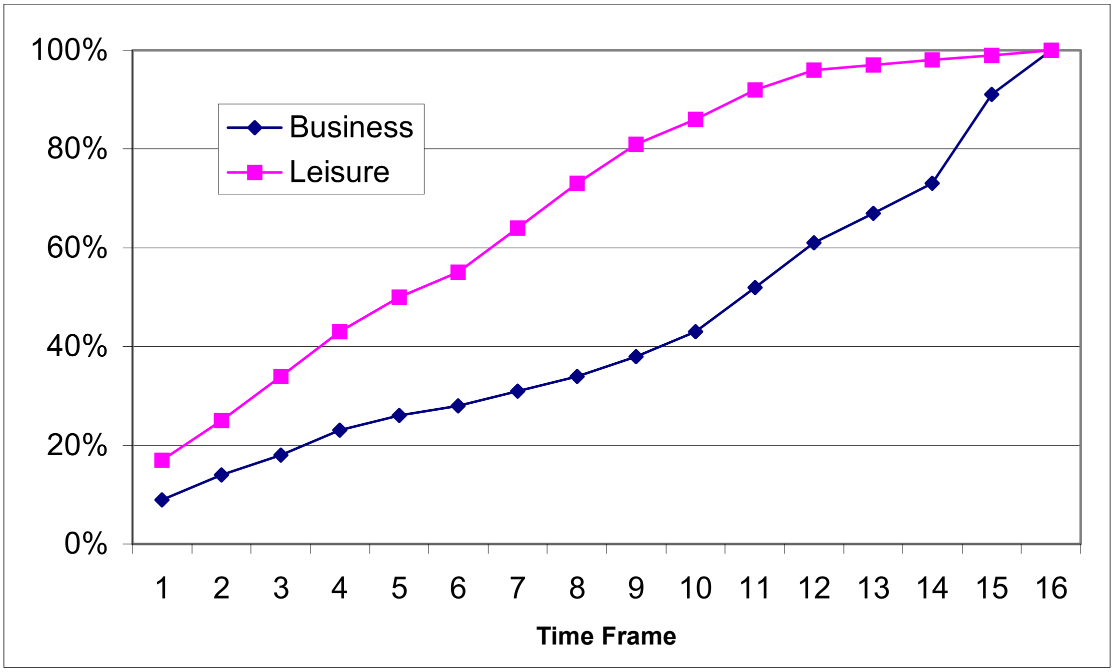


**Figure 3.5.** : PODS booking curves

Once the number of passengers of both types wishing to travel in all O-D markets on a particular date has been calculated, the PODS Passenger Choice Model must determine the characteristics of every single passenger, characteristics that will have a large influence on the final selection of a travel alternative.

### 3.4.2. Passenger Characteristics

Each passenger that intends to travel and makes a booking request in PODS gets assigned a set of characteristics to represent his preferences for various elements of the air transportation service. Based on those characteristics and the attributes of all available travel alternatives, a passenger will select a travel option. These characteristics have then a major influence on passenger choice behavior in PODS. They can be divided into three main elements: each individual traveler gets assigned a decision window, a maximum willingness to pay and a set of passenger disutilities.

#### 3.4.2.1. Decision Window

As already mentioned, the PODS Passenger Choice Model is based on the Boeing Decision Window Model. In order to represent his preferences for a particular schedule, each passenger gets assigned a time decision window. The boundaries of this time window represent respectively the earliest departure time and the latest convenient arrival time that fulfill a passenger's original schedule constraints. The concept is that most travelers' schedule plans and time constraints do not require looking for a specific single departure time. Each passenger's schedule plan and constraints can be satisfied through a range of convenient travel times that ensure a satisfactory solution to his own time-space problem.

Two parameters define a decision window: its width and its position during the day. The width of a decision window is the sum of the minimum travel time in the market and a random element called the schedule tolerance. The value of the schedule tolerance is defined randomly for each single traveler but depends on the market stage length and the passenger trip purpose. Decision windows are on average shorter for business travelers than for leisure travelers to reproduce the emphasis people traveling on business place on time and schedule. In addition, schedule tolerance is smaller in short-haul markets than in long-haul markets. The position of the decision window is defined such as to reproduce the typical time of the day distribution of demand in every market with typically peak demand in the morning and late afternoon.

All path/classes that fit into a decision window (i.e. for which departure time is after the beginning of the time window and arrival time before the end of the time window) are equally schedule-attractive to the air traveler. All other path/classes are also equally unattractive to the air traveler and the difference between the two categories is modeled through a specific cost called the replanning disutility (see below, section 3.4.2.3.).

#### 3.4.2.2. Maximum willingness to pay

In addition to his decision window, each traveler gets assigned a maximum willingness to pay. This dollar value represents the maximum amount this passenger is willing to pay for his ticket. If the fare of a path/class is above that value, the passenger will not accept to travel on that path/class and will look for other alternatives to fulfill his travel needs.

Maximum willingness to pay values are assigned randomly to passengers but there are set to reproduce willingness to pay curves designed to represent some expected passenger characteristics. The willingness to pay curves are not the same for business and leisure travelers. On average, business travelers have a much higher willingness to pay than leisure travelers. A very large proportion of them is ready to pay the most expensive fare (Y fare), if necessary. On the contrary, the willingness to pay curves are designed such that most leisure travelers will accept to pay only the most inexpensive fare (Q fare).

The equation of the maximum willingness to pay curves is the following:

P(pay at least f) = min [1, exp(- $\frac{.6931*(f - basefare)}{(emult - 1)*basefare}$ )]

Where f = the fare in question
Basefare = fare at which all potential travelers would travel
e-mult = elasticity multiplier (of the base fare where 50% of the passengers are willing to travel)

To differentiate the behavior of business and leisure travelers, the value of the basefare and the e-mult are different for the two categories. For leisure travelers the basefare is equal to the lowest fare (Q fare) but for business travelers, it is set at 2.5 times the Q fare. The e-mult is set to 1.2 and 3 for leisure and business travelers respectively.

This means that all business travelers are ready to pay a fare up to 2.5 times the lowest fare whereas all leisure travelers are willing to accept anything but that Q fare. In addition, 50% of the leisure travelers cannot afford a fare higher than 1.2 times the Q fare whereas 50% of the business travelers can afford up to 7.5 times the Q fare.

As a result, almost all business travelers are typically willing to purchase the most expensive fare but a large proportion of the leisure travelers cannot afford anything but the cheapest fare. The design of the willingness to pay curves reproduces fundamental behavioral differences between business and leisure travelers and the dichotomy between a largely inelastic demand for business travel and a very price-sensitive demand for leisure travel.

Table 3.2. describes an example of the willingness to pay curves used in PODS:

| | Probability that a random passenger will pay at least | | | |
|---|---|---|---|---|
| | Q | M | B | Y |
| Pax Type | $100 | $150 | $200 | $400 |
| Business | 100% | 100% | 100% | 93% |
| Leisure | 100% | 18% | 3% | 0% |

**Table 3.2.** : Maximum willingness to pay

### 3.4.2.3. Passenger Disutilities

In addition to a time decision window and a maximum willingness to pay, an additional set of characteristics is generated for each passenger and is used to calculate disutilities. Disutilities are used to represent additional non-monetary costs that depend on the attributes of each path/class and influence the choice of air travelers. There are two types of disutilities: the disutilities associated with the restrictions and other disutilities.

Except for the Y full fare, all lower fare-classes have restrictions: the B fare has one restriction, the M fare has two of them and the Q fare three of them. Such a fare/restriction structure is typical of the pricing structure found in the airline industry since deregulation. In PODS, each passenger gets assigned a random value for the disutility or cost associated with each restriction. The average disutility is different for business and leisure travelers. For business travelers, the restrictions disutilities are designed such that on average, the passenger will prefer an unrestricted Y full-fare, and then respectively a B, M and Q fare. This means that, on average, for a business traveler, the dollar cost of a Y fare will be less than the sum of a B fare and its restriction disutility cost.

Airlines usually design restrictions to achieve market segmentation: Restrictions are used to offer different products designed to fulfill the needs of different market segments. For instance, restrictions are designed such that business travelers cannot usually meet them and then prefer to choose the full-fare most of them are willing to pay for. For instance, lower fares require usually a Saturday night stay, a restriction most passengers traveling on business cannot fulfill as they typically travel on short trips during the week and want to return home to spend the weekend with their families. As a result, there would be little stimulation in business traffic if the lowest fares in the market were reduced because the restriction disutilities are high enough to make business travelers prefer a Y unrestricted full-fare.

On the contrary, most leisure travelers are able to meet the restrictions requirements associated with the lower fares and then in PODS, the average restriction disutilities are lower for leisure than for business travelers. On average leisure travelers will prefer the most restricted Q fare, then the M, B and finally Y full fare. In addition, these restrictions are independent and their values are not correlated for a single passenger. Table 3.3. gives an example of the average restriction disutility costs for business and leisure travelers in a market with a Q fare of $100.

| Pax Type | Fare Class | Y | B | M | Q |
|---|---|---|---|---|---|
| | Avg Res. 1 | N/A | $225.00 | $225.00 | $225.00 |
| | Avg Res. 2 | N/A | N/A | $75.00 | $75.00 |
| | Avg Res. 3 | N/A | N/A | N/A | $75.00 |
| | Fare | $400.00 | $200.00 | $150.00 | $100.00 |
| Business | **Average Total cost** | **$400.00** | **$425.00** | **$450.00** | **$475.00** |
| | Avg Res. 1 | N/A | $175.00 | $175.00 | $175.00 |
| | Avg Res. 2 | N/A | N/A | $25.00 | $25.00 |
| | Avg Res. 3 | N/A | N/A | N/A | $25.00 |
| | Fare | $400.00 | $200.00 | $150.00 | $100.00 |
| Leisure | **Average Total cost** | **$400.00** | **$375.00** | **$350.00** | **$325.00** |

**Table 3.3.** : Restriction disutility cost

In addition to the restriction disutilities, there are three other disutilities included in the PODS Passenger Choice Model: replanning disutility, path quality index and unfavorite airline. These disutilities are once again independent for a single passenger and on average higher for business travelers than for leisure travelers.

If a particular path/class is outside the decision window of a passenger, a replanning disutility is added to the cost of the path/class to take into account the inconvenience of the schedule. This cost is higher on average for business travelers that are known to place more emphasis on frequency of service and schedule but does not depend directly on the size of the decision window. For more information about passenger preference for schedule and replanning disutilities, the reader is referred to Chapter 4.

Passengers usually prefer non-stop flights to connecting paths. In order to take this into account, each path/class gets assigned a path quality index. This index is equal to 1 for non-stop flights and 2 for connecting paths. The PQI disutility of each path/class is equal to a base disutility multiplied by the value of the index. As a result, for the same passenger, the PQI disutility of a connecting path will be twice as much as that of a non-stop service.

Finally, air travelers tend to have some preferences for a particular airline, in part due to frequent flyer programs the airlines have developed over the years. The frequent flyer programs are often non-linear and target primarily business passengers that tend to give them more importance than leisure travelers. To model these preferences, each passenger gets assigned in PODS a preferred airline and a randomly drawn unfavorite airline disutility. If a path/class is not operated by its favorite airline, the unfavorite airline disutility is added to the total cost of the path. The probability $pfav_a$ that a passenger considers Airline A as his favorite carrier is determined by the airline coefficient of preference $calp_a$

$$pfav_a = \frac{calp_a}{\sum_i calp_i}$$

In PODS Network D, each airline has a coefficient of preference of 0.5, which means that 50% of the passengers will consider each airline as their favorite airline.

All disutilities are defined randomly for each air traveler. However, on average, for both business and leisure travelers, they are a linear function of the market basefare that depends to some extent on the length of haul. The intercept and the slope of the disutility functions have been calculated to reproduce some expected passenger behavior. These passenger behaviors have been defined through a survey of the marketing expertise of the member airlines. For instance, the path quality disutility is designed to decrease in relative terms with length of haul, as passengers flying to medium/long-haul destinations are supposed to be more willing to accept a connection than passengers flying to a short-haul destination.

In addition, as already mentioned, all disutilities are on average higher for business than for leisure passengers: business passengers are assumed to place more emphasis on non-monetary elements like the quality of the path, unrestricted fares, airline preference, schedule convenience than leisure passengers and to be less sensitive to price. Finally, all disutilities are assumed to be independent and to follow a normal distribution with a 0.3 k-factor, typical of air transportation demand according to marketing research conducted by Boeing. For more information on passenger disutilities, the reader should refer to Lee (2000). Table 3.4. summarizes the three latter average disutility costs for a market with a Q fare of $100.

| Pax type | Market Base fare | Replanning | Unfavorite Airline | Path quality |
|---|---|---|---|---|
| Business | $250.00 | $61.56 | $30.21 | $22.23 |
| Leisure | $100.00 | $11.90 | $9.02 | $7.01 |

**Table 3.4.** : Replanning, unfavorite airline and path quality disutility costs

### 3.4.2.4. The Passenger Choice Set

Once the simulator has generated the characteristics of a specific traveler, each traveler will make travel decisions through a two-step process: in the first step, the simulator will define the passenger choice set. In PODS network D, in most markets, up to 25 alternatives are potentially available: the air travelers have the choice among two airlines, three paths per airline, four fare classes per path (Y, B, M and Q classes) plus the no go alternative. Alternatives are also called path/classes and they are characterized by these two elements: the path (which includes the airline identity) and the fare class. The no go alternative is included in the choice set of all air travelers. However, some of the other 24 alternatives will be excluded from the passenger choice set for three reasons:

- The revenue management controls: the alternative is unavailable and will be excluded from a passenger choice set if the airline inventory indicates that no availability remains for this path/class. Availability is based on the booking limits set by the optimizer at the beginning of each time frame minus the bookings that occurred so far during the time frame.

- The advance purchase requirements associated with lower fare classes: in PODS B, M and Q classes have advance purchase requirements of respectively 21, 14 and 7 days. If a booking request occurs between 21 and 14 days before the flight departure date, only Y, B and M classes will remain available and will be included in the passenger choice set.

- The willingness to pay: PODS will exclude from a passenger choice set any path/class that has a fare higher than the passenger maximum willingness to pay.

Now that the passenger choice set has been defined, let us finally consider the decision rule used in PODS to determine which alternative will be chosen.

### 3.4.2.5. The Passenger Choice

As mentioned in the previous section, the passenger choice set contains in most markets in PODS network D between 1 and 25 alternatives depending on the passenger willingness to pay, the date of the booking request and the state of the airline inventory. Let us now consider the decision rule used to choose among path/classes that are included in a passenger choice set.

As already mentioned, the choice set of any air traveler contains at least one alternative, the no go alternative. There are then two cases: if the passenger choice set contains only one alternative i.e. the no go alternative, the passenger will "choose" not to go. However, if there are at least two alternatives in the passenger choice set, the no go alternative will never be chosen and PODS will use the following decision rule to select the most preferred alternative: the air traveler will consider additional non-monetary elements and a generalized cost is calculated for each alternative (except the no go alternative). This generalized cost is the sum of the fare and all additional disutilities that depend on:

- the characteristics of this specific air traveler (O-D market, business or leisure, time window)
- the fare/class (restrictions)
- the path (airline, quality of the path, schedule)

In order to choose among all path/classes considered, the generalized cost of each alternative is calculated by adding the fare and the sum of all six disutilities associated with the path/class for that particular passenger. The air traveler will then choose the path/class, which has the lowest generalized cost. As already mentioned business travelers tend to put more emphasis on the convenience of the path/class attributes and leisure travelers on the fare as all disutilities are on average largely higher for business travelers.

Once an air traveler has chosen a path/class, the seat availability is updated by decreasing the airline inventory by one seat on the legs traversed by the path/class. In addition, this travel decision is recorded in the historical database that feeds the forecaster and the optimizer components of the simulator used to set revenue management controls for the entire network at the end of each time frame.

As a result, the decision rule currently used in PODS makes a distinction between monetary elements (willingness to pay) and non-monetary elements (disutilities). The choice of a particular path/class is based both on monetary (fare) and non-monetary (restrictions, passenger disutilities) considerations but this choice is conditional on the maximum willingness to pay requirement. Only path/classes which fare is lower than the passenger maximum willingness to pay are effectively included in a passenger choice set.

## 3.5. Discrete Choice Models and the PODS Passenger Choice Model

In this section, we examine the relationship between the PODS Passenger Choice Model and the discrete choice models usually found in the literature to study consumer choice and described in Chapter 2. We will investigate the similarities and differences between the PODS Passenger Choice Model and respectively the logit model, the nested logit model and finally the more flexible mixed logit model.

### 3.5.1. PODS and the Logit Model

The PODS Passenger Choice Model and logit models share three fundamental common characteristics. First, both models are utility maximizing models. In PODS, passengers choose the alternative included in their choice set that has the lowest generalized cost. This is equivalent to a utility maximizing model as the generalized cost multiplied by minus 1 can be considered as a utility measure and multiplying the generalized cost by minus 1 for all alternatives does not change the outcome of the choice process for any decision-maker. Choosing the alternative that has the lowest generalized cost or the highest utility is then equivalent.

In addition, the second fundamental common characteristic between the PODS Passenger Choice Model and discrete choice models is that the number of alternatives in the consumer/passenger choice set is finite. As outlined in Chapter 2, this property defines the class of discrete choice models. Furthermore, as described in the previous section, the number of alternatives in a passenger choice set in PODS is finite as well: in network D, the passenger choice set can include up to 25 alternatives in most markets and at the maximum 49 in the more frequently served inter-hub market.

Finally, the logit model is also particularly well suited to deal with choice sets that vary in size and composition from one decision-maker to the next. The PODS Passenger Choice Model creates such a situation as we have shown in the previous section that the passenger choice set varies from one passenger to the next based on the passenger characteristics (willingness to pay, time decision window), the date of the booking request and the state of the airline inventory. The ability of the logit model to accommodate variable choice sets could prove to be very useful in that respect.

However, there are two fundamental differences between the PODS Passenger Choice Model and the logit model. First, the logit model cannot accommodate random taste variation i.e. heterogeneity in response in the population to a particular element of the decision-maker utility because random taste variation violates one of the fundamental assumptions of the logit model, the independence between the error terms across alternatives.

As we have shown in the previous section, all disutilities that are included in an alternative's cost/utility in PODS vary from one passenger to the next and are randomly drawn from independent normal distributions. Only the fare component remains constant across all passengers. As a result, the PODS Passenger Choice Model is based on random taste variation and differs fundamentally from logit models that cannot accommodate random taste variation.

To better understand the difference between PODS and logit models, let us take the example of restriction 1 that applies in PODS to all alternatives that involve a B, M or Q fare class. As mentioned earlier, the value of restriction 1 is drawn randomly for each passenger from a pre-defined normal distribution. As a result, for each passenger, the value of restriction 1 will deviate from the average of the restriction's distribution. If we use a logit model, this deviation from the mean restriction disutility will be included in the random part of the utility function since a logit specification requires all coefficients to be fixed. As a result, for the same passenger, the random part of the utility function will include that same deviation for all alternatives that involve a B, M and Q fare class: error terms cannot then be considered to be independent across alternatives. As mentioned in Chapter 2, if tastes vary with unobserved parts of the utility, then the logit model is not appropriate as the error terms will necessarily be correlated across alternatives. A logit model is then a misspecification and a more flexible model able to accommodate random taste variation is then needed to represent the passenger choice problem modeled in PODS.

In addition, the logit model assumption that error terms are independent across alternatives leads to a very specific substitution pattern: the proportional substitution pattern. Such a substitution pattern is not observed in PODS because the PODS Passenger Choice Model assumes random taste variation in some elements of the passenger utility function and we have shown that random taste variation creates correlation between error terms across alternatives.

To show the difference in substitution patterns between the PODS Passenger Choice Model and the logit model, let us take an example. Let us assume that a business passenger has three alternatives in his choice set:

- The no go alternative
- A Y fare-class on airline 1 in a path that fits within his decision window
- A Y fare-class on airline 2 in a path that also fits within his decision window

Let us further assume that the B fare class becomes available only on airline 2. We will assume in that example that the Q, M, B, and Y fares in that market are set to respectively \$100, \$150, \$200 and \$400. Table 3.5. describes the market share of all three travel alternatives (the no go alternative is never chosen since there are some other travel alternatives in the passenger choice set):

| | *PODS PCM* | | | *MNL* | | |
|---|---|---|---|---|---|---|
| | Y1 | Y2 | B2 | Y1 | Y2 | B2 |
| No | 50% | 50% | N/A | 50% | 50% | N/A |
| Yes | **40%** | **32%** | 28% | **36%** | **36%** | 28% |

**Table 3.5.** : The PODS Passenger Choice Model and logit models

This example illustrates the difference in substitution patterns between the PODS Passenger Choice Model and a multinomial logit model: in the logit model, the proportion of passengers shifting to the new alternative – a B fare class on airline 2 – comes in similar proportion from passengers that chose initially airline 1 and 2. However, in PODS, the passengers shifting to the B class come primarily from passengers already traveling on airline 2. The PODS Passenger Choice Model exhibits a more realistic substitution pattern as one would expect passengers to be more willing to shift to a different fare class on the same airline than to modify both fare class and carrier identity at the same time. As a result, the use of a multinomial logit approach in PODS would under-estimate airline 1 market share in such a situation.

In fact, the substitution pattern of the PODS Passenger Choice Model reflects the correlation between the Y and B fare class alternatives offered by the same carrier, airline 2. This substitution pattern reflects the underlying assumption that the utility of a Y and a B fare class on the same airline are correlated as they share some common elements associated with airline 2 characteristics like for instance the value a passenger gives to his preference for traveling on a particular airline. As a result, such a substitution pattern violates the logit fundamental assumption of independence in error terms across alternatives. We will see in the next section how a nested logit model is able to accommodate such correlation and substitution patterns.

### 3.5.2. PODS and the GEV Family of Models

As shown in Chapter 2, the GEV family is a class of models that enables the researcher to accommodate more complex substitution patterns than the logit model and keeps part of the mathematical simplicity of logit, especially a closed-form expression for the choice probabilities. The most widely used GEV family model is the nested logit model described in Chapter 2. Let us examine the relationship between the nested logit model and the PODS Passenger Choice Model.

The nested logit model takes into account the correlation between the utilities of some alternatives by grouping them into nests. In the nested logit model, correlation is allowed between alternatives that belong to the same nest but not between alternatives that belong to different nests. Referring to the example above, the nested logit model could accommodate such a correlation pattern by grouping into a nest the Y and B fare class alternatives offered by airline 2. However, this example is only a partial description of the complex substitution/correlation pattern that exists in PODS. Indeed, the passenger choice problem in PODS implies a three-dimensional choice situation: the choice of an airline, a flight schedule and a fare class. Some correlation might exist between two fare classes on the same flight offered by the same airline and they should be grouped into a nest but there might also be some correlation between two different flights of the same airline for the same fare class. So they also should be grouped into another nest. If a travel alternative can belong to only one nest, it is then impossible to represent such a correlation/substitution pattern. In the nested logit model, alternatives belong only to a single nest so the nested logit model is not able to accommodate the complexity of the choice situation modeled in PODS, which involves more complex correlation patterns.

In fact, the nested logit model requires the researcher to establish a hierarchy between the various levels of the choice problem and only select correlation patterns are allowed at the lower level of the hierarchy that can be symbolized by a tree like in Chapter 2. Independence is assumed between nests at all nodes at the upper level of the tree. In PODS, no hierarchy is assumed between the three levels of the passenger choice problem: passengers do not choose an airline first and then a flight schedule and then a fare class. In PODS, correlation might exist between the utility of any combination of alternatives. As a result, a nested logit approach is not adapted to represent the passenger choice problem as it is modeled in PODS.

In addition, if the GEV family of models is able to accommodate more complex substitution patterns than the logit model, it cannot accommodate random taste variation and all coefficients of a decision-maker utility have also to remain constant. Only more flexible models like the mixed logit model can accommodate random coefficients to represent the heterogeneity in response in the population to some elements of the decision-maker utility. Let us then examine the relationship between the PODS Passenger Choice Model and mixed logit models.

### 3.5.3. PODS and Mixed Logit Models

As already mentioned in Chapter 2, mixed logit models are a highly flexible class of discrete choice models. They can accommodate any form of substitution patterns and are particularly amenable to incorporate random taste variation. Mixed logit models are defined on the basis of the functional form of the choice probabilities: mixed logit choice probabilities are the integral of standard logit probabilities over a density of parameters. Using a mixed logit specification to represent random taste variation is straightforward. The utility specification is the same as for standard logit except that the parameters are supposed to vary across decision-makers rather than being fixed (the parameters are random variables). The researcher has then to specify a distribution for each coefficient of the systematic utility and estimate the parameters of this distribution.

In the case of the PODS Passenger Choice Model, as already mentioned all disutility costs are assumed to be normally distributed except for the fare coefficient that is specified to be fixed. In addition, all the disutility costs are assumed to be independent. As a result, the joint distribution of all disutility costs is the product of six normal distributions. Then, to turn the PODS Passenger Choice Model into a mixed logit specification only requires that the utility of each alternative given the disutility costs take a standard logit form. This is achieved if an iid extreme value term is added to the utility of each alternative. However, as already mentioned in Chapter 2, by scaling up the utilities appropriately, the researcher can ensure that adding an extreme value term to the true utility will never change the outcome of the choice process for any decision-maker or air traveler.

As a result, the PODS Passenger Choice Model can be considered as the equivalent of a mixed logit model with the mixing distribution equal to the product of six independent normal distributions, one for each of the disutility cost coefficients. As the disutility cost distributions depend on the passenger type and the basefare of each O-D market, the current PODS Passenger Choice Model can in network D be approximated by 964 mixed logit choice models (482 O-D markets, two passenger types per market).

The PODS Passenger Choice Model can then be approximated by a set of mixed logit models. This result is consistent with the fact that mixed logit choice models can approximate any utility maximizing model: as already mentioned, the PODS Passenger Choice Model is actually a utility maximizing model. The use of a mixed logit model to approximate the PODS Passenger Choice Model is especially straightforward as mixed logit models are extremely convenient for random coefficient choice problems like the passenger choice problem modeled in PODS.

## *3.6. Conclusion*

In this chapter, we have described first the general architecture of the PODS simulator and then in detail one of its four major components, the PODS Passenger Choice Model. Finally, we have shown the similarities and differences between the choice model used in PODS and the models generally found in the consumer choice theory literature. Combined with Chapter 2, we have provided in this chapter a general framework to understand the PODS Passenger Choice Model as a whole and establish that it can be approximated by a series of mixed logit models with a random coefficient specification. In the next two chapters, we will focus in more detail on one of the component of the Passenger Choice Model: How passenger preference for flight schedules is modeled in PODS. In a first part, we will concentrate on a more detailed description of the decision window model used in PODS and compare it to alternative solutions found in the literature. In a second step, we will use the simulator to test the impact of alternative approaches to model passenger preference for schedule in PODS.

# Chapter 4 Passenger Preference for Schedule

## 4.1. Introduction

In the previous chapters, we have described the general structure of the PODS Passenger Choice Model and its relationship with the discrete choice models found in the consumer choice theory literature. In this chapter, we are going to focus in more detail on how passenger preference for schedule is modeled in PODS and compare the PODS approach to the literature on traveler preference for schedule.

The development of low-cost competition in the United States and Europe in the recent years has become a growing challenge for full-service network carriers. As a result, network carriers need to focus on their strengths including network coverage and frequency of service. These industry trends have raised the interest for a review of how preference for schedule is modeled in PODS and an investigation of the impact of schedule asymmetry on PODS simulation results.

In the first part of this chapter, we will review the literature on how to model preference for schedule in inter-city travel in general and air transportation in particular. In the subsequent section, we will take a closer look at how preference for schedule is modeled in PODS including a more detailed description of the Boeing Decision Window Model with its strengths and shortcomings. Based on the literature review and the current PODS approach, we will finally propose alternative approaches to model passenger preference for schedule in PODS.

## 4.2. Passenger Preference for Schedule: Literature Review

The demand for air travel is a derived demand that reflects travelers' need to participate in activities at their destination. As a result, the scheduling of these activities determine to a great extent a traveler's preferred departure and arrival time. All else being equal, a passenger will choose the flight departure that offers the best solution to his own individual time-space problem, i.e. participate to some activities at his destination. From an airline perspective, the objective is to design a flight schedule that accommodates the departure time preferences of the largest possible number of travelers based on the distribution of demand by time of the day observed in each market subject to some constraints like for instance the size of their fleet or aircraft rotations. Schedule convenience is an essential part of the choice by potential passengers of a particular air travel itinerary and the models used to represent passenger preference for schedule are an essential component of any attempt to simulate air traveler choice among various travel alternatives.

The review of the literature on passenger preference for schedule in the air travel industry reveals two different approaches to model schedule convenience: the schedule delay vs. the decision window models. In the air travel literature, schedule delay has been defined as a measure of schedule convenience related to the difference between a passenger ideal departure time and his actual flight departure time (Douglas and Miller, 1974). A similar concept has been used in empirical studies to estimate travelers' sensitivity to the average time between scheduled departures (Morrison and Winston, 1985, 1986).

On the other hand, another approach found in the literature challenges the assumption that air travelers may have a unique ideal departure time. Flexibility in the schedule of their activities at destination implies that a range of departure times is a convenient solution to their time-space problem. This approach has been developed by the Boeing Airplane Company in the Boeing Decision Window Model (Boeing Commercial Airplane Group, 1997). According to Boeing research on passenger behavior, each individual air traveler does not have a single ideal departure time but a decision window. The passenger decision window represents the time frame that the traveler considers convenient for travel. It is bounded by the passenger earliest convenient departure time and latest convenient arrival time. Any flight schedule that fits entirely within the decision window is equivalent to the air traveler from a scheduling point of view. Before describing in more detail in the next section the Boeing Decision Window Model and its influence on how passenger preference for schedule is modeled in PODS, let us examine how passenger preference to schedule is applied in two recent studies on passenger choice in air transportation.

As already mentioned in Chapter 2, Prossaloglou and Koppelman (1999) investigate the choice of a carrier, a flight and a fare-class. Based on stated preference data collected through a survey, they use a logit model to estimate the impact of several factors like fare or carrier identity on the selection of a particular travel alternative by individual air travelers. For more detail about the general framework and results of their study, the reader is referred to Chapter 2, Section 5.

Regarding more specifically the choice of a flight schedule, the authors use a schedule delay approach and include a schedule delay variable in the passenger utility function. They tested first a linear schedule delay model where the schedule delay is equal to the difference between a passenger ideal departure time obtained from the survey and the actual flight departure time. Their findings indicate that business travelers are more reluctant to deviate from their ideal departure time than leisure travelers. The difference in the schedule delay coefficient by trip purpose reflects business travelers' expected greater sensitivity to schedule delay. Based on the values of the fare and schedule delay coefficients included in the model, they estimated that the average values of one hour of schedule delay were $60 and $17 for business and leisure travelers respectively. These findings support the usual segmentation of air travel demand between time-sensitive business travelers and less time-sensitive leisure travelers.

In addition, they further explored traveler's sensitivity to schedule delay by using more complex non-linear formulations of the schedule delay function. Under this more complex approach, they collected data for each passenger on both an ideal departure time and a decision window representing non-ideal but convenient and acceptable departure times. As expected, these models indicate a greater sensitivity to schedule delays associated with flights that depart outside the decision window, suggesting a non-linear sensitivity to schedule delays.

In his doctoral dissertation, Mehndiratta (1996) studies the impact of time of day preferences on the scheduling of business trips in the domestic US focusing mainly on trips involving air transportation. His study includes an exploratory survey of recent business trips by a small group of ten San-Francisco based professionals. One of the conclusions from the interviews with these business travelers is that there was always a flight that fulfilled their travel schedule preferences, even in markets where only a few travel alternatives were available, sometimes only a single non-stop flight. That suggests that business travelers adapt their travel plans to the travel schedules offered by the airlines and supports the assumption that air travelers do not have a specific ideal departure time but a range of preferred travel times. Most of the business travelers included in this sample actually had some flexibility in the design of their schedule at their destination. Mehndiratta states that "there was substantial evidence suggesting that supply and destination related constraints were non-binding in a majority of cases". In addition, when asked to state when they would have preferred to depart under ideal circumstances, most respondents reported times close to the times when flights were actually scheduled. These results support the decision window concept: Most business travelers do not have in mind a unique ideal departure time but a range of convenient travel schedules reflecting their ability to adapt their schedule at destination to the flight schedules offered by the airlines. In that perspective, each traveler decision window can be viewed as the result of the passenger time preferences mitigated by his time constraints like flying time, meeting schedule at destination. The time decision window can be viewed as the traveling schedule that minimizes the disutility associated with the disruption of the traveler's regular schedule under some constraints (work schedule at destination etc.)

In addition, Mehndiratta focuses in his work on the difference between the valuation of time across different periods of the day. As already mentioned in Chapter 2, he divided a regular 24-hour schedule into three periods: work, leisure and sleep time. He proposed and formulated a theory to accommodate variations in the value of time among these three periods of the day. His work suggests that deviating from the passenger decision window can be costly to the traveler, especially if this involves a disruption of some activities like sleeping or spending leisure time at home. This suggests that time matters and that different travel alternatives outside the decision window are very unlikely to be equivalent to the business traveler from a scheduling point of view. This study could have several implications on how we can view preference for schedule in PODS. But before discussing the impact of Mehndiratta's results on the approach to scheduling in PODS, let us describe in the next section in more detail the current approach to schedule convenience in PODS.

## 4.3. Passenger Preference for Schedule in PODS

### 4.3.1. The Boeing Decision Window Model

As already mentioned in Chapter 3, the PODS Passenger Choice Model has been developed as an extension of the Boeing Decision Window Model. As a result, the approach used in PODS to model passenger preference for schedule is directly inspired by the options developed in the Boeing Decision Window Model. Before coming back to the PODS Passenger Choice Model and its schedule component, let us then first examine in more detail its foundations, i.e. the Boeing Decision Window Model. This section is based on *Decision Window Path Preference Methodology Description*, The Boeing Commercial Airplane Group (1997).

The Decision Window Model was developed at Boeing originally as a scheduling decision support tool for the airlines. The objective of the Decision Window Model is to assist an airline in designing attractive schedules in a particular origin-destination market in order to maximize market coverage and market share. Thanks to the Decision Window Model, airlines can assess the potential impact of alternative schedules on their own and competitor's loads. This tool is designed to help airlines build schedules that are attractive to the demand and increase the number of people considering traveling (market coverage) and their own market share.

As a result, to determine each airline market share, the Decision Window Model requires modeling the choice of individual passengers among various travel alternatives or paths based primarily on one characteristic: the schedule of each path. The approach used by Boeing to model this choice process is based on the concept of a decision window. As already mentioned, each passenger gets assigned a decision window that represents his range of convenient travel times, given his own time-space problem. In designing the decision window model, Boeing modelers assumed that individual passengers do not have in mind a single ideal departure time to fulfill their travel needs but that a range of travel times are convenient to them thanks to some flexibility in their schedule plans at their destination. As a result, any path that wholly fits within a passenger decision window is equally attractive to the passenger from a solely schedule perspective.

The first step of the model is then to define for each passenger a decision window. As already mentioned in Chapter 3, a decision window is characterized by two parameters: its width and its location. The width of a decision window is the sum of the minimum travel time in the market and a random element called the schedule tolerance. Schedule tolerance represents the amount of flexibility a traveler has. The value of the schedule tolerance is defined randomly for each single traveler and varies from one passenger to the next but depends on the market stage length and the passenger trip purpose. Indeed, decision windows are on average shorter for business travelers than for leisure travelers to reproduce the emphasis people traveling on business place on time and schedule. In addition, schedule tolerance is smaller in short-haul markets than in long-haul markets. The location of the decision window is defined such as to reproduce the typical demand distribution during the day in every market with for instance the usual peaks in demand in the morning and late afternoon.

The second step of the model is to define each passenger choice set. Only paths that fit completely within the boundaries of a passenger decision window will be considered and included in the passenger choice set. In addition, only the best service for each airline is included in the passenger choice set. For instance, if airline A has two paths that fit within the passenger decision window, one non-stop path and one connecting path, only the non-stop path will be included in the passenger choice set as the model assumes that the passenger will always choose that path as they are equivalent from a schedule perspective but the non-stop service is preferred from a path quality point of view. Then, the last step of the model is the decision rule and the passenger choice. There are three cases:

- No path fits within the passenger decision window. The passenger must then re-plan his trip: a new decision window will be generated and a new decision process will be initiated
- If only one path fits within the passenger decision window, the passenger will choose that path
- If several paths fit within the decision window and are then equally attractive to the passenger from a schedule perspective, the passenger will select a path based on the carrier identity and the quality of the path. For more information about this choice process, the reader is referred to Boeing Commercial Airplane Group (1997)

After this description of the Decision Window Model, let us now examine the similarities and differences between the Boeing Model and the PODS Passenger Choice Model focusing in particular on the approach to passenger preference for schedule.

### 4.3.2. Comparative Analysis of the PODS Passenger Choice and the Decision Window Models

The PODS Passenger Choice Model is an extension of the Boeing Decision Window Model: It shares a large number of similarities with the Boeing Model. However, additional elements beyond schedule convenience, path quality and airline identity are taken into account in PODS like for instance fare and fare class restrictions. In addition, the PODS Passenger Choice Model uses at the last step a utility-maximizing decision rule based on the calculation for each path of a generalized cost.

But strictly from a schedule perspective, there is one fundamental difference between the Boeing Decision Window Model and PODS. This difference involves the definition of the passenger choice set. In the Boeing Decision Window Model, only paths that fit within the passenger decision window are included in the choice set and if there is none, a new decision window is defined and only paths that fit within this new window are considered. In PODS, all paths are included in the passenger choice set, whether or not they fit into the passenger decision window and the difference between the paths inside the decision window and outside the decision window is the replanning disutility. For more information on replanning disutilities, see Chapter 3.

As a result, in both the Boeing Decision Window Model and PODS, all paths that fit within a passenger decision window are equally attractive from a schedule point of view. However, in PODS, all paths that do not fit within a passenger decision window are also equally attractive to the passenger from a schedule perspective. The replanning disutility is the same for all paths outside the decision window whatever their position might be, whether there is a slight violation of the decision window boundaries or the path is completely outside the decision window. In the Boeing Decision Window Model, paths outside a passenger decision window are only considered if no path fits within a passenger initial decision window and the passenger needs to re-plan. Even, when replanning occurs, all paths are not considered equally schedule-attractive as only paths that fit within the new decision window are now considered.

Both the PODS and the Boeing Decision Window Model approaches might have drawbacks: The Decision Window Model approach is justified within the objectives of this model that focuses mainly on the influence of airline scheduling decisions on airline market share. However, it might not be appropriate within the PODS framework that goes beyond dealing only with airline scheduling decisions and is focused on airline revenue management decisions and takes into consideration additional elements of the passenger choice problem, especially fare and fare class restrictions. As a result, it seems reasonable that passengers would consider paths that do not fit within the boundaries of their original schedule plans if they are associated with some other substantial benefits like a lower fare or fewer restrictions.

In PODS, additional paths that do not fit wholly within a passenger decision window should then be included in a passenger choice set. However, it might not be reasonable to assume that the schedule of all paths outside the decision window is equally attractive to the air traveler. The approach currently used in PODS, which we will call the constant replanning disutility model might lead to some unrealistic decisions, especially for business travelers that are assumed to be more time-sensitive than leisure travelers.

For instance, it seems really unlikely that a business traveler that needs to travel in the morning, attend a meeting in the destination city during the day and come back in the evening would accept an evening departure. He might just decide not to travel at all if no path is available in the morning. In addition, it is also unlikely that such a business passenger would give the same value to a midday and evening departure. This problem is much more an issue for business travelers that have short decision windows and put a lot of emphasis on schedule convenience than for leisure travelers, which have wider decision windows that might cover in some cases the entire length of the day.

In fact, under the current PODS approach, because passengers consider all paths that are both inside and outside their decision windows, it can happen that a business traveler with a 6 a.m.-12 p.m. decision window prefers a B fare on an evening flight completely outside its decision window over a Y fare on the morning flight that would fit into it because the fare difference is higher than the sum of the restriction and replanning disutilities.

For instance, let us consider a business passenger willing to pay the Y fare (under the current PODS default settings, 93% of business passengers are willing to do so) and that has the following choice set: either a Y fare on the morning non-stop path that fits into his decision window or a B and Y fare on an evening non-stop flight outside his decision window. Both paths are operated by the same airline. The passenger will never choose the Y/evening combination as the cost of this path/class is always higher than the Y/morning path/class. As a result, his choice set is reduced to the following alternatives: Y/morning or B/evening. Under the current PODS design, the probability of choosing the B/evening path/class is between 10 and 15% for values of the market Q fare varying between $85 and $268 (Q fares in PODS network D and E are within that range in most markets). Such a figure might seem too high and unrealistic for a path that is located completely outside a passenger decision window but also too low for a path that deviates hardly from one of the decision window boundaries.

As a result, the constant replanning disutility model currently used in PODS might under-estimate the importance of schedule convenience, especially for business travelers. However, before going further into designing and testing alternatives to the constant replanning disutility model, let us examine the potential impact of schedule convenience on PODS simulation results. In order to assess the importance of passenger preference for schedule in PODS, let us do a sensitivity analysis of PODS simulation results with regard to the level of the replanning disutility for both business and leisure passengers

### 4.3.3. Sensitivity Analysis of PODS Simulation Results

In order to evaluate the impact of the schedule component of the PODS Passenger Choice Model on simulation results, let us do a sensitivity analysis of PODS results with regard to the value of the replanning disutility. For this analysis, we will consider both the schedule-symmetric Network D and the more schedule-asymmetric international alliance Network E. For a more complete description of the network characteristics, the reader is referred to Chapter 3. However, as a reminder, airline 1 has a small schedule advantage over its competitor in symmetric Network D thanks to a better geographical location of its hub with regard to the bulk of traffic flows. In less symmetric Network E, airline 1 has a large schedule advantage over its US competitor thanks to a wider schedule coverage.

For both networks, we will vary the replanning disutility from 0 to 200% of the current base replanning disutility with 25% increments. In addition, we also tested an extremely high replanning disutility called here the infinite case: under this assumption, like in the original Boeing Decision Window Model, only paths that fit within a passenger decision window are considered. The cost of all paths outside the decision window is so high that a potential passenger will never choose such a path, except if there are no paths in his choice set that fit in his decision window. Figures 4.1. and 4.2. below show the results of the PODS simulations carried out for this sensitivity study.

The simulation results reveal common characteristics for both networks: in both Network D and E, a decrease in the value of the replanning disutility leads to a decrease in revenues for the schedule dominant airline and an increase in revenues for other airlines. Similarly, an increase in the value of the replanning disutility leads to an increase in revenues for the schedule dominant airline and a decrease for the other carriers. This is consistent with our expectations as the schedule advantage of an airline has more impact as the influence of schedule convenience on passenger choice increases. When all replanning disutilities are set to 0 and schedule convenience is supposed to have no impact on passenger choice, airline revenues tend to converge if all airlines use the same revenue management method.

However, the impact of a variation in the replanning disutility is much greater in schedule-asymmetric network E than in schedule-symmetric network D. While significant, the impact of a change in replanning disutility value on airline revenues remain relatively limited in network D. However, the level of the replanning disutility has a dramatic impact on airline revenues in the more complex and less symmetric network E. Its impact is multiplied by at least a factor of 3 compared to network D.

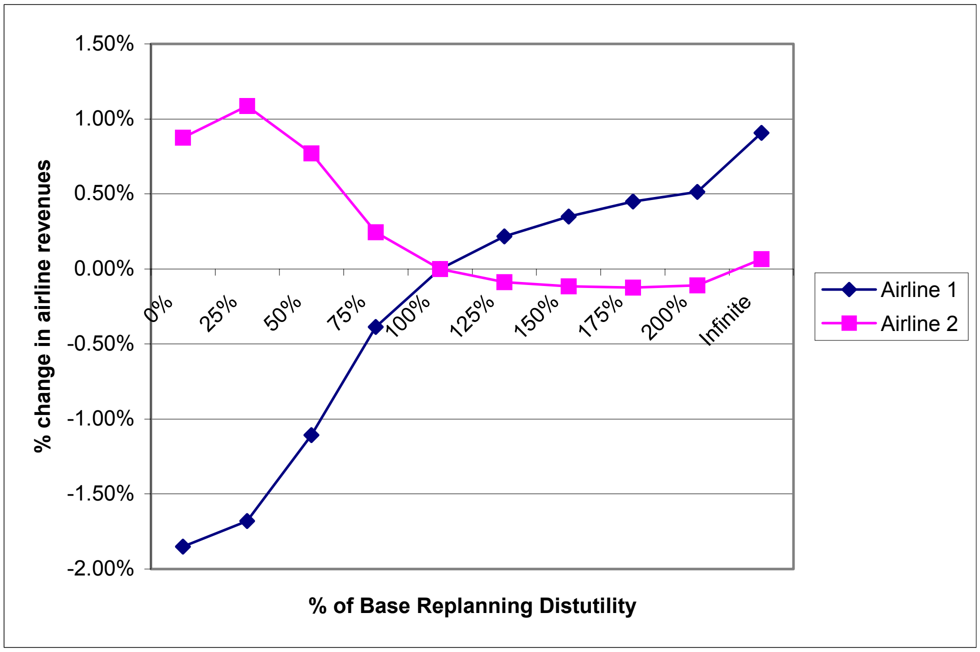


**Figure 4.1.** : Sensitivity of airline revenues w.r.t. replanning disutility in PODS Network D

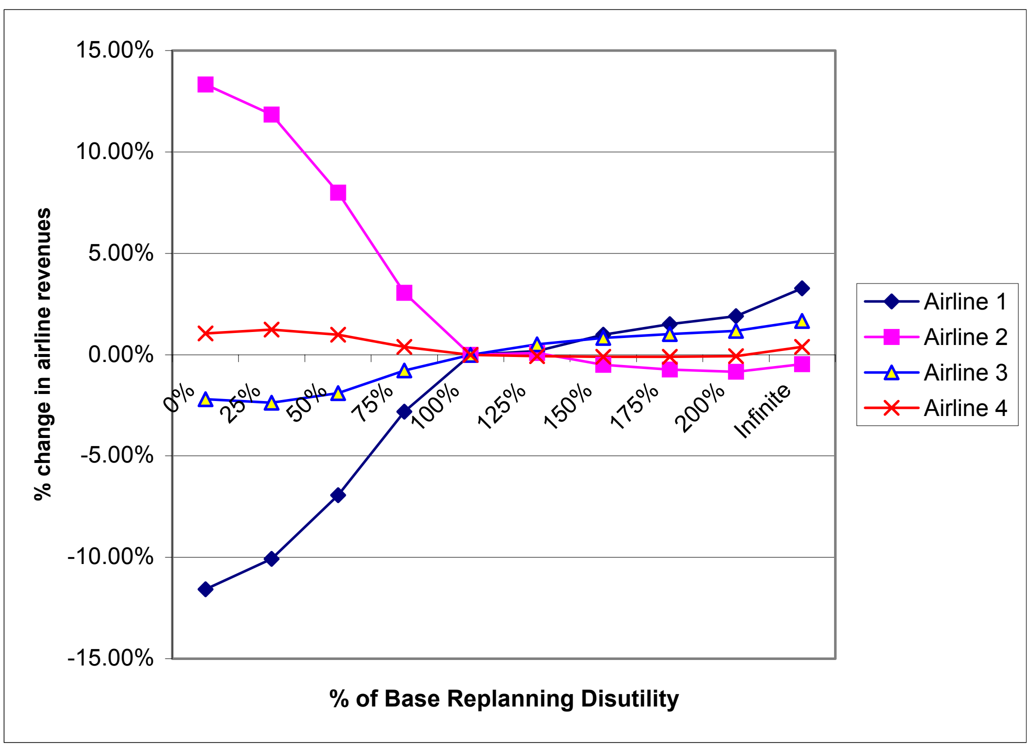


**Figure 4.2.** : Sensitivity of airline revenues w.r.t. replanning disutility in PODS Network E

Similarly, Figures 4.3. and 4.4. show that varying the replanning disutility has a different impact on load factors in Network D and E. In Network D, changes in load factor are relatively limited for both airlines: If the replanning disutility is very low, the schedules offered by both carriers are more attractive to passengers, load factors are higher and are similar for both airlines. However, when schedule convenience has more impact on passenger choice, airline 1 load factor starts increasing and becomes higher than airline 2 load factor due to the slight schedule advantage related to the geographical location of airline 1 hub.

In network E, if we set the replanning disutility to zero, airline 2 has a much higher load factor as it is as attractive as airline 1 from a schedule perspective but offers less capacity as aircraft capacity has been initially calibrated based on the base value of the replanning disutility and not on a zero replanning disutility. However, as passengers give a higher value to schedule convenience and the replanning disutility increases, airline 1 and 2 load factors follow two very different patterns: airline 1 load factor increases rapidly thanks to its attractiveness to schedule-conscious travelers due to its wider schedule coverage and airline 2 load factor decreases rapidly. Airline 3 and 4 load factors experience more moderate change that can be mainly explained by the evolution of their US codeshare partner load factor.

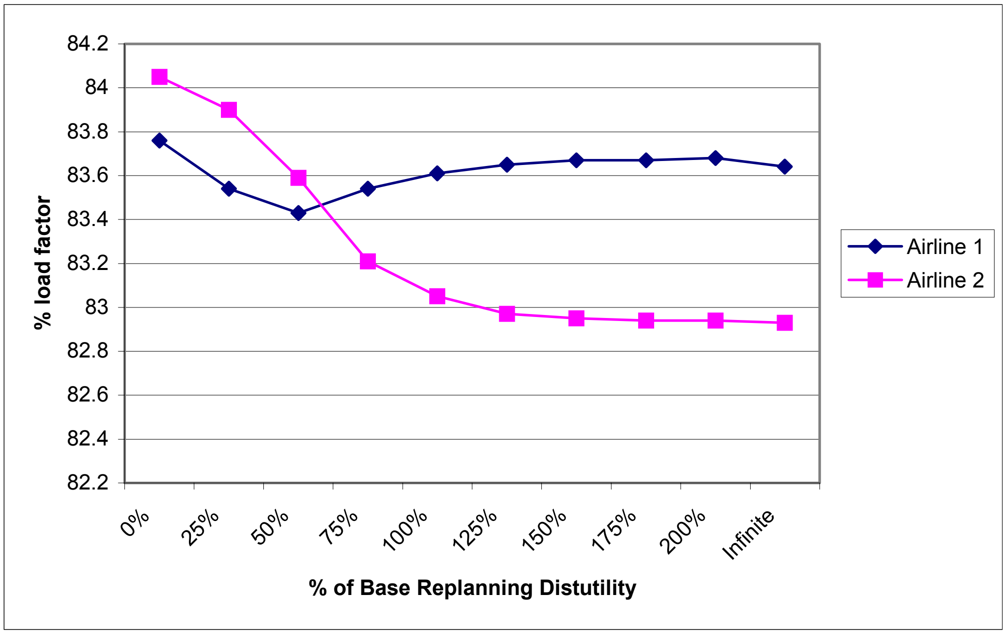


**Figure 4.3.** : Sensitivity of airline load factor w.r.t. replanning disutility in PODS Network D

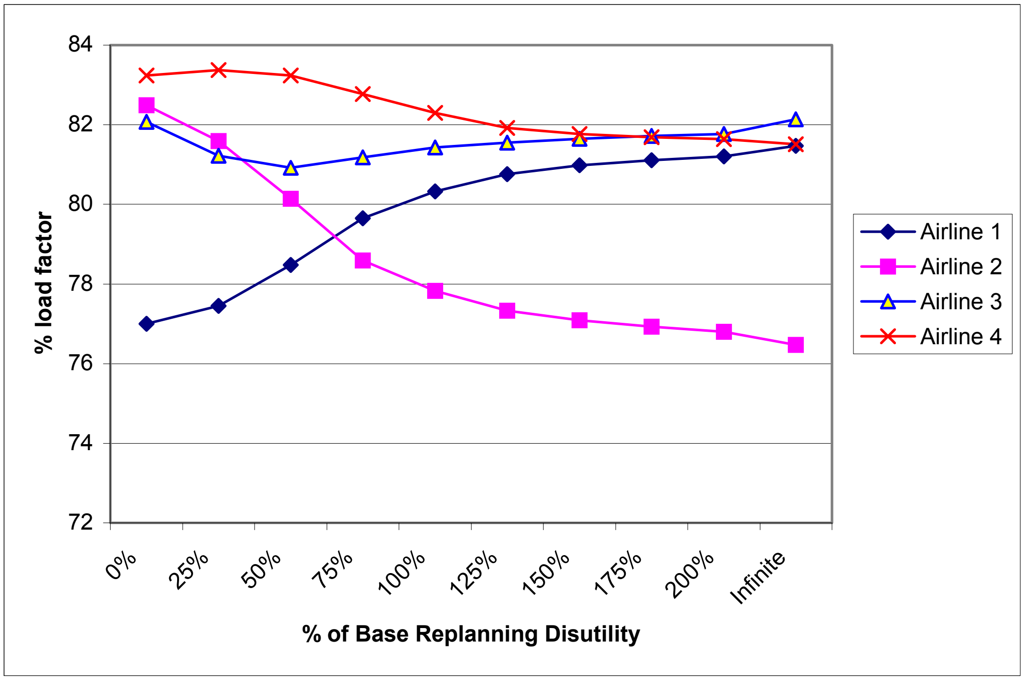


**Figure 4.4.** : Sensitivity of airline load factor w.r.t. replanning disutility in PODS Network E

To get a better understanding of the differences between Network D and E, let us examine the revenue per category figures. In PODS, we collect data, on passenger choice behavior when their initial most preferred travel alternative is no longer available. These data are called actual choice given first choice data. The first choice of a passenger is defined as the alternative with the lowest generalized cost that meets the passenger willingness to pay and the advance purchase requirements. If this alternative is available, we say that the passenger got his first choice satisfied. If it is not available due to the revenue management controls, the passenger has the four following options: Travel on the same path but on a higher fare class (sell-up), travel on the same airline but on a different path (recapture), travel on another airline (spill) or decide not to travel at all.

Unlike in the real world, since PODS is a simulation of a booking process, it is possible to track passenger behavior when denied booking of their first choice. As a result, it is possible to calculate in PODS the proportion of revenues that comes from passengers that had their first choice satisfied or were denied booking. We divide the airline total revenues into the following four categories:

- First choice revenues are revenues from all passengers that had their first choice satisfied
- Sell-up revenues are revenues from passengers that were denied booking of their first choice and decided to sell-up to a higher fare class on the same path
- Recapture revenues are revenues from passengers that were denied booking of their first choice and decided to shift to another path of the same airline
- Spill-in are revenues from passengers that were denied booking of their first choice on another airline and decided to shift to a path offered by this airline

Figures 4.5. to 4.8. display the revenues by category for the two airlines competing in PODS Network D:

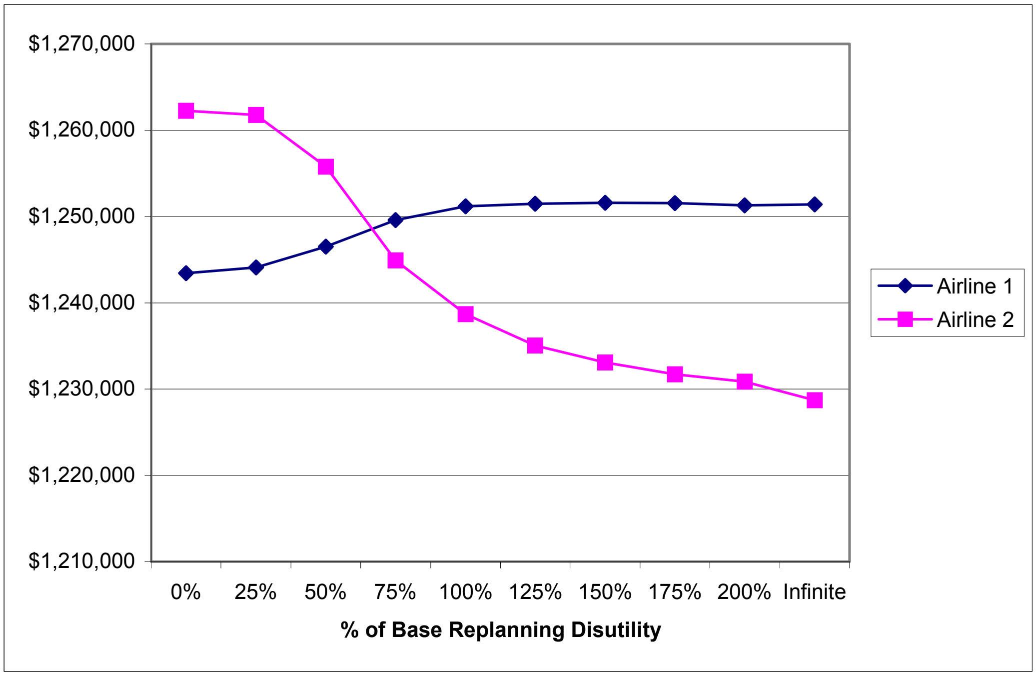


**Figure 4.5.** : First choice revenues in Network D

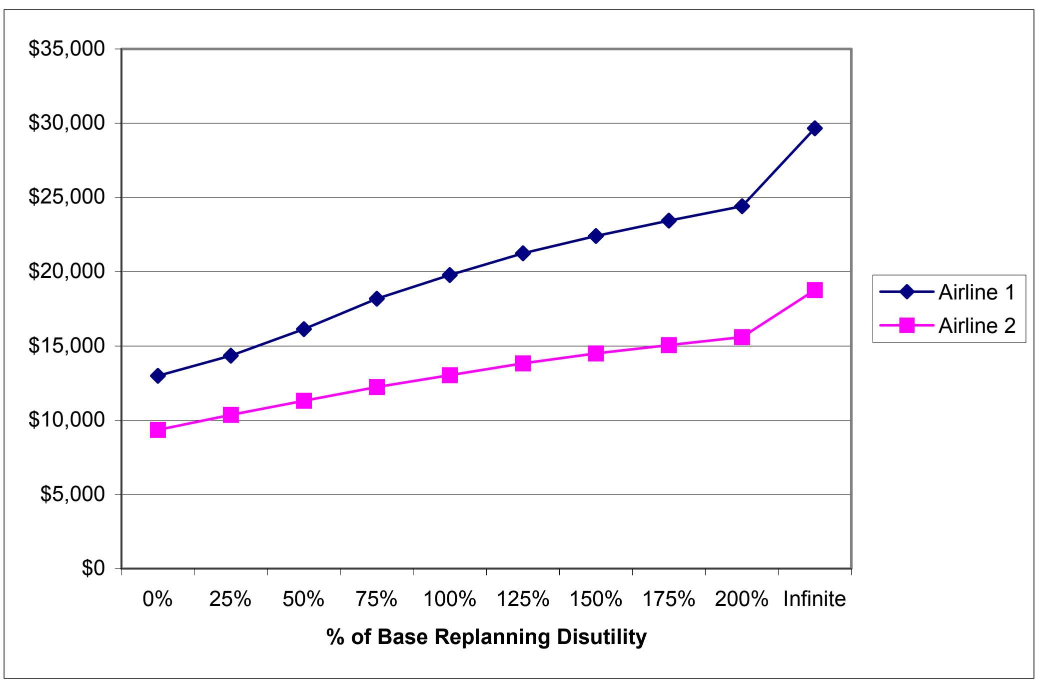


**Figure 4.6.** : Sell-up revenues in Network D

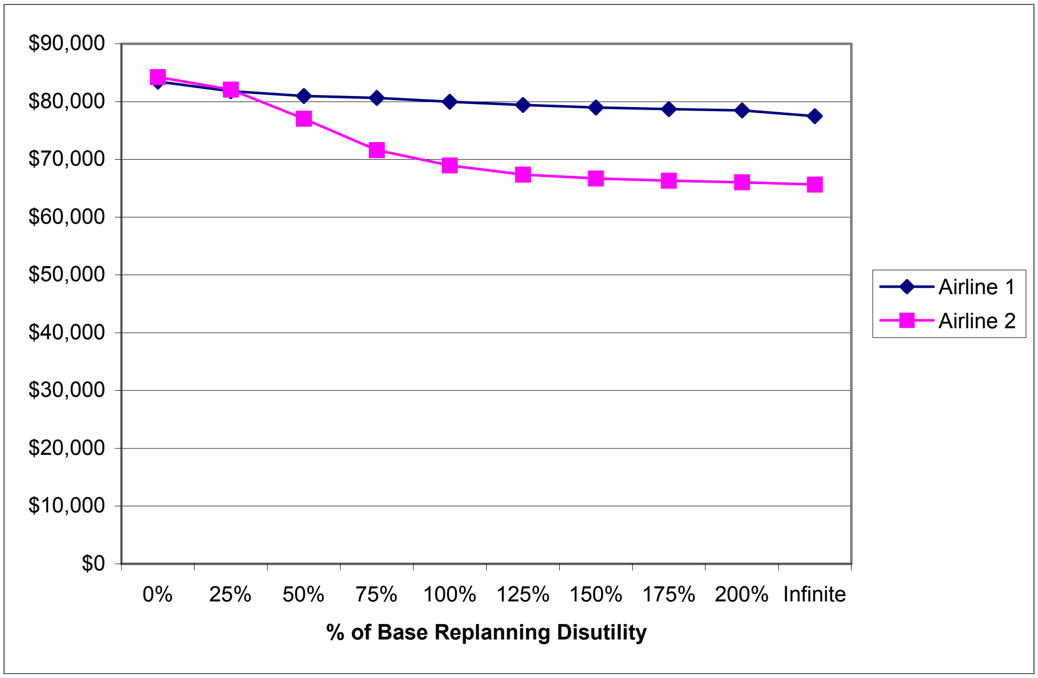


**Figure 4.7.** : Recapture revenues in Network D

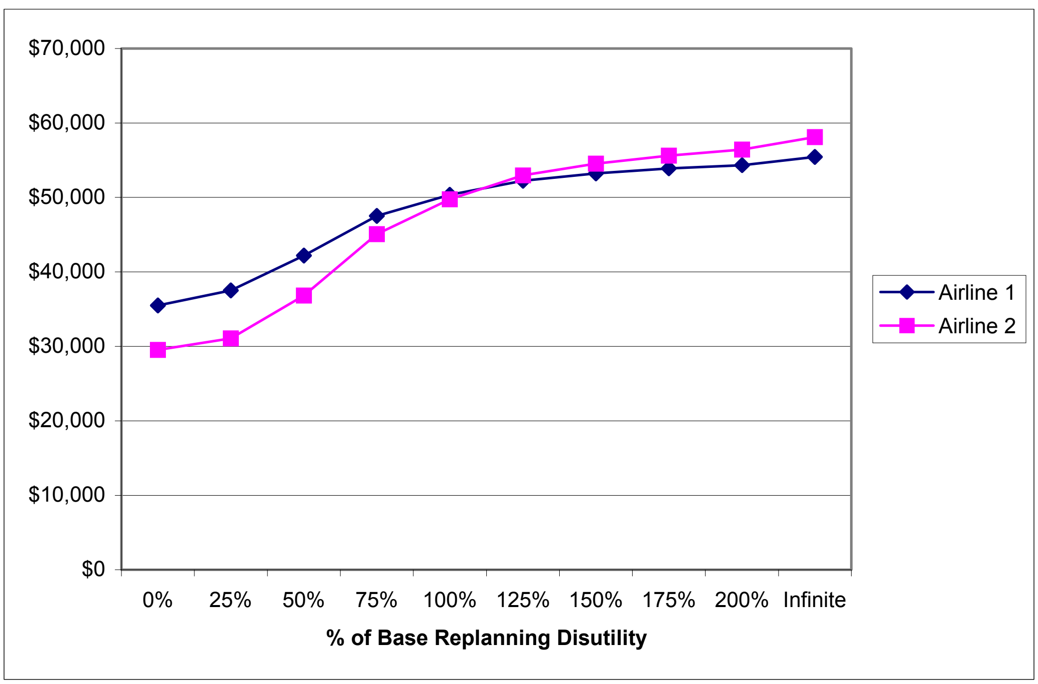


**Figure 4.8.** : Spill-in revenues in Network D

In PODS Network D, three categories of revenue follow a similar pattern for both airlines when the replanning disutility varies. As the replanning disutility increases and passengers tend to give a higher value to schedule convenience, sell-up and spill-in revenues increase as more passengers are willing to shift to a more expensive fare class or a different airline to select a path that fits within their decision window. On the contrary, more passengers are reluctant to shift to another path of the same airline as such paths might be located outside their decision window and recapture revenues decrease.

However, for the largest revenue category, first choice revenues, the replanning disutility sensitivity analysis reveals different patterns: First choice revenues increase with the replanning disutility for airline 1 but decrease for airline 2. As the replanning disutility increases and represents a higher proportion of a path/class generalized cost, some additional passengers will have airline 1 as their first choice thanks to his slight schedule advantage: For some passengers, airline 1 will be the only one for which a path is located within their decision window and a convenient schedule becomes more important as the replanning disutility increases. For those passengers, it is less and less likely for airline 2 to be their first choice as airline 1 better schedule can less and less be compensated by some other elements like path quality or the premium associated with a path on a passenger favorite airline.

However, Figures 4.9. to 4.12. show that the evolution of revenues by category is very different in PODS network E:

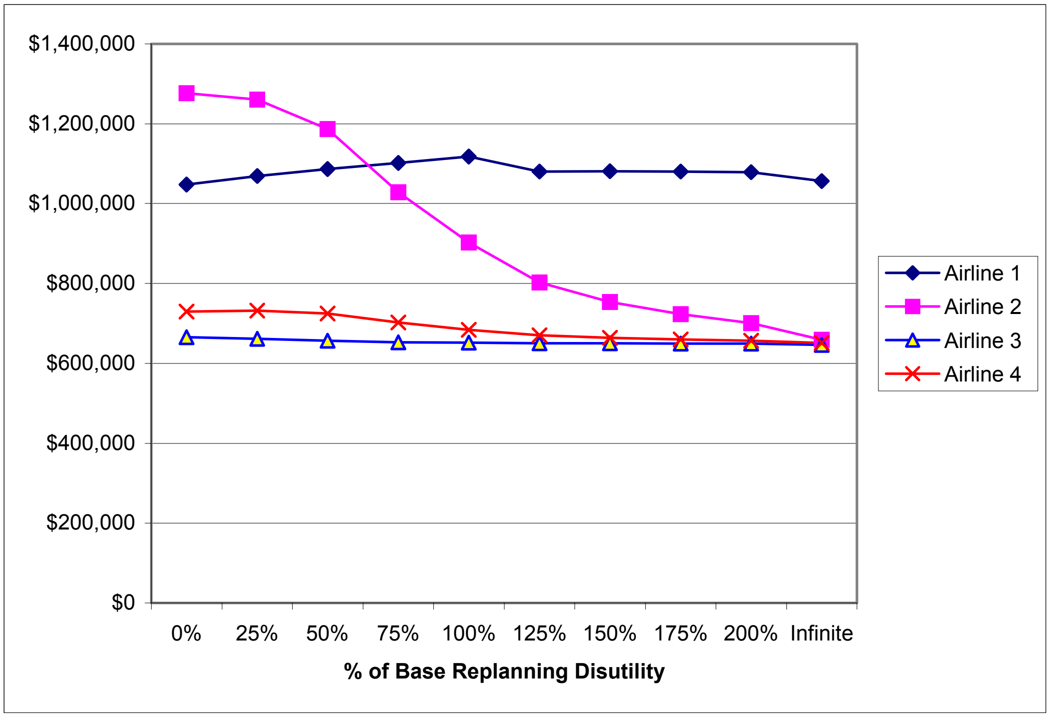


**Figure 4.9.** : First choice revenues in Network E

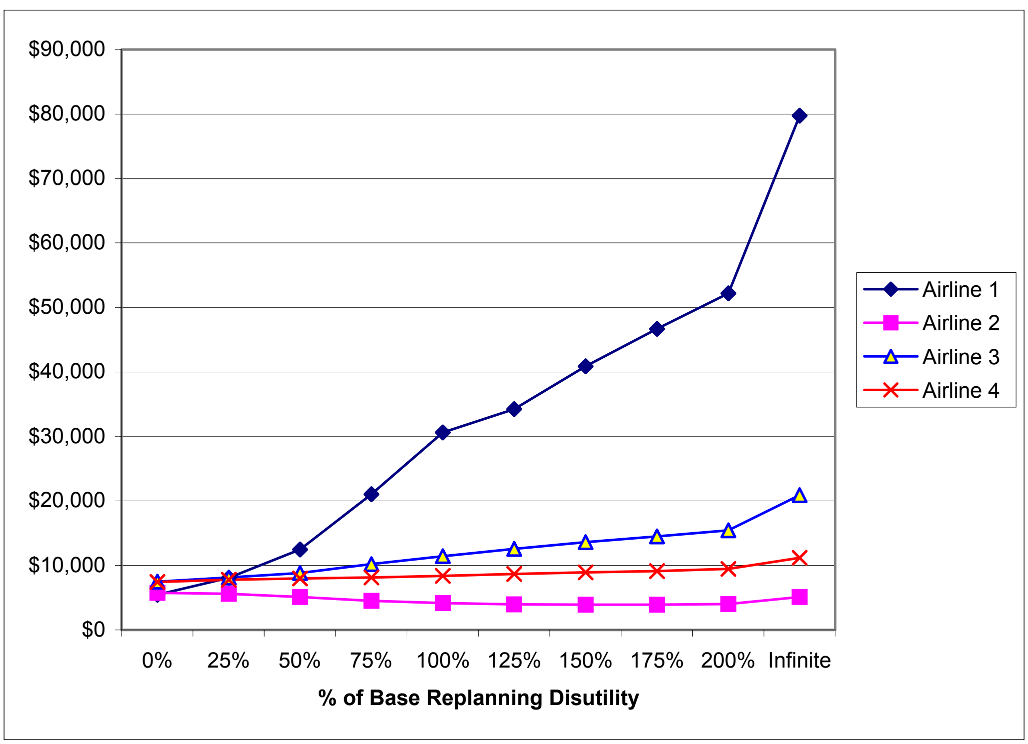


**Figure 4.10.** : Sell-up revenues in Network E

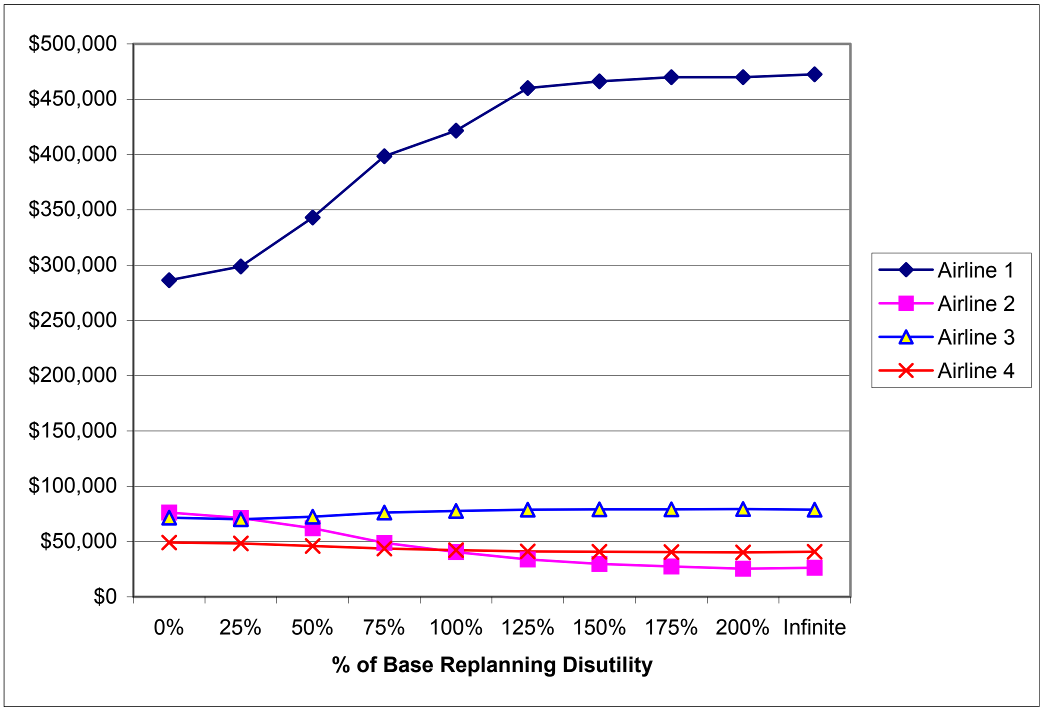


**Figure 4.11.** : Recapture revenues in Network E

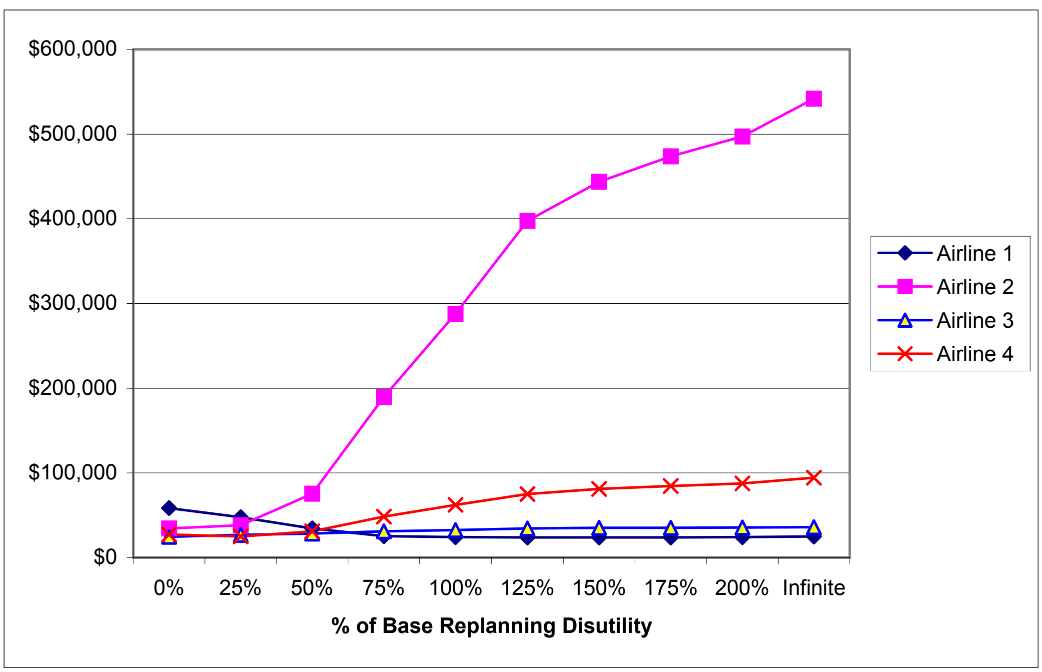


**Figure 4.12.** : Spill-in revenues in Network E

The evolution of the revenue per category figures reveal in Network E the impact of the replanning disutility on the choice of passengers in the presence of a significant asymmetry between the schedules offered by the competing airlines. The sensitivity of revenues in each category is very much driven by the impact of schedule asymmetry.

For airline 1, as in Network D, first choice revenues initially increase slightly as the replanning disutility increases. However, as the replanning disutility becomes larger, more and more passengers prefer to travel on airline 1 that offers the most convenient schedule. Faced with a very high demand, airline 1 load factor starts to increase sharply and the airline has less and less space available to accommodate the demand. As a result, a higher proportion of passengers do not get their first choice satisfied and first choice revenues start to decrease. Recapture and sell-up revenues increase sharply as the airline is faced with very high demand and cannot satisfy the first choice requirements of many passengers: A significant proportion of them accept to shift to a higher fare class on the same path or to another airline 1 path that might also be located in their decision window. However, some also get spilled to the competitor and this explains the large increase in airline 2 spill-in revenues.

For airline 2, first choice revenues decrease sharply. As schedule convenience becomes more important, less and less passengers are willing to travel on a path located outside their decision window and so less and less passengers consider airline 2 as their first choice since it has a shorter schedule coverage and offers paths that fit within the decision window for fewer passengers than airline 1. In addition, for very few passengers, airline 2 offers several paths that fit in their decision window and as a result, recapture revenues also decrease.

For airline 3 and 4, the revenue per category figures display a more stable pattern and their evolution is combination of both the forces described in the PODS network D case and the influence of their US codeshare partners. For instance, due to the lack of attractiveness of its US partner, airline 4 first choice revenues tend to decrease but airline 3 first choice revenues remains more stable. Similarly, airline 4 spill-in revenues tend to increase like for airline 2 and airline 3 recapture revenues tend to grow following the pattern observed for its US codeshare partners.

To conclude, the results of this sensitivity analysis show the large impact of the replanning disutility and how to value schedule convenience in the presence of a significant schedule asymmetry. As a result, the approach used to model schedule convenience in the Passenger Choice Model is an important factor that needs to be considered if PODS is to be used to study the impact of asymmetry in general and schedule asymmetry in particular. In order to study the impact of schedule asymmetry in PODS, the drawbacks of the constant replanning disutility model should be addressed and alternative models should be developed and tested.

## *4.4.* Alternative approaches to Passenger Preference for Schedule in PODS

To design a new approach to passenger preference for schedule, we will restrict our analysis in this section within the limits of the schedule delay and decision window concepts as these are the two approaches found in the air transportation literature that have been applied in research on passenger preference for schedule in the past.

### 4.4.1. Decision window vs. Schedule delay?

Within this framework, our first option could be to design an approach based entirely on the concept of schedule delay without any reference to a decision window. Under such an approach, instead of a decision window, each passenger would get assigned a unique ideal departure time. The ideal departure time would be assigned randomly for each passenger but would be designed such as to reproduce a pre-determined typical time of day distribution of air travel demand in each market. A disutility cost would be added to the cost of each path based on the difference between the actual path departure time and the ideal departure time of the passenger. This disutility cost could be a linear or non-linear function of the difference between actual and ideal departure time and would vary based on trip purpose and market stage length.

However, even if such an approach could be implemented in PODS, it does not seem very attractive. As already stated, we believe that most air travelers have some flexibility in the schedule of their activities at their destination and that a single ideal departure time does not exist for most passengers. Most passengers have in mind a range of acceptable travel times that is a convenient solution to their time-space problem. In addition, the results of the exploratory survey carried out by Mehndiratta supports the concept of a decision window. As a result, we will prefer alternative approaches that remain based on the decision window concept and are not solely based on a schedule delay approach.

An alternative to the PODS approach to passenger preference for schedule would be to keep the current decision window model but to modify the passenger choice set. Under that approach, like in the initial Boeing Decision Window Model, only paths that fit entirely within a passenger decision window would be considered and included in the passenger choice set. All paths outside the passenger decision window would be excluded. If there were no paths included in his decision window, the passenger would need to re-plan and would be assigned a second decision window and would pick among the available paths that fit into that new window and satisfy the passenger willingness to pay requirements.

However, as already mentioned in the previous section, such an option might not be appropriate as the purpose of the PODS Passenger Choice Model is to represent the decision process of individual air travelers faced with multiple trade-offs like for instance between fare and schedule convenience. Such a solution, which gives a primary role to schedule convenience in passenger choice and is close to an infinite replanning disutility model could lead to major counter-intuitive effects.

For instance, with the introduction of asymmetric schedules in PODS, such a solution might overestimate the competitive advantage of the carrier offering the best schedule. The definition of short decision windows for business travelers to reproduce the high value they place on schedule convenience would lead some of them to select the schedule-dominant and only carrier that offer a flight schedule that fits within their decision window without considering alternative paths offered by competing carriers that do not fit in the decision window but do not deviate largely from it. This might lead to an over-estimation of passenger preference for the carrier that offers the most convenient schedule to a large number of business travelers.

As a result, our preferred approach draws on the advantages of both the decision window and the schedule delay approaches. As in the current PODS Passenger Choice Model, a decision window would be defined for each passenger and all path/classes that fit into the window would be equivalent to the air traveler from a schedule point of view. However, unlike under the current constant replanning disutility model, path/classes that do not fit in the decision window would not be equally attractive to the passenger anymore. This would be obtained through a variable replanning disutility approach. Similar to a schedule delay model, the replanning disutility of each path would vary based on its deviation from the passenger decision window.

The behavioral assumption is that passengers have some flexibility in their schedule plans at destination and this is reflected through their decision window, but beyond this initial flexibility, any deviation from their original plans comes at a cost and induces an inconvenience to the passenger modeled through the replanning disutility. In addition, this inconvenience is growing with the deviation of the path schedule with regard to their initial plans or decision window: The replanning disutility is then a function of the deviation from the decision window. The next question is then: What should be the form of the replanning disutility function?

As already mentioned, Prossaloglou and Koppelman study suggests that the relationship between the cost to the passenger of the schedule inconvenience and the deviation from his initial schedule plans is non-linear. According to their research, the cost of schedule inconvenience is an increasing function of the deviation from the passenger ideal departure time. In addition, Mehndiratta study suggests that there exists a difference in the valuation of time by business travelers between different periods of the day like work, leisure and sleep time. This suggests that the replanning disutility of a path that does not fit within a passenger decision window depends on the amount of disruption it creates in the passenger usual work, leisure and sleep schedule. As a result, Mehndiratta work suggests that the value of the replanning disutility should be a piece-wise linear function of the deviation from the decision window based on which part of the passenger schedule gets disrupted.

### 4.4.2. The Variable Replanning Disutility Approach

As mentioned in the last section, Mehndiratta study of the impact of time of the day effects on the demand for intercity travel suggests a piece-wise linear form for the replanning disutility function. The implementation in PODS is then based on that principle. The variable replanning disutility is equal to the product of a common base replanning disutility and a replanning disutility index that depends on the deviation of the path schedule from the passenger decision window. The base replanning disutility is equal to the current PODS replanning disutility described in Chapter 3 to keep, for the same deviation from the passenger decision window, the scale in the value of replanning disutilities across O-D markets established by Lee (2000).

To allow for a high level of flexibility in the specification of various piece-wise linear replanning disutility functions representing a wide range of possible passenger behavior patterns, each segment of the piece-wise linear function has only a one-hour duration. The value passengers give to the first hour of deviation from the decision window can be specified to be different from the value they give to the second hour of deviation, the third hour of deviation etc. As a result, for each passenger type, the replanning disutility index is then composed of 24 input parameters, one for each incremental hourly deviation from the passenger decision window.

In addition, the replanning disutility index is different for business and leisure passengers as numerous studies including the work of Prossaloglou and

Koppelman have shown that the two categories exhibit different behavioral patterns with regard to passenger preference for schedule. As a result, the difference between the value of the replanning disutility for a business and leisure traveler comes from the difference in the value of both the base replanning disutility and the replanning disutility index.

The values of the replanning disutility function input parameters are designed to reproduce expected passenger behavior for both passenger types. As business passengers choose in PODS primarily between Y and B fares, the replanning disutility index is calibrated based on the proportion of business travelers that prefer a Y fare class itinerary inside their decision window to a B fare class itinerary outside their decision window all else being equal (same path quality and travel on the same airline). For leisure travelers, as they primarily selects Q and M fares, the replanning disutility is calibrated based on the proportion of passengers that prefer a M fare inside their decision window over a Q fare outside their decision window all else being equal.

Tables 4.1. and 4.2. below describe one example of a replanning disutility index for respectively business and leisure passengers in a market with a $100 Q fare, $150 M fare, $200 B fare and $400 Y fare. In this market, the base average replanning disutility is equal to $61.56 for business passengers and 11% of them prefer a B fare outside their decision window over a Y fare inside their decision window, all else being equal. Similarly, the base average replanning disutility is $11.9 for leisure passengers and 94.3% of them prefer a Q fare outside their decision window over an M fare inside their decision window.

| Hours of deviation | Replanning Disutility | % pax prefer B outside |
|---|---|---|
| 1 | $15.4 | 27.5% |
| 2 | $38.5 | 17.8% |
| 3 | $77.0 | 7.9% |
| 4 | $119.3 | 3.2% |
| 5 | $157.0 | 1.6% |
| 6 | $192.4 | 0.9% |
| 7 | $226.5 | 0.6% |
| 8 | $259.7 | 0.4% |
| 9 | $292.4 | 0.3% |
| 10 | $324.7 | 0.3% |
| 11 | $356.8 | 0.2% |
| 12 | $388.6 | 0.2% |
| 13 | $420.3 | 0.2% |
| 14 | $451.8 | 0.2% |
| 15 | $483.2 | 0.2% |
| 16 | $514.6 | 0.2% |
| 17 | $545.9 | 0.1% |
| 18 | $577.1 | 0.1% |
| 19 | $608.3 | 0.1% |
| 20 | $639.5 | 0.1% |
| 21 | $670.6 | 0.1% |
| 22 | $701.6 | 0.1% |
| 23 | $732.7 | 0.1% |
| 24 | $763.7 | 0.1% |

**Table 4.1.** : Variable replanning disutility (business travelers)

| Hours of deviation | Replanning Disutility | % pax prefer B outside |
|---|---|---|
| 1 | $11.9 | 94.3% |
| 2 | $13.4 | 91.4% |
| 3 | $14.9 | 87.7% |
| 4 | $16.4 | 83.2% |
| 5 | $17.9 | 78.1% |
| 6 | $19.3 | 72.5% |
| 7 | $20.8 | 66.6% |
| 8 | $22.3 | 60.5% |
| 9 | $23.8 | 54.6% |
| 10 | $25.3 | 48.9% |
| 11 | $26.8 | 43.6% |
| 12 | $28.3 | 38.7% |
| 13 | $29.8 | 34.2% |
| 14 | $31.2 | 30.2% |
| 15 | $32.7 | 26.6% |
| 16 | $34.2 | 23.4% |
| 17 | $35.7 | 20.7% |
| 18 | $37.2 | 18.2% |
| 19 | $38.7 | 16.1% |
| 20 | $40.2 | 14.3% |
| 21 | $41.7 | 12.7% |
| 22 | $43.1 | 11.3% |
| 23 | $44.6 | 10.0% |
| 24 | $46.1 | 9.0% |

**Table 4.2.** : Variable replanning disutility (leisure passengers)

Figure 4.13. describes the difference in expected passenger behavior between the initial constant replanning disutility and the new variable replanning disutility approaches

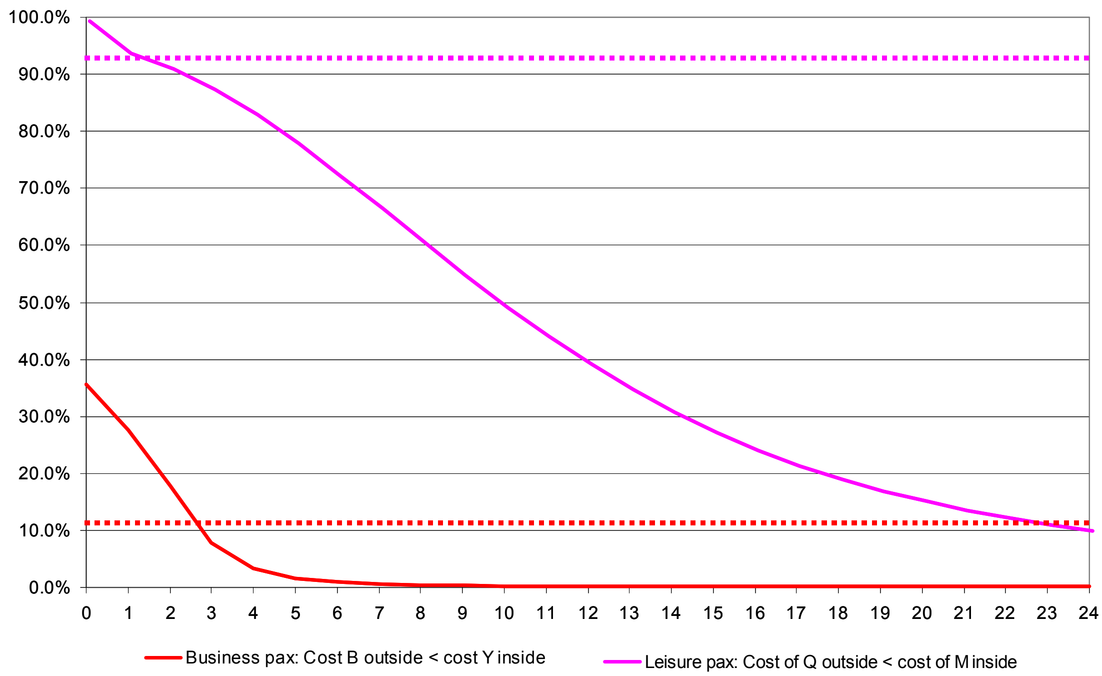


**Figure 4.13.** : Expected passenger behavior under the initial and new approach to passenger preference for schedule

Finally, Figure 4.14. below describes the value of the replanning disutility under the two schemes for a business passenger with a decision window that spans from 9 a.m. to 3 p.m..

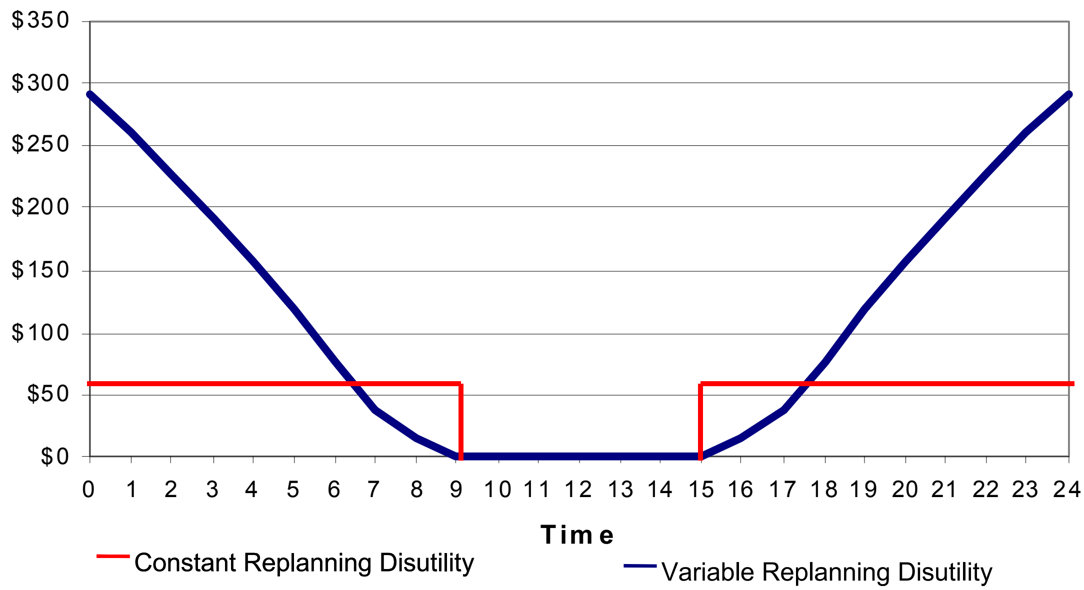


**Figure 4.14.** : Value of the replanning disutility for a business passenger with a 9 a.m.-3 p.m. decision window

As a result, the variable replanning disutility approach takes into account that if a passenger has a range of acceptable schedules that are equally attractive to him, unlike in the initial PODS design, everything outside that range is not equivalent and he does not give the same value to a path close to his schedule requirements and one that deviates largely from it. We expect the variable replanning disutility approach to have a greater impact on business travelers as they have shorter decision windows and more path/classes are likely to be outside their decision window. With such an approach, we are able to model the fact that business travelers might be extremely reluctant to accept path/classes that deviate largely from their desired schedule but might still consider flying on an airline that offers a better price/restriction value on a path that deviates only a little from the passenger decision window.

## 4.5. Conclusion

In this chapter, we have examined how passenger preference for schedule is currently modeled in PODS and compared the PODS approach to the concepts found in the air transportation literature. Based on our analysis of the strengths and weaknesses of the constant replanning disutility model used in PODS, we have defined an alternative, the variable replanning disutility model.

In the next chapter, we will test the variable replanning disutility approach through PODS simulations. Based on simulation results, we will assess the impact of the variable replanning disutility models on PODS results in both a schedule-symmetric and a schedule-asymmetric environment.

# Chapter 5 PODS Simulation Results

## 5.1. Introduction

In the previous chapter, we have described a new alternative approach to model passenger preference for schedule in PODS called the variable replanning disutility model. In this chapter, we will use the simulator to test the impact of this approach on the revenue performance of airlines in a competitive environment.

In particular, we will show how passenger preference for schedule affects some key airline performance characteristics like airline revenues and load factor. We will also study the differential impact of passenger preference for schedule in two different competitive environments, one called schedule-symmetric where all competitor airlines offer similar schedules and one called schedule-asymmetric where the competing airlines differentiate themselves by offering different schedules.

This chapter is organized as follows: Section 5.2. is an explanation of the base case settings with an emphasis on some key inputs of the simulator. Section 5.3. presents the simulation results of the new approach to model passenger preference for schedule in PODS, followed by a summary in section 5.4.

## 5.2. Simulation Set-up

### 5.2.1. Base Case Settings

Before presenting the results of PODS simulations, it is necessary to describe the base case used as a basis for comparison of all simulations performed for this thesis. Since the objective of this thesis is to understand preference for schedule in the airline industry and enhance how passenger preference for schedule is modeled in PODS, our base case will be the current approach to passenger preference for schedule in PODS described in Chapter 4 and called the constant replanning disutility model. As a result, in the base case, a constant replanning disutility is added to the cost of all path/classes that do not fit entirely within a passenger decision window irrespective of the time position of the path relative to the passenger decision window. As described in Chapter 4, the value of the replanning disutility depends on trip purpose and the market basefare.

### 5.2.2. Variable Replanning Disutility Functions

In Chapter 4, we defined a new approach to passenger preference for schedule in PODS based on a variable replanning disutility. Under this model, the value of the replanning disutility added to the cost of every path located outside a passenger decision window depends on the position in time of the path relative to the decision window. As described in Chapter 4, the value of the replanning disutility is calibrated to represent expected passenger behavior for both business and leisure passengers. For business passengers (respectively leisure passengers), the value of the replanning disutility is calibrated based on the proportion of passengers that prefer a Y path/class (respectively a M path/class) inside their decision window over a B path/class (respectively a Q path/class) outside their decision window, all else being equal. We will consider three alternatives, labeled low, medium and high for the value of passenger replanning disutilities. Tables 5.1. and 5.2. and Figures 5.1. and 5.2. describe the three levels of replanning disutility and the associated passenger preferences in a market with a lowest fare (Q fare) of $100. In addition, in order to isolate the contribution of each type of passenger to the change in airline revenue performance, we will consider in this chapter three simulation scenarios:

- Scenario 1: The replanning disutility is constant for leisure passengers and variable for business passengers

- Scenario 2: The replanning disutility is constant for business passengers and variable for leisure passengers

- Scenario 3: The replanning disutility is variable for both types of passengers

| | Low | | Medium | | High | |
|---|---|---|---|---|---|---|
| Hours of deviation | Replanning Disutility | % pax prefer B outside | Replanning Disutility | % pax prefer B outside | Replanning Disutility | % pax prefer B outside |
| 1 | $15.4 | 27.5% | $15.4 | 27.5% | $15.4 | 27.5% |
| 2 | $30.8 | 20.7% | $38.5 | 17.8% | $53.9 | 12.9% |
| 3 | $51.3 | 13.6% | $77.0 | 7.9% | $97.5 | 5.1% |
| 4 | $77.0 | 7.9% | $119.3 | 3.2% | $150.1 | 1.8% |
| 5 | $110.8 | 3.8% | $157.0 | 1.6% | $193.9 | 0.9% |
| 6 | $143.6 | 2.0% | $192.4 | 0.9% | $233.4 | 0.6% |
| 7 | $167.1 | 1.4% | $226.5 | 0.6% | $270.4 | 0.4% |
| 8 | $184.7 | 1.0% | $259.7 | 0.4% | $305.9 | 0.3% |
| 9 | $198.4 | 0.9% | $292.4 | 0.3% | $340.3 | 0.3% |
| 10 | $209.3 | 0.8% | $324.7 | 0.3% | $374.0 | 0.2% |
| 11 | $218.3 | 0.7% | $356.8 | 0.2% | $407.1 | 0.2% |
| 12 | $225.7 | 0.6% | $388.6 | 0.2% | $439.9 | 0.2% |
| 13 | $232.0 | 0.6% | $420.3 | 0.2% | $472.4 | 0.2% |
| 14 | $237.4 | 0.6% | $451.8 | 0.2% | $504.6 | 0.2% |
| 15 | $242.1 | 0.5% | $483.2 | 0.2% | $536.6 | 0.1% |
| 16 | $246.2 | 0.5% | $514.6 | 0.2% | $568.5 | 0.1% |
| 17 | $249.9 | 0.5% | $545.9 | 0.1% | $600.2 | 0.1% |
| 18 | $253.1 | 0.5% | $577.1 | 0.1% | $631.8 | 0.1% |
| 19 | $256.0 | 0.5% | $608.3 | 0.1% | $663.4 | 0.1% |
| 20 | $258.6 | 0.5% | $639.5 | 0.1% | $694.9 | 0.1% |
| 21 | $260.9 | 0.4% | $670.6 | 0.1% | $726.3 | 0.1% |
| 22 | $263.0 | 0.4% | $701.6 | 0.1% | $757.6 | 0.1% |
| 23 | $265.0 | 0.4% | $732.7 | 0.1% | $788.9 | 0.1% |
| 24 | $266.8 | 0.4% | $763.7 | 0.1% | $820.2 | 0.1% |

**Table 5.1.** : Variable replanning disutility (business travelers)

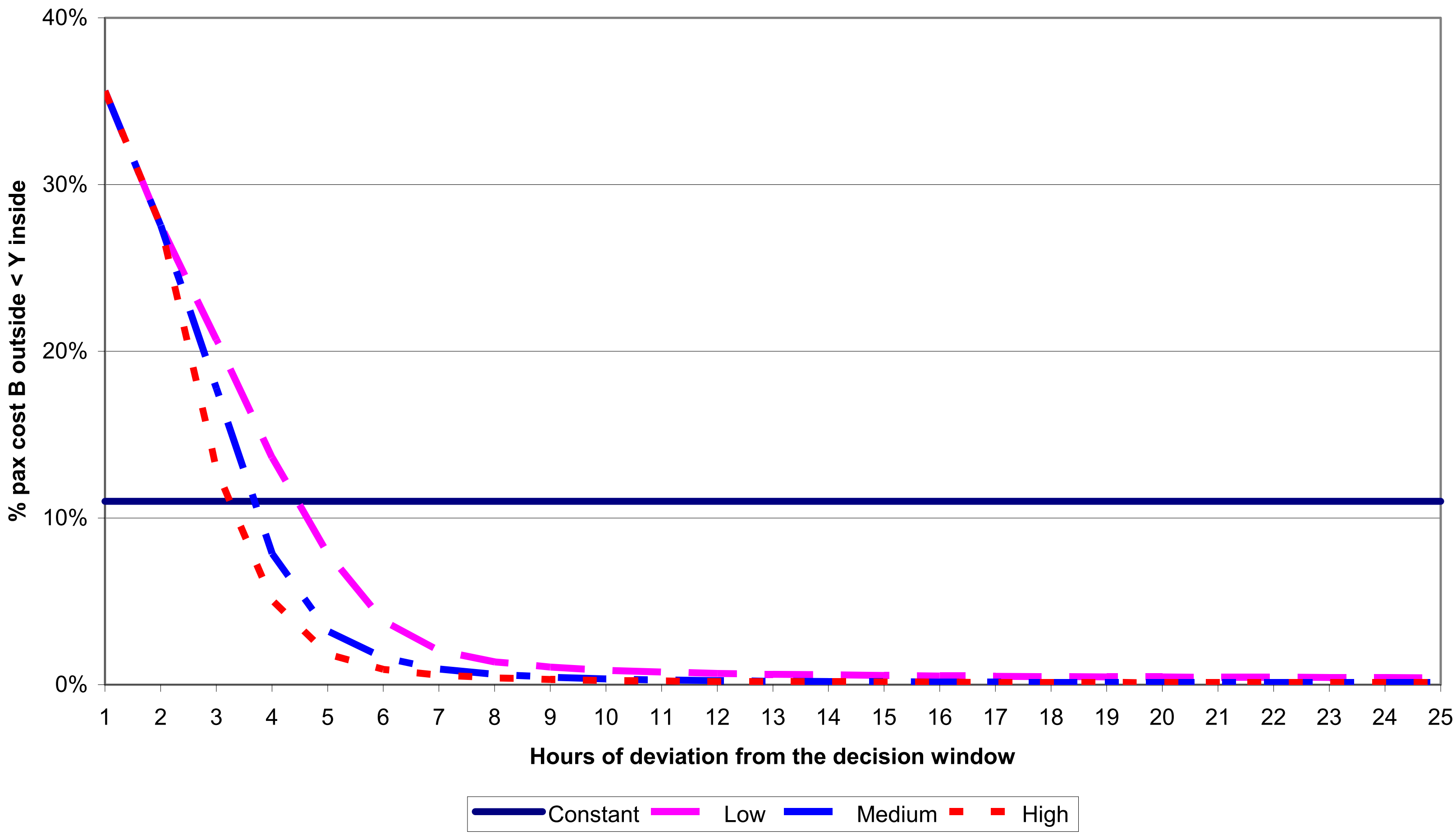


**Figure 5.1.** : Expected passenger behavior under the constant and variable replanning disutility models (business passengers)

| | Low | | Medium | | High | |
|---|---|---|---|---|---|---|
| Hours of deviation | Replanning Disutility | % pax prefer Q outside | Replanning Disutility | % pax prefer Q outside | Replanning Disutility | % pax prefer Q outside |
| 1 | $3.0 | 99.8% | $11.9 | 94.3% | $11.9 | 94.3% |
| 2 | $4.5 | 99.6% | $13.4 | 91.4% | $13.4 | 91.4% |
| 3 | $6.0 | 99.3% | $14.9 | 87.7% | $14.9 | 87.7% |
| 4 | $7.4 | 98.8% | $16.4 | 83.2% | $16.4 | 83.2% |
| 5 | $8.9 | 97.8% | $17.9 | 78.1% | $17.9 | 78.1% |
| 6 | $10.4 | 96.4% | $18.8 | 74.4% | $19.3 | 72.5% |
| 7 | $11.9 | 94.3% | $19.6 | 71.6% | $20.8 | 66.6% |
| 8 | $13.4 | 91.4% | $20.1 | 69.5% | $22.3 | 60.5% |
| 9 | $14.5 | 88.6% | $20.5 | 67.9% | $23.8 | 54.6% |
| 10 | $15.5 | 86.0% | $20.8 | 66.6% | $25.3 | 48.9% |
| 11 | $16.2 | 83.7% | $21.1 | 65.5% | $26.8 | 43.6% |
| 12 | $16.9 | 81.6% | $21.3 | 64.6% | $28.3 | 38.7% |
| 13 | $17.4 | 79.7% | $21.5 | 63.8% | $29.8 | 34.2% |
| 14 | $17.9 | 78.1% | $21.7 | 63.1% | $31.2 | 30.2% |
| 15 | $18.2 | 76.6% | $21.8 | 62.5% | $32.7 | 26.6% |
| 16 | $18.6 | 75.3% | $21.9 | 62.0% | $34.2 | 23.4% |
| 17 | $18.9 | 74.2% | $22.1 | 61.6% | $35.7 | 20.7% |
| 18 | $19.2 | 73.1% | $22.1 | 61.2% | $37.2 | 18.2% |
| 19 | $19.4 | 72.2% | $22.2 | 60.9% | $38.7 | 16.1% |
| 20 | $19.6 | 71.3% | $22.3 | 60.5% | $40.2 | 14.3% |
| 21 | $19.8 | 70.5% | $22.4 | 60.3% | $41.7 | 12.7% |
| 22 | $20.0 | 69.8% | $22.4 | 60.0% | $43.1 | 11.3% |
| 23 | $20.2 | 69.2% | $22.5 | 59.8% | $44.6 | 10.0% |
| 24 | $20.3 | 68.6% | $22.6 | 59.5% | $46.1 | 9.0% |

**Table 5.2.** : Variable replanning disutility (leisure travelers)

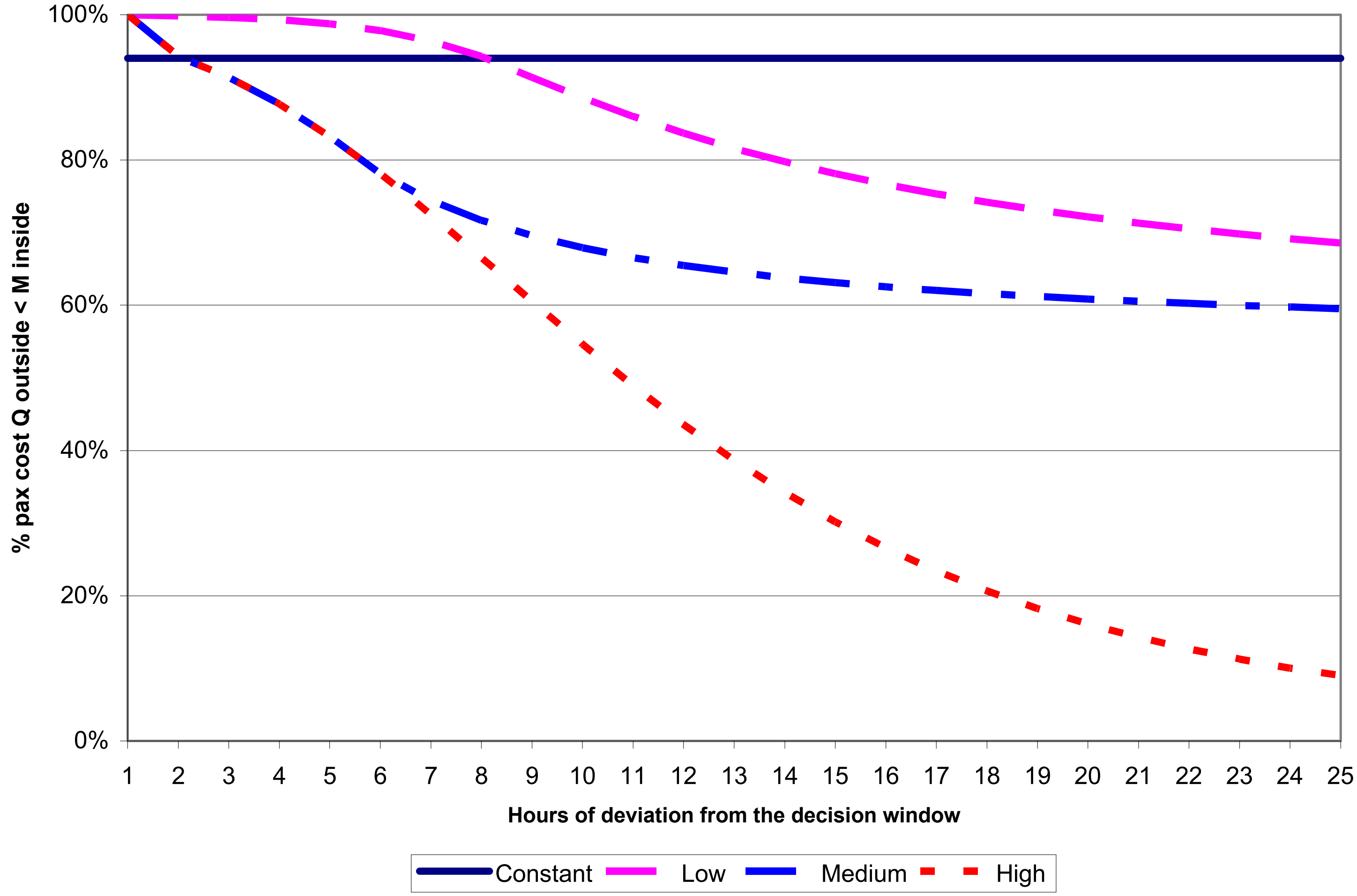


**Figure 5.2.** : Expected Passenger behavior under the constant and variable replanning disutility models (leisure passengers)

### 5.2.3. Other PODS Inputs

Apart from the replanning disutility, other major PODS inputs include other passenger preference inputs and forecasting, detruncation and revenue management method inputs. As the objective of this thesis is to model passenger preference for schedule in the airline industry and understand its impact on PODS simulation results, all these inputs will remain constant across all simulations.

All other passenger preference inputs like restriction, path quality index and unfavorite airline disutility are equal to the settings described in Chapter 3. In addition, we will use the current default PODS detruncation and forecasting techniques, i.e. booking curve detruncation and pick-up forecasting methods. For a description of these methods, the reader is referred to Darot (2001).

In addition, since revenue management is not the primary focus of this thesis, all airlines will use the standard Expected Marginal Seat Revenue (EMSR) algorithm to control seat allocation among fare classes. EMSR was first introduced by Belobaba (1987) and has been used since then by a large number of airlines for seat allocation purposes. For a description of EMSR, the reader is also referred to Darot (2001).

Finally, the simulations were performed in both “schedule-symmetric” Network D and “schedule-asymmetric” network E. For a description of these two network environments, the reader is referred to Chapter 3. The average load factor is equal to 83% and 80% in network D and E respectively.

## 5.3. Simulation Results

### 5.3.1. Revenues

Let us first look at the impact on airline revenues following the introduction of a variable replanning disutility in Network D and E.

**Network D**

Figures 5.3., 5.4. and 5.5. below show the change in airline revenues for the three simulation scenarios.

First of all, the impact of switching from the current constant replanning disutility to the new variable disutility is much greater for business passengers (scenario 1) than for leisure passengers (scenario 2). This is expected since business and leisure passengers have different behavioral patterns with business passengers giving more emphasis to non-monetary elements like preference for schedule and leisure passengers' decisions being more influenced by fare levels. In PODS, business passengers have on average a much higher replanning disutility than leisure passengers and the replanning disutility represents a higher proportion of the total cost of a path for business than for leisure passengers. As a result, we expect airline revenues to be more sensitive to a change in the way business passengers value preference for schedule. The introduction of a variable replanning disutility in PODS Network D can lead to change in revenues over 1% if this new valuation of passenger preference for schedule is introduced for business passengers only versus a maximum change in revenues of no more than 0.15% if it is used by leisure passengers only.

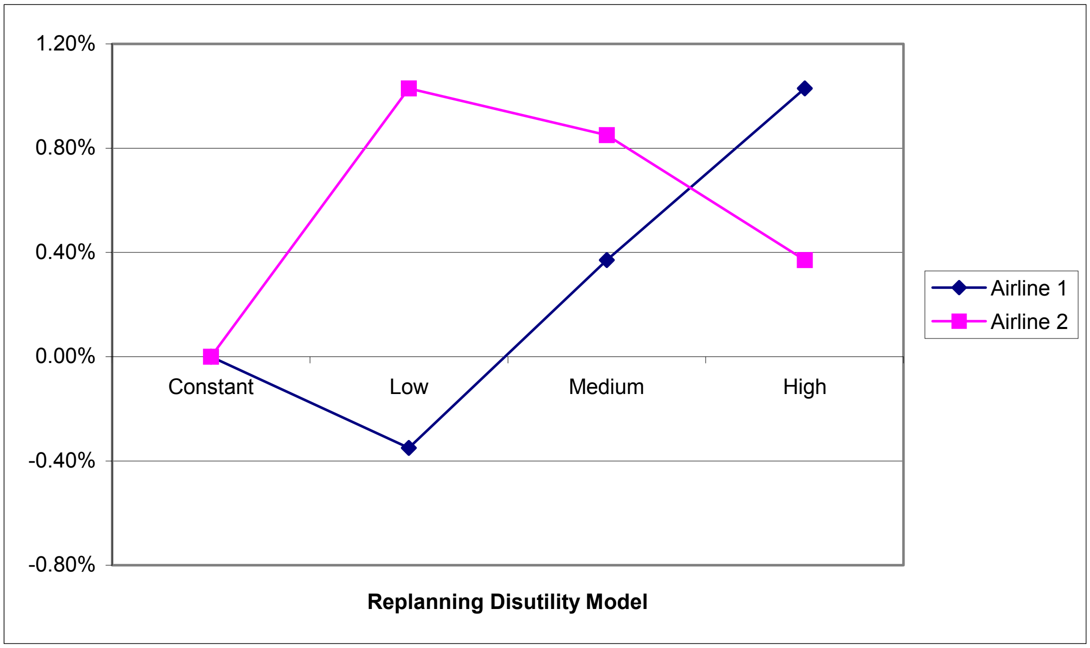


**Figure 5.3.** : % change in simulated airline revenues when business passengers only use a variable replanning disutility (Scenario 1)

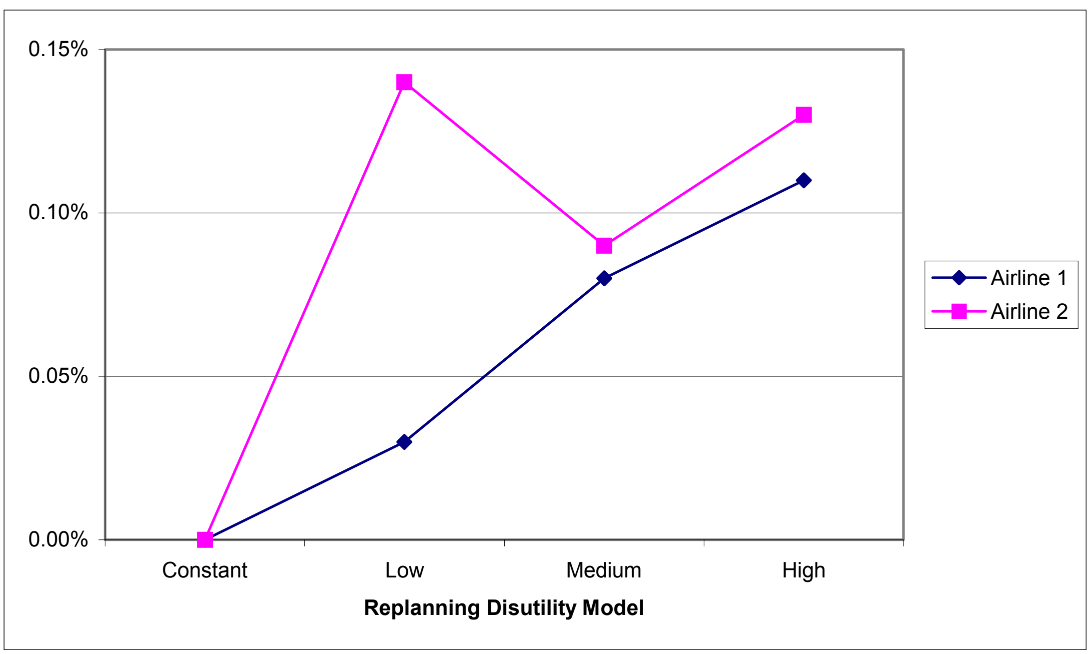


**Figure 5.4.** : % change in simulated airline revenues when leisure passengers only use a variable replanning disutility (Scenario 2)

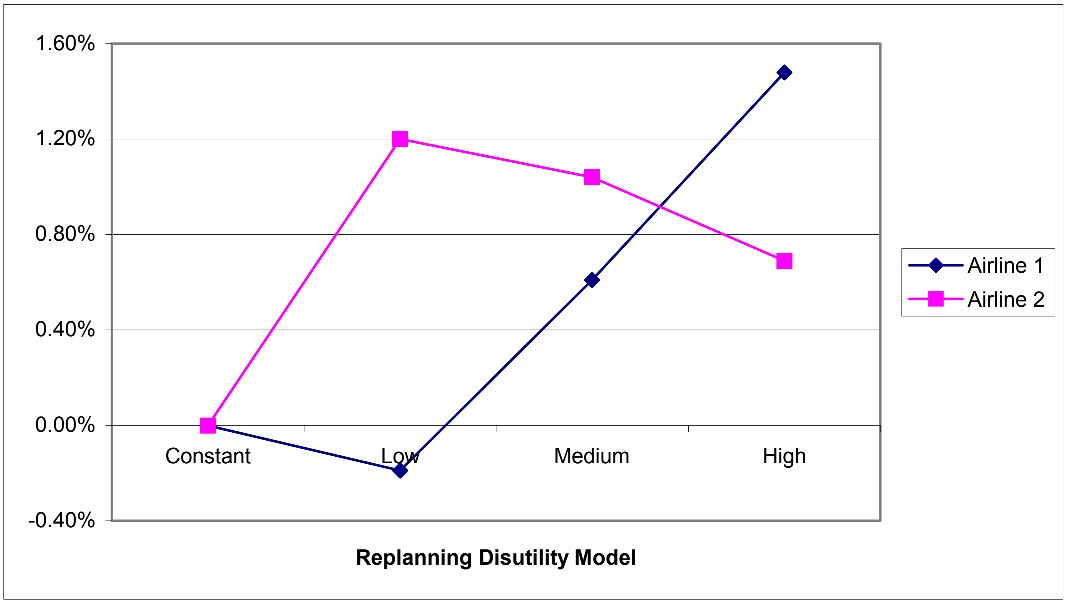


**Figure 5.5.** : % change in simulated airline revenues when both business and leisure passengers use a variable replanning disutility (Scenario 3)

In addition, simulation results also show that simulated airline revenues at the system level (sum of airline 1 and 2 revenues) tend to increase for all scenarios and all values of the replanning disutility. This increase in simulated total revenues for the industry is larger for the medium and large variable replanning disutility alternatives than for the low alternative. Table 5.3. below shows the change in revenues at the industry level in PODS Network D. The base case for this table is industry revenues with a constant replanning disutility model

| | *Scenario 1* | *Scenario 2* | *Scenario 3* |
|---|---|---|---|
| Low | 0.33% | 0.09% | 0.50% |
| Medium | 0.61% | 0.09% | 0.83% |
| High | 0.70% | 0.12% | 1.09% |

**Table 5.3.** : % change in simulated airline revenues at the system level in Network D

As already stated in Chapter 4, the major drawback of the constant replanning disutility model is that the value of a path located outside a passenger decision window is independent of the time location of the path relative to the decision window. As a result, the constant replanning disutility model under-estimates the preference of a passenger for a path located close to his decision window but over-estimates his preference for a schedule-inconvenient path located far outside his decision window. When a variable replanning disutility model is introduced, passengers become very reluctant to travel on schedule-inconvenient paths located far outside their decision window and would rather pay a higher fare to switch to a more schedule-convenient path located inside or close to their decision window. As a result, simulated revenues at the system level increase and this trend is strengthened as the value of the replanning disutility increases and passengers are less and less willing to travel on schedule-inconvenient paths. This is especially true for business passengers that have on average shorter decision windows and higher replanning disutilities.

Finally, let us consider how these incremental revenues at the system level are split between the two competing airlines. Simulation results show that when the variable replanning disutility is relatively low, airline 2 revenues tend to increase while airline 1 revenues decrease. However, when the replanning disutility is high, we observe higher revenues gains for airline 1 and airline 2 revenues tend to decrease. Actually, as stated in Chapter 3, despite very similar schedules, airline 1 has a small schedule advantage over airline 2 due to the geographical location of its hub relative to the bulk of traffic flows in Network D. When the variable replanning disutility is low, business passengers are more willing to switch to airline 2: Since airline 2 offers slightly more inconvenient schedules with longer travel times in most markets, airline 2 paths might lie outside the decision window for some business passengers but relatively close to it. As a result, if the variable disutility is very low for paths close to a passenger decision window, some airline 1 passengers might be willing to switch to airline 2 to take advantage of a lower fare/restriction product, a better path quality or probably more frequently to travel on their favorite airline.

However, as the value of passenger preference for schedule increases and the variable replanning disutility becomes larger even for paths close to a passenger decision window, business passengers are less and less willing to switch airlines and prefer to travel on the airline offering the most convenient schedule. As a result, revenue gains become higher for airline 1 and revenues start to decrease for airline 2.

To summarize, the use of a variable replanning disutility model leads to an increase in PODS simulated revenues at the system level but also to a more realistic evaluation of the benefits of airline 1 schedule advantage: the variable replanning disutility model leads to a decrease in the cost of paths located outside but close to a passenger decision window and this benefits airline 2 that offers slightly less attractive schedules.

**Network E**

As shown by the sensitivity analysis of airline revenues with regard to the value of a constant replanning disutility performed in Chapter 4, the impact of how passengers value preference for schedule on airline revenues is much greater in schedule-asymmetric network E than in schedule-symmetric network D. Similarly, the magnitude of the change in revenues is much greater in network E than in network D when a variable replanning disutility is introduced. Figures 5.6., 5.7. and 5.8. below show the evolution of airline revenues in network E for the three simulation scenarios.

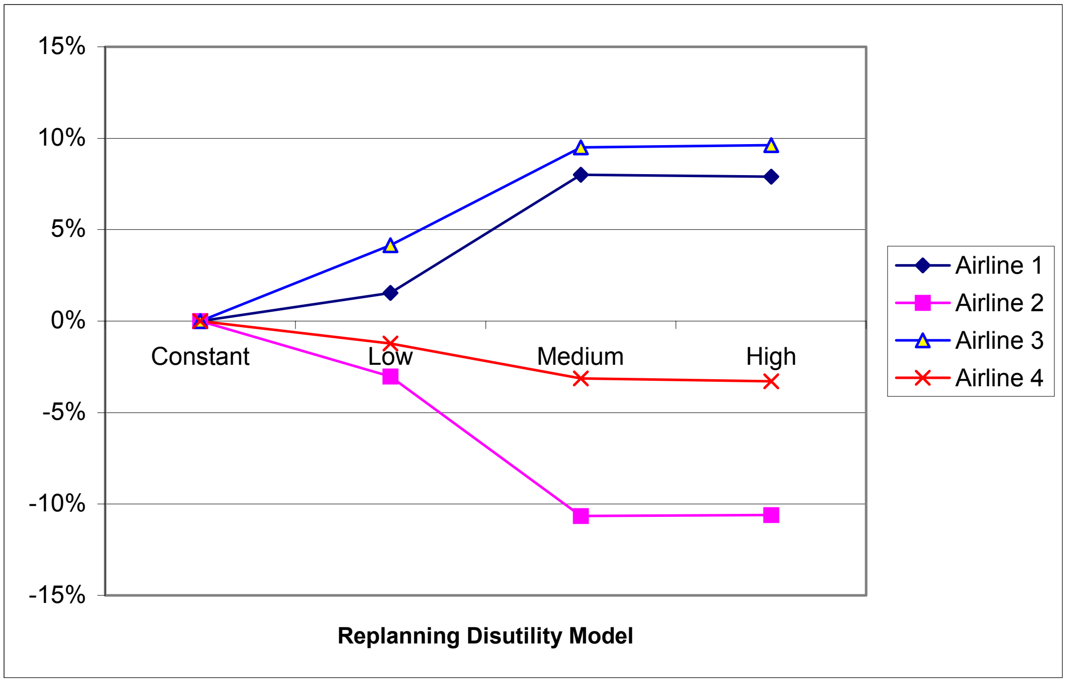


**Figure 5.6.** : % change in simulated airline revenues when business passengers only use a variable replanning disutility (Scenario 1)

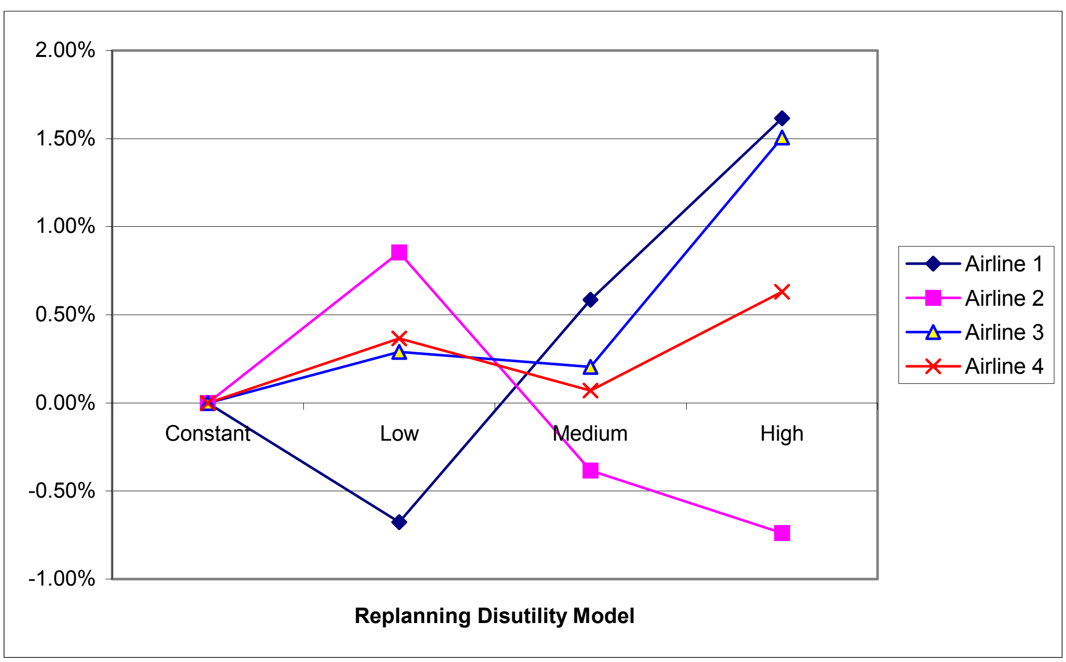


**Figure 5.7.** : % change in simulated airline revenues when leisure passengers only use a variable replanning disutility (Scenario 2)

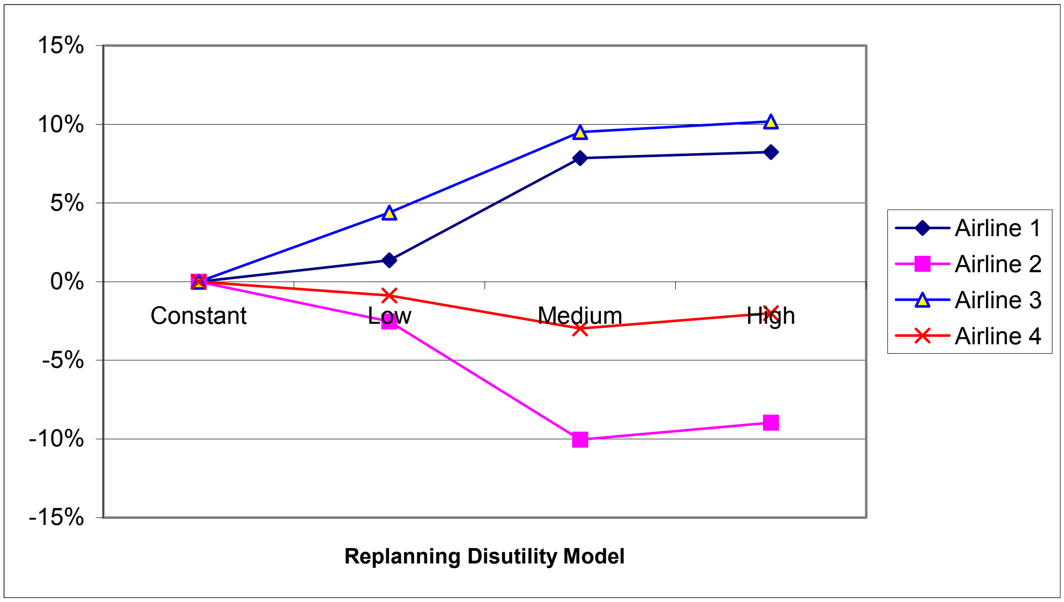


**Figure 5.8.** : % change in simulated airline revenues when both business and leisure passengers use a variable replanning disutility (Scenario 3)

As in Network D, the introduction of a variable replanning disutility has a greater impact for business than for leisure passengers. When the variable replanning disutility is used for both types of passengers, the change in airline revenues comes primarily from business passengers.

In addition, like in Network D, revenues at the system level tend to increase for all scenarios and all values of the replanning disutilities. However, this increase is more moderate than in Network D for scenarios 1 and 3 as airline 1 and 3 do not have sufficient capacity to accommodate all the passenger demand that prefer to travel on schedule–convenient paths and a large number of passengers are unable to book their most preferred alternative as confirmed by the analysis of more detailed revenue data in the next subsection.

| | *Scenario 1* | *Scenario 2* | *Scenario 3* |
|---|---|---|---|
| Low | 0.02% | 0.12% | 0.22% |
| Medium | 0.27% | 0.17% | 0.45% |
| High | 0.25% | 0.23% | 0.56% |

**Table 5.4.** : % change in simulated airline revenues at the system level in Network E

Finally, like in the sensitivity analysis of last chapter, the schedule-dominant airline (airline 1) and its alliance partner (airline 3) benefit from the variable replanning disutility approach, especially if the replanning disutility is assumed to be relatively high and passengers view preference for schedule as an important element in the choice of a travel alternative. A large number of passengers, especially business passengers prefer to travel on airline 1 and 3 that offer the most attractive schedule.

On the contrary, revenues of airline 2 that offers a less convenient schedule decrease as a variable replanning disutility is introduced, especially for the medium and high alternatives and this also affects negatively the revenues of its alliance partner, airline 4. For a significant number of passengers, airline 2 paths are located relatively far from the boundaries of their decision window. For these passengers, the value of the replanning disutility increases when the variable replanning disutility model is used instead of the constant replanning disutility, especially for the medium and high alternatives. These passengers consider airline 2 schedules as very unattractive and most of them prefer then to travel on airline 1. This explains the large decline in airline 2 revenues when a variable replanning disutility is introduced.

As a result, the introduction of a variable replanning disutility leads to a more realistic evaluation of the revenue advantage associated with attractive schedules as the constant replanning disutility approach tends to under-evaluate the cost of paths located far from the passenger decision window, something occurring often in Network E due to the difference between the airline schedules.

### 5.3.2. Load factor and Loads by Fare Class

Let us now examine the impact of the variable replanning disutility model on airline load factor and fare class mix (loads by fare class).

**Network D**

Table 5.5. below compares the load factor for the constant and variable replanning disutility models.

| | *Airline 1* | | | *Airline 2* | | |
|---|---|---|---|---|---|---|
| | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 1 | Scenario 2 | Scenario 3 |
| Constant | | 83.73% | | | 83.46% | |
| Low | 83.27% | 83.53% | 83.13% | 83.66% | 83.64% | 83.75% |
| Medium | 83.38% | 83.60% | 83.29% | 83.53% | 83.43% | 83.53% |
| High | 83.54% | 83.51% | 83.37% | 83.30% | 83.54% | 83.55% |

**Table 5.5.** : Airline load factor in Network D

As expected, when a low variable replanning disutility is introduced, airline 1 load factor tends to decrease and airline 2 load factor tends to increase as the variable replanning disutility model reduces the benefit of airline 1 schedule advantage: More passengers, especially business passengers, now select airline 2 as their first choice because the cost associated with paths located outside a passenger decision but close to it has been lowered. However, as replanning disutilities become larger and we shift from the low to the medium and high alternatives, this effect tends to progressively disappear and airline 1 load factor increases again.

This trend is confirmed by the analysis of the loads by fare class or fare class mix. In Figures 5.9. and 5.10. below, fare classes are grouped in two categories: high-yield fare classes (Y and B fare classes) that are selected primarily by business passengers and low-yield fare classes (M and Q fare classes) selected primarily by leisure passengers.

When a variable replanning disutility is introduced, fare class mix at the industry level improves: The number of high-yield passengers increases while the number of low yield passengers decreases. This is expected, as with a variable replanning disutility, passengers are more willing to purchase higher fares in order to travel on a schedule convenient path.

However, the change in fare class mix is not similar for airline 1 and airline 2. If the variable replanning disutility is low, fare class mix improves for airline 2 and deteriorates for airline 1. With a low replanning disutility, more business passengers will choose airline 2 as their first choice due to a lower and more accurate valuation of the schedule inconvenience associated with airline 2 slightly less convenient schedules. But, as the cost of schedule inconvenience increases for the medium and high alternatives, convenient schedules become more valuable and airline 1 loads in the high-yield classes increase as well.

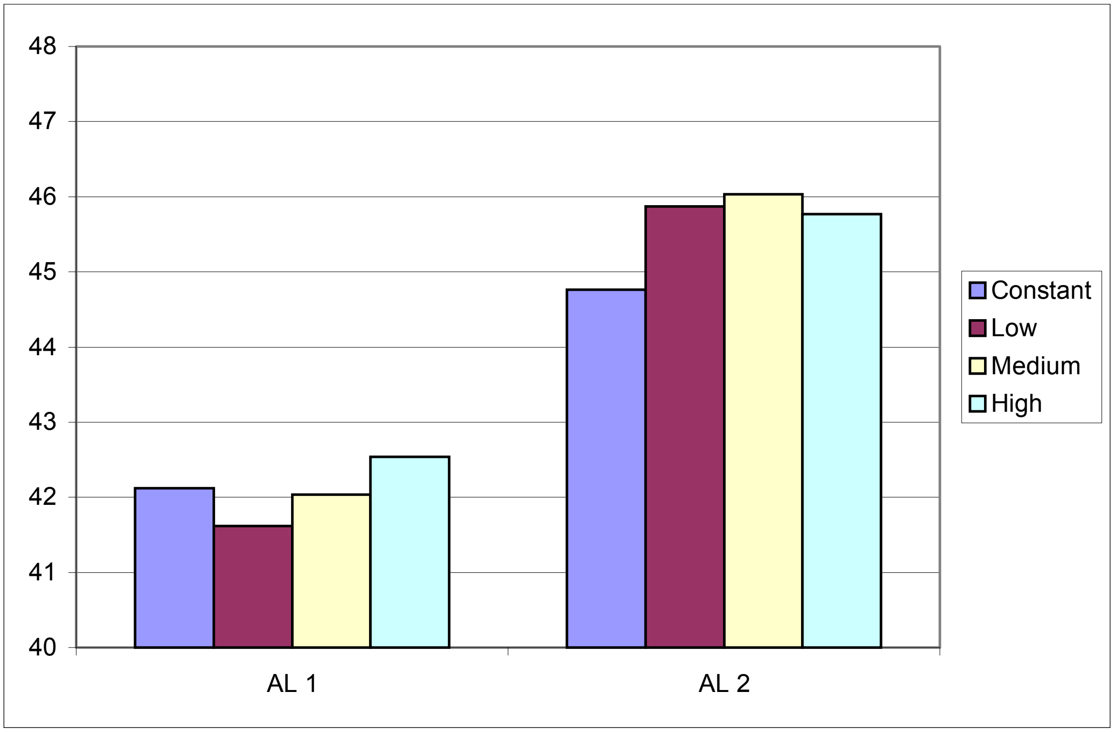


**Figure 5.9.** : Loads in the high-yield fare classes (Network D, Scenario 1)

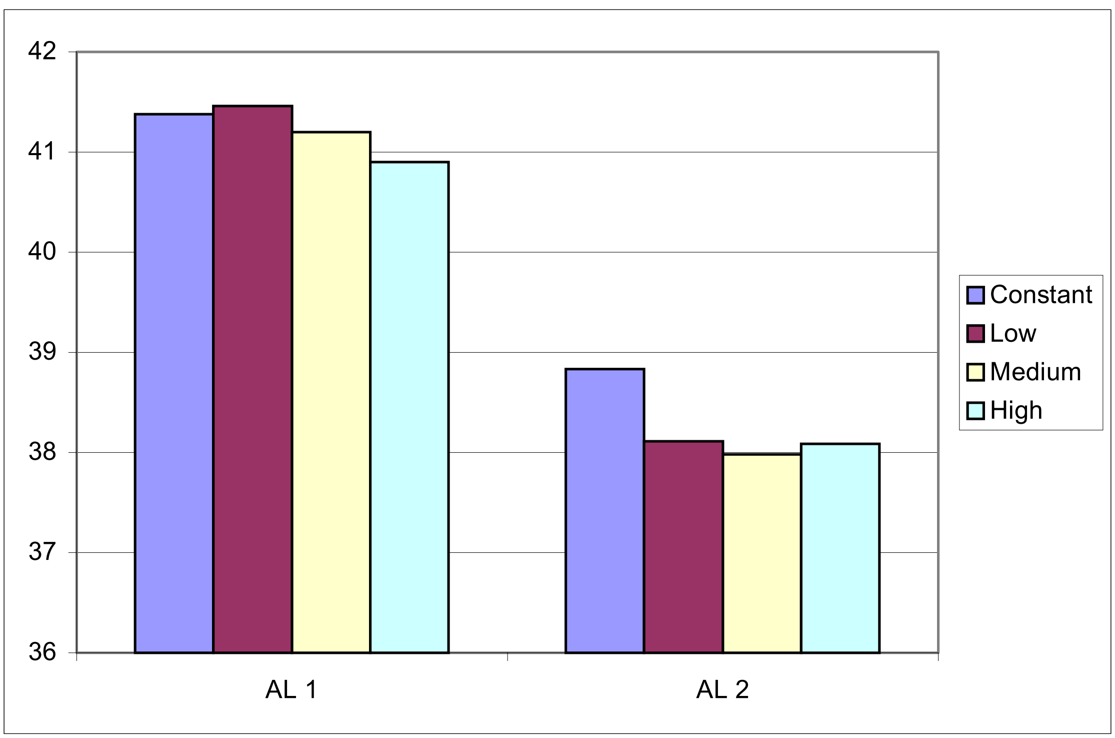


**Figure 5.10.** : Loads in the low-yield fare classes (Network D, Scenario 1)

**Network E**

Table 5.6. below shows that the change in load factor is influenced by the schedule asymmetry of Network E. Unlike in network D, airline 1 load factor does not decrease when a variable replanning disutility is introduced but increases very substantially, especially for scenarios 1 and 3. When a variable replanning disutility is used, the paths offered by airline 2 that lie far outside the decision window for a significant number of passengers appear very unattractive and more passengers prefer to travel on airline 1. This is especially true for business passengers that consider schedule convenience as a very important criterion in the choice of a path and have both shorter decision windows and higher replanning disutilities.

| | *Airline 1* | | | *Airline 2* | | |
|---|---|---|---|---|---|---|
| | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 1 | Scenario 2 | Scenario 3 |
| Constant | | 78.08% | | | 82.15% | |
| Low | 78.84% | 78.33% | 78.53% | 81.74% | 81.52% | 81.49% |
| Medium | 80.10% | 78.01% | 79.51% | 80.35% | 81.98% | 80.56% |
| High | 80.12% | 77.78% | 78.40% | 80.33% | 80.55% | 80.20% |

**Table 5.6.** : Airline load factor in Network E

The effect of schedule asymmetry is confirmed by the analysis of the fare class mix. In Network E, high-yield loads increase substantially for Alliance A partners as business passengers try to take advantage of convenient flight schedules. However, as airline 1 and 3 are unable to accommodate all of the demand, they spill low-yield passengers to their competitors and the number of low-yield passengers increases for airline 2 and 4 as shown in Figures 5.11 and 5.12. below.

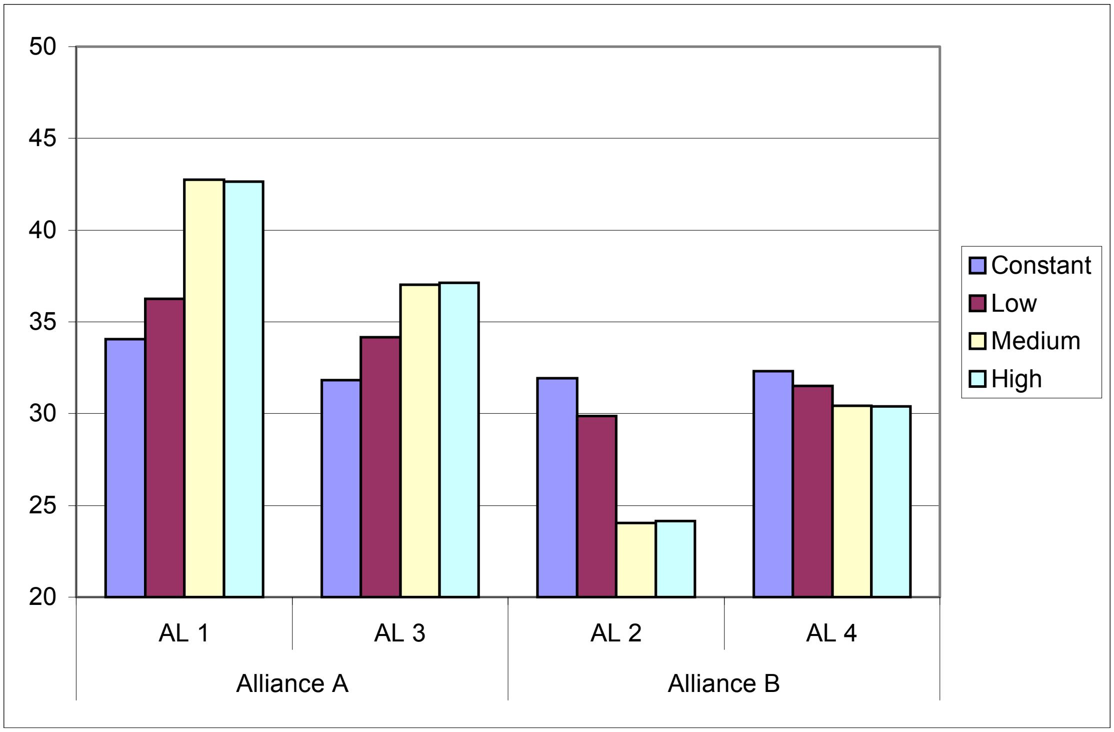


**Figure 5.11.** : Loads in the high-yield fare classes (Network E, Scenario 1)

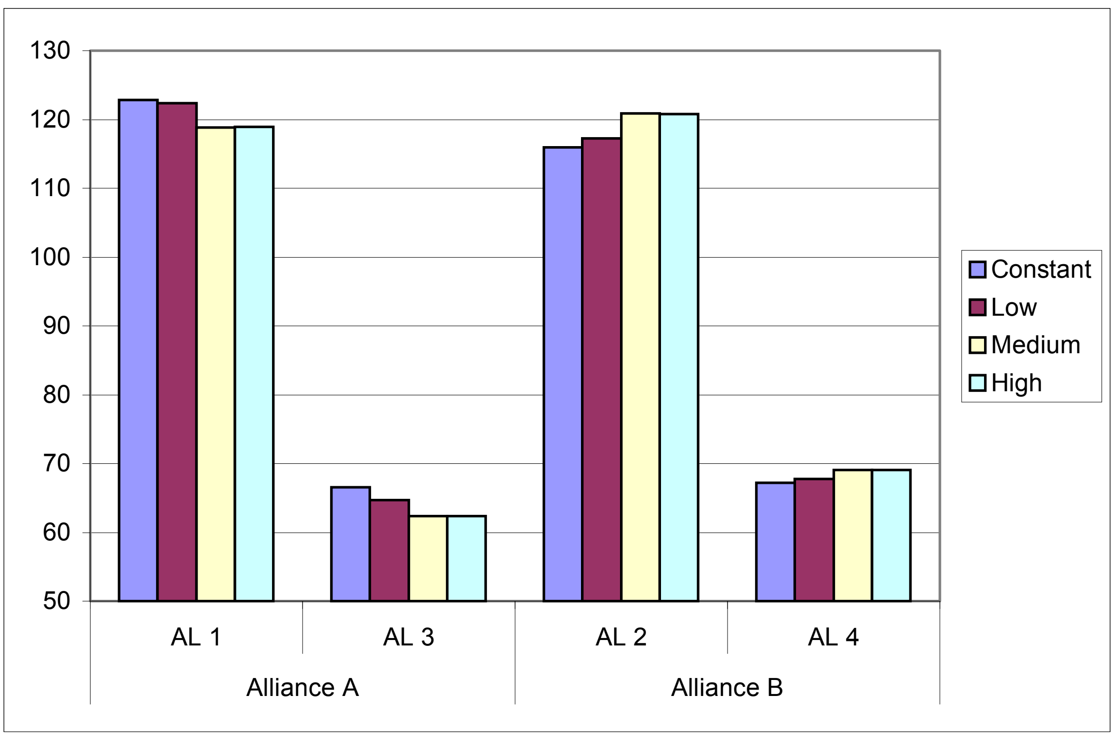


**Figure 5.12.** : Loads in the low-yield fare classes (Network E, Scenario 1)

### 5.3.2. Revenues per Category

Finally, to confirm our analysis so far, let us now look at more detailed revenue data, revenues by category in Network D and E. Since we established in the previous section that most of the change in airline revenues was associated with the introduction of a variable replanning disutility for business passengers, we will present in this section simulation results when only business passengers use a variable replanning disutility (scenario 1). Results for the other two scenarios are largely similar. For all graphs in this section, the base case is revenues with a constant replanning disutility for both business and leisure passengers.

**Network D**

Figures 5.13. to 5.16. show the change in first choice, sell-up, recapture and spill-in revenues for both airlines. For a definition of these revenue categories, the reader is referred to Chapter 4.

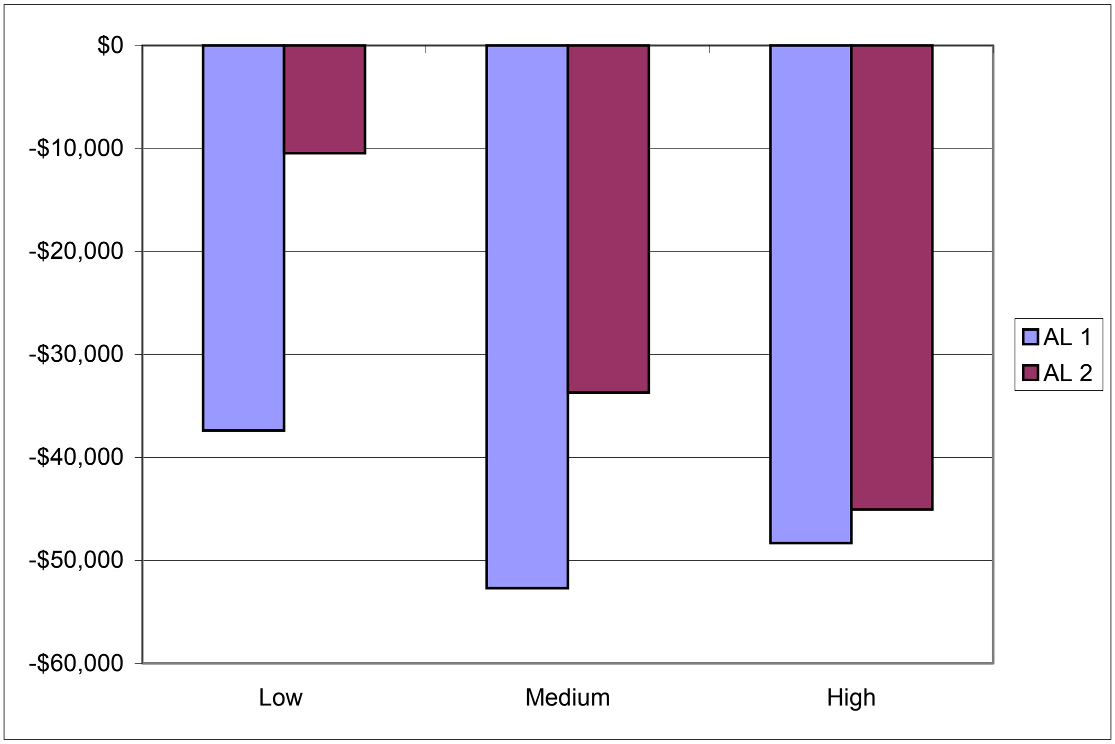


**Figure 5.13.** : Change in first choice revenues (Network D)

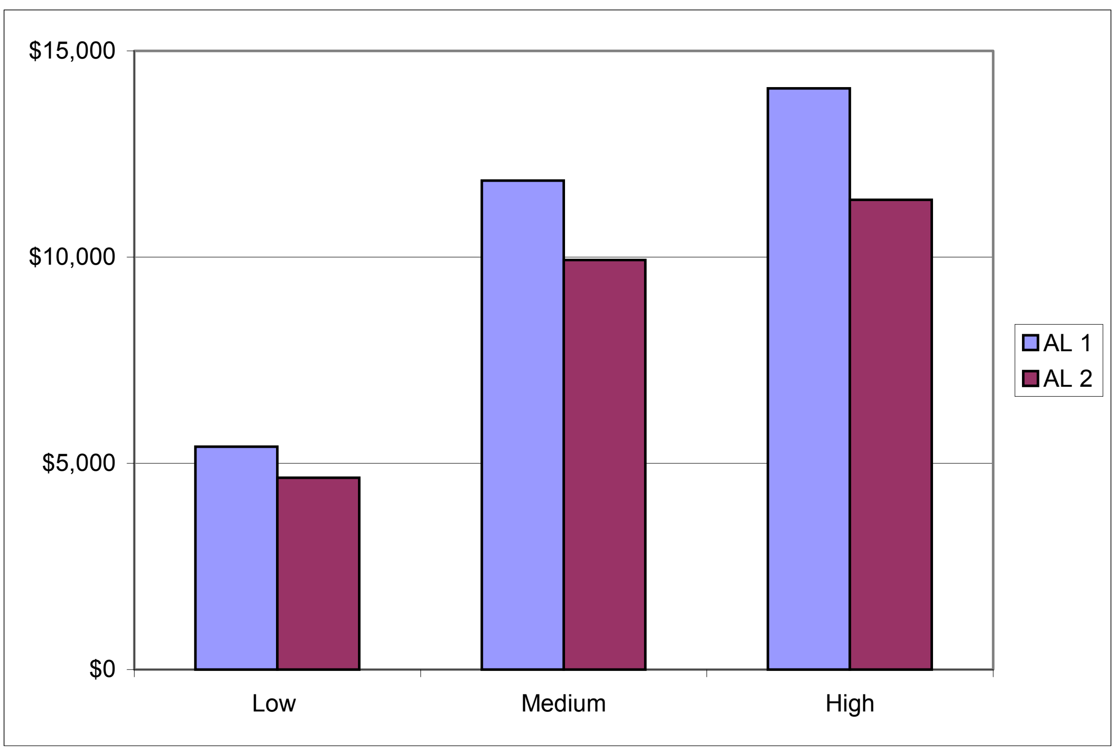


**Figure 5.14.** : Change in sell-up revenues (Network D)

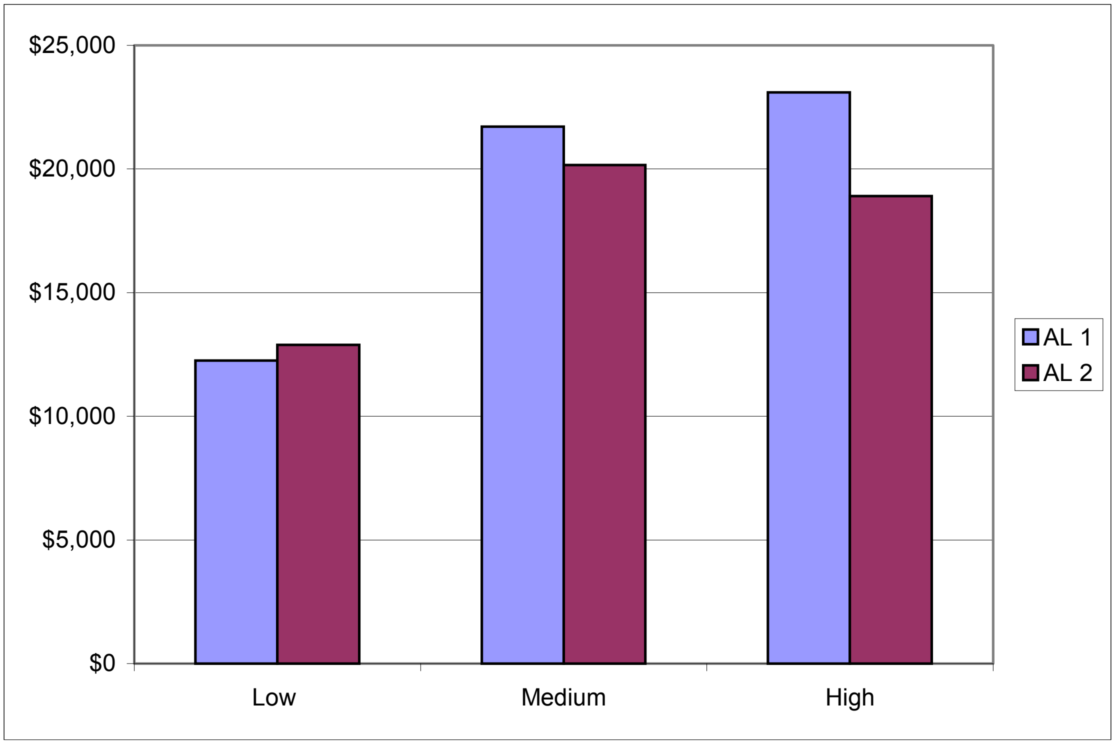


**Figure 5.15.** : Change in recapture revenues (Network D)

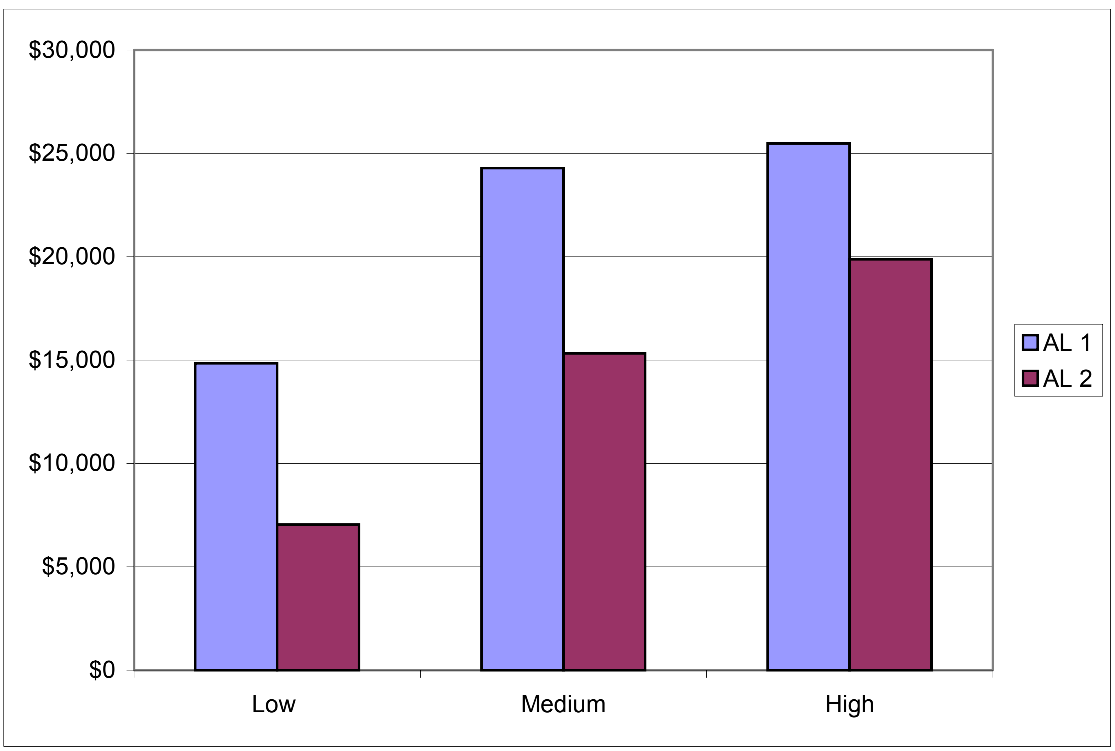


**Figure 5.16.** : Change in spill-in revenues (Network D)

Let us first consider sell-up revenues, i.e. revenues from passengers switching to a higher fare class on the same flight itinerary. As expected, sell-up revenues increase as business passengers are more willing to switch to a higher fare class to travel on a schedule-convenient path and are not willing to accept to shift to a path that is located far from their decision window due to its level of schedule inconvenience. Also, as expected, sell-up revenues are higher for the medium and high alternatives for both airlines. As the value of schedule convenience increases, business passengers are more and more reluctant to shift to alternative paths and are more willing to accept to sell-up to a more expensive fare class in order to travel on a schedule-convenient path.

In addition to sell-up revenues, recapture revenues also increase when a variable replanning disutility is introduced. While this may seem counter-intuitive, it is the result of two opposite effects. With a variable replanning disutility, business passengers are more reluctant to shift to alternative paths offered by the same airline located far outside their decision window. However, they are also more willing than before to shift to an alternative path that is located close to their decision window. When a variable replanning disutility is introduced, the second effect seems to dominate and recapture revenues increase. However, as the value of schedule convenience increases, business passengers should be less and less willing to shift to paths located outside their decision window, even if they are located relatively close to their decision window. Indeed, as we shift from the medium to the high valuation of replanning disutilities, recapture revenues increase only slightly for airline 1 and start to decrease for airline 2.

As is the case with sell-up and recapture revenues, spill-in revenues also tend to increase when business passengers use a variable replanning disutility to evaluate the schedule convenience of a path located outside their decision window. As outlined in Chapter 4, this is expected since passengers are assumed to be more willing to shift to another airline to travel on a schedule-convenient path, even if this airline is not their favorite carrier and this trend is also expected to increase with the value of schedule convenience as observed.

Finally, let us analyze the evolution of first choice revenues. Since business passengers are more willing shift to a higher fare class on the same path, alternative paths located close to their decision window or to the competitor airline, the supply of seats is constrained by the aircraft capacity and the load factor is relatively high (83%), this means that the proportion of business passengers that get their first choice satisfied will decrease and first choice revenues will be negatively impacted for both airlines. Table 5.7. below shows data for the proportion of passengers that had their first choice satisfied by airline and fare class:

| | Fare Class | Constant | Low | Medium | High |
|---|---|---|---|---|---|
| AL 1 | Y | 0.909 | 0.904 | 0.898 | 0.896 |
| | B | 0.916 | 0.907 | 0.886 | 0.879 |
| | M | 0.875 | 0.857 | 0.817 | 0.81 |
| | Q | 0.724 | 0.665 | 0.632 | 0.619 |
| AL 2 | Y | 0.91 | 0.903 | 0.896 | 0.895 |
| | B | 0.915 | 0.903 | 0.875 | 0.87 |
| | M | 0.848 | 0.809 | 0.756 | 0.745 |
| | Q | 0.733 | 0.657 | 0.629 | 0.629 |

**Table 5.7.** : Proportion of passengers that had their first choice satisfied

For all fare classes and both airlines, the proportion of passengers that had their first choice satisfied decreases as a variable replanning disutility is introduced and the value of schedule convenience increases. As a result, this leads to a decrease in first choice revenues for both airlines.

However, the decrease in first choice revenues is larger for airline 1 than for airline 2, especially for the low replanning disutility case. As mentioned earlier, airline 1 has a small schedule advantage over airline 2 due to the geographical location of its hub relative to the bulk of traffic flows in Network D. As a result, some business passengers may have in their decision window only an airline 1 path. However, for these passengers, there is probably an airline 2 path that does not fit within their decision window but is located very close to it. If a constant replanning disutility is added to the cost of all paths located outside a passenger decision window, PODS will over-estimate the cost of the airline 2 path and an airline 1 path will be the passenger's first choice, even if airline 2 is his favorite airline. However, if we use a variable replanning disutility and this disutility is assumed to be relatively low, the same passenger may now choose the airline 2 path as his first choice, especially if this path has a better fare/restriction combination, a better path quality or more frequently if airline 2 is the passenger's favorite airline. As a result, if we introduce a variable replanning disutility, more passengers will have airline 2 as their first choice and first choice revenues will decrease less for airline 2 than for airline 1, especially if the value of the variable replanning disutility is relatively low.

**Network E**

Figures 5.17. to 5.20. show the change in revenues for each airline by category in Network E when a variable replanning disutility is introduced for business passengers only.

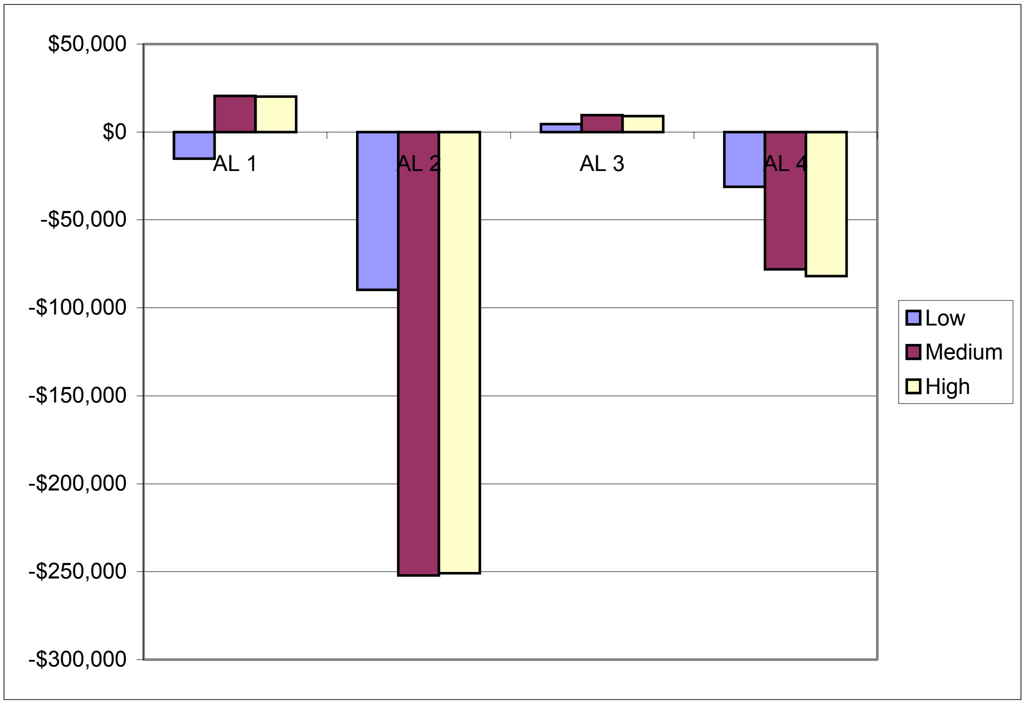


**Figure 5.17.** : Change in first choice revenues (Network E)

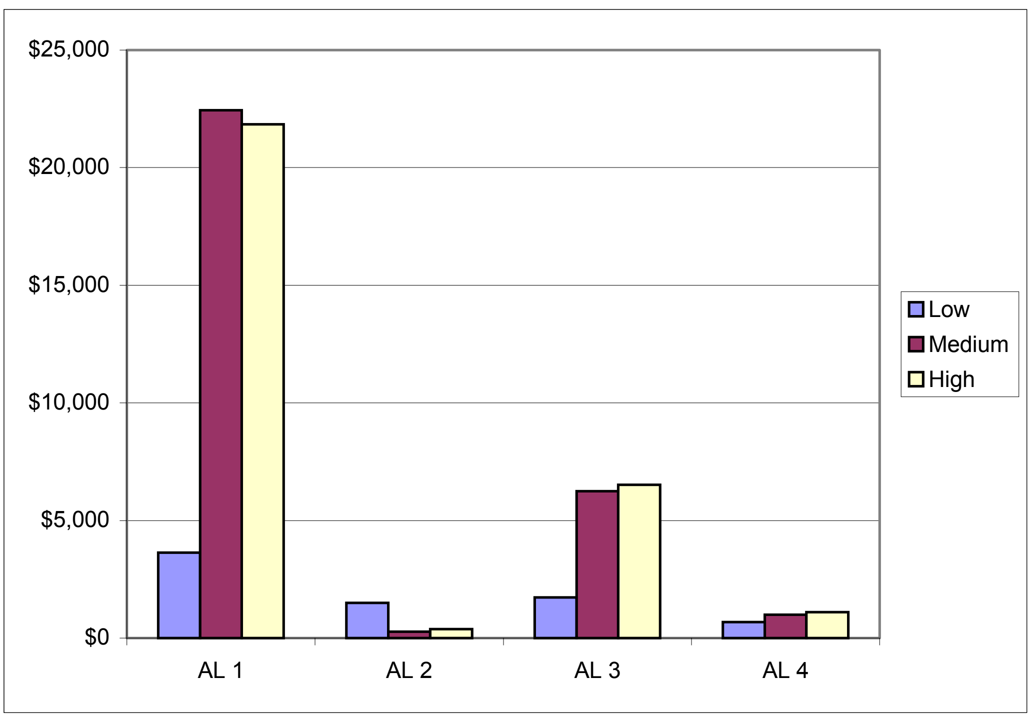


**Figure 5.18.** : Change in sell-up revenues (Network E)

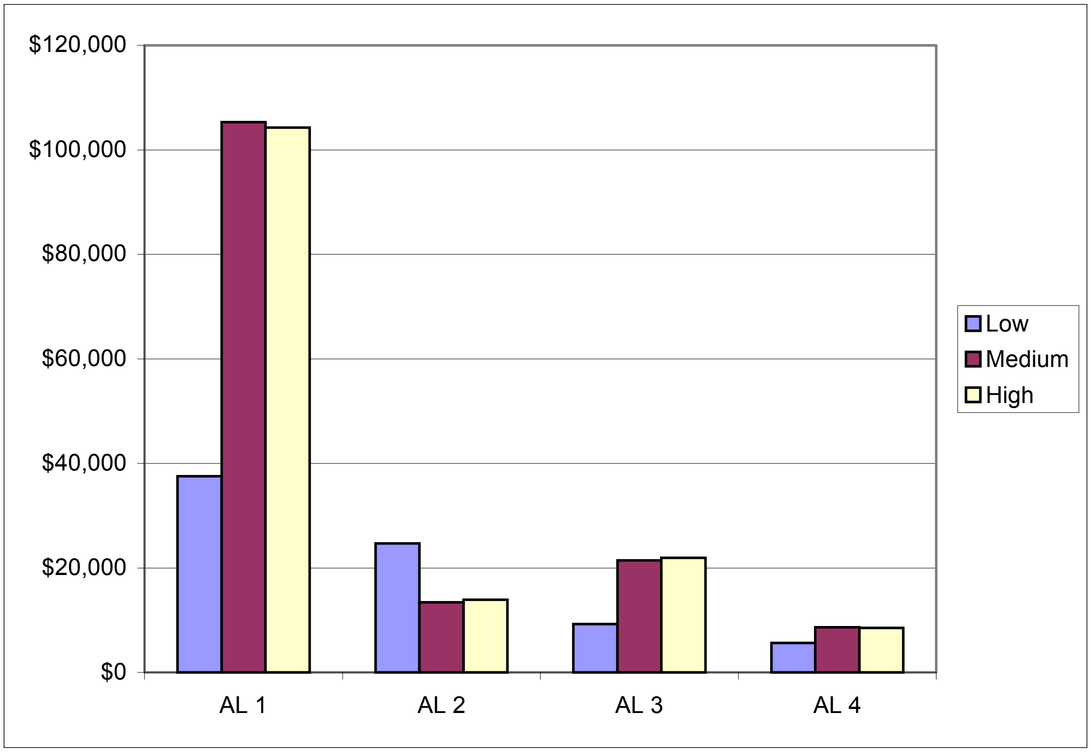


**Figure 5.19.** : Change in recapture revenues (Network E)

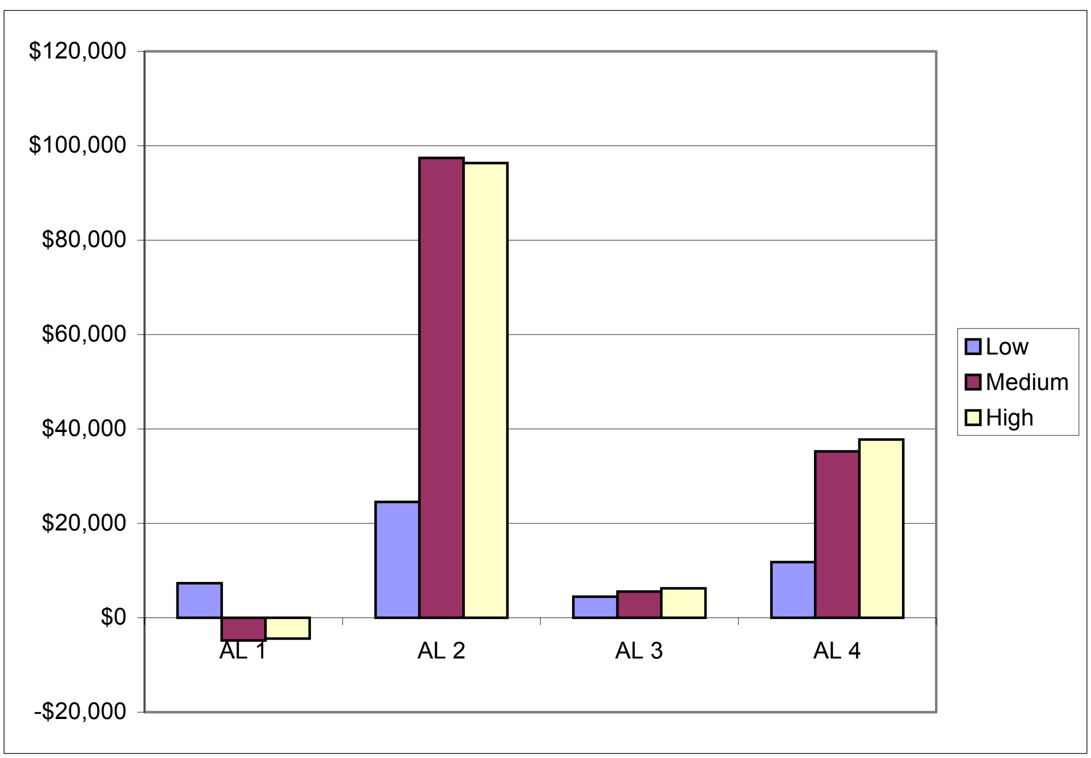


**Figure 5.20.** : Change in spill-in revenues (Network E)

Like in Network D, sell-up revenues increase for all airlines as business passengers are more willing to switch to a higher fare class in order to travel on a schedule-convenient path. With a variable replanning disutility, they prefer to pay a higher fare than travel on a schedule-inconvenient path located far outside their decision window. However, sell-up revenues also reflect the asymmetry between airlines in Network E. Sell-up revenues increase far more for the schedule-dominant airline 1 than for its weaker competitor airline 2. In addition, as schedule convenience becomes a more important criterion in passenger choice, the schedule disadvantage penalizes airline 2 even more and sell-up revenues are lower for airline 2 for the medium and high than for the low alternative. Finally, for airline 3 and 4, the change in sell-up revenues depends on their alliance partner and sell-up revenues increase more for airline 3 than for airline 4.

Also like in Network D, recapture revenues increase when a variable replanning disutility is introduced. As explained above, this increase is the result of two opposite effects. With a variable replanning disutility, business passengers may be more or less reluctant to shift to alternative paths depending on the location of the path relative to their decision window. The increase in recapture revenues suggests that for most business passengers, the location of alternative paths offered by the same airline is generally relatively close to their decision window. However, like for the sell-up category, the change in recapture revenues also reflect the structure of Network E and its schedule asymmetry between airline 1 and airline 2: Recapture revenues increase more for airline 1 thanks to its schedule attractiveness and this also influences positively recapture revenues of its European partner.

However, the influence of schedule asymmetry in Network E is the largest on the following two categories: first choice and spill-in revenues. Due to the attractiveness of its schedule relative to its competitor, airline 1 is much more attractive to a large number of passengers. A large proportion of business passengers consider it as their first choice and the introduction of a variable replanning disutility tends to strengthen that trend. As a result, unlike in Network D, airline 1 first choice revenues increase slightly, especially for the medium and high alternatives. On the other hand, as passenger preference for schedule increases under the variable replanning disutility scheme, airline 2 seems less and less attractive and airline 2 first choice revenues decrease sharply, especially for the medium and high alternatives. For airline 3 and 4, the change in first choice revenues is once again influenced by the evolution of their US partner and they increase slightly for airline 3 and decrease for airline 4 but less than for airline 2.

Faced with a very high demand, airline 1 does not have enough space to accommodate all passengers and is unable to satisfy their first choice requirements for a significant number of them. Some get recaptured and this explains the high level of airline 1 recapture revenues. Others switch to its weaker competitor and this is reflected in the high level of airline 2 spill-in revenues. In addition, the use of a variable replanning disutility tends to strengthen these trends compared to the constant replanning disutility approach. This explains the sharp increase in airline 1 recapture and airline 2 spill-in revenues, especially for the medium and high alternatives.

### 5.4. Summary

In this chapter, we have analyzed the impact of introducing a variable replanning disutility in both a schedule-symmetric and schedule-asymmetric environment. The analysis of the revenue performance at the industry, airline and category levels reveals that the impact of how passenger preference for schedule is modeled in PODS can be very significant and depends largely on the structure of the network environment.

In the next chapter, we will synthesize all the results and the lessons learned during this research and we will develop further research directions on passenger preference for schedule and traveler choice in the airline industry.

# Chapter 6 Conclusion

As mentioned in the introduction, the objective of this thesis was to review the PODS Passenger Choice Model and evaluate the relevance of its assumptions relative to the current state of the airline industry, to the issues studied by the PODS consortium and the recent advancements in consumer choice theory. In addition, as the development of a transatlantic alliance network was associated with the introduction of schedule asymmetry in PODS, we have focused as a case study on one particular component of the PODS passenger choice model, passenger preference for a flight schedule.

## *6.1. Summary of Findings and Contributions*

In the first part of this thesis, we have described the current PODS Passenger Choice Model and compared its assumptions to the models found in the consumer choice literature. We have established in Chapter 3 that the PODS generalized cost function can be compared to the specification of a mixed logit model with normally distributed independent random coefficients. As a result, the PODS Passenger Choice Model can be approximated by a series of mixed logit models, one for each market and passenger type (964 market-types in Network D). Based on data collected on actual passenger choice or through surveys simulating a booking process, it would then be possible to estimate the coefficients of the PODS generalized cost function, i.e. their mean and standard deviation using available estimation techniques.

In the second part of this thesis, we have focused on how passenger preference for schedule is modeled in PODS and its impact of the revenue performance of competing airlines both in a schedule-symmetric and schedule-asymmetric environment. We have established that the constant replanning disutility model used in PODS was not relevant to study the impact of schedule asymmetry because it tends to under-estimate or over-estimate the value of a path class depending on its location relative to the passenger decision window. As a result, we have developed an alternative model called the variable replanning disutility model that determines the value of the replanning disutility based on the deviation of each path from the passenger decision function. Based on a review of the literature, we have proposed to use a piece-wise linear function to calculate the replanning disutility of each path based on its offset from the nearest boundary of the passenger decision window.

We have then used the simulator to evaluate the impact of the variable replanning disutility model in both a schedule-symmetric and schedule-asymmetric network environment. From the results of the simulation study, we can draw the four following conclusions.

First, the variable replanning disutility model has a larger impact in PODS on business than on leisure passengers. This was expected since business travelers are generally assumed in the industry to be more sensitive to non-monetary elements like for instance fare class restrictions or schedule than leisure passengers that are assumed to be mainly concerned about fares and this behavior is incorporated in the PODS Passenger Choice Model. As a result, based on the proportion of business passengers in the traveler population in each market, one can determine whether the impact of offering a wider schedule coverage will be rather large or small.

Second, the use of a variable replanning disutility model leads to an increase in simulated revenues at the system level. Due to this more realistic representation of the schedule convenience of each path, passengers are more reluctant to travel on schedule-inconvenient paths located far outside their decision window and some of them prefer to pay a higher fare to travel on a more schedule-convenient path. As a result, simulated industry revenues in general and sell-up revenues in particular increase.

Third, the impact of the variable replanning disutility model is much larger in a schedule-asymmetric than in a schedule-symmetric environment. PODS simulation results show that the revenue advantage of offering a better schedule can be very significant, something that was largely under-estimated when the constant replanning disutility model is used. As a result, PODS could be used to estimate the potential revenue benefits of offering improved schedules or the potential revenue losses incurred when an airline decides to cut its schedule coverage at the market and network levels.

Finally, the detailed analysis of the simulation results suggests that the introduction of the variable replanning disutility model establishes a better balance between the different components of the PODS generalized cost function. For instance, with the variable replanning disutility model, some passengers are more willing to accept a flight schedule outside but close to their decision window in order to travel on their most preferred airline.

To conclude, simulation results suggest that the introduction of the variable replanning disutility model is a significant enhancement of the PODS simulator: This more realistic representation of the value of the schedule convenience of each path allows using PODS to study with more accuracy scheduling issues that are relevant to the current industry environment like for instance the impact on revenues of schedule reductions large network carriers have been implementing in the recent months.

## *6.2. Future Research Directions*

From this study of airline passenger choice and this review of the PODS Passenger Choice Model, two categories of future research directions could be explored.

First, one could investigate further the passenger preference for schedule issue. In particular, data collection of actual passenger choice behavior or surveys based on the simulation of a booking process involving the choice among several alternative flight schedules could enable researchers to calibrate the variable replanning disutility function and determine whether to use a function close to one of the three alternatives proposed in this thesis.

Furthermore, it would also be interesting to study the relative impact of the variable replanning disutility model on various revenue management methods routinely used in PODS and in the airline industry. Such a study would enable to determine whether sophisticated revenue management techniques can be useful to leverage the benefits of offering a better schedule or mitigate the revenue losses associated with a less attractive schedule.

In addition to improving the schedule component of the PODS Passenger Choice Model, further research could be useful on other parts of the choice model prior to starting significant research and simulation work on some issues involving specific parts of the passenger choice model. For instance, in order to simulate the competition between a network airline in PODS and a low-cost competitor, it might be interesting to review how fare class restrictions are modeled in PODS and if the assumptions used in the simulator are a realistic representation of the current state of the industry. Furthermore, besides preference for schedule, it might be relevant to investigate some elements of the model related to the strengths of network carriers like for instance passenger loyalty and frequent flyer programs modeled in PODS through the unfavorite airline disutility.

Finally on a more extended scale, extensive data collection on passenger booking choice behavior would enable to use advanced estimation techniques to determine the value of the coefficients of the PODS generalized cost function based on the specification of a mixed logit model with independent, normally distributed random coefficients or any other mixed logit specification.

# REFERENCES


Belobaba, P. P. (1987), *Air Travel Demand and Airline Seat Inventory Management*, Ph.D Thesis, Massachusetts Institute of Technology, Cambridge MA, 1987

Ben-Akiva M.E., Lerman S. R. (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge MA, 1985

Ben-Akiva M., McFadden D., Train K., Walker J., Bhat C.R., Bierlaire M., Bolduc D., Boersch-Supan A., Brownstone D., Bunch D.S., Daly A., de Palma A., Gopinath D., Karlstrom A., and Munizaga M.A. (2002), *Hybrid Choice Models: Progress and Challenges*, Marketing Letters, Vol. 13, No. 3, August 2002

Bhat, C.R. (2000), *Flexible Model Structures for Discrete Choice Analysis*, Handbook of Transport Modelling, Edited by David A. Hensher and Kenneth J. Button, Elsevier Science, 2000

Boeing Airplane Company (1997), *Decision Window Path Preference Methodology Description*, Seattle WA, 1997

Chandrashaker V. S. (1989), *A Study of Passenger Behavior when Denied Booking, S.M. Thesis*, Northwestern University, Evanston IL, 1989

Darot J. (2001), *Revenue Management for Airline Alliances: Passenger Origin-Destination Simulation Analysis*, S.M. Thesis, Massachusetts Institute of Technology, Cambridge MA, 2001

Douglas G.W. and Miller J.C. (1974), *Economic Regulation of Domestic Air Transport: Theory and Policy*, The Brookings Institution, Washington D.C., 1974

Hausman J. and McFadden D. (1984), *Specification Tests for the Multinomial Logit Model*, Econometrica, Vol. 52, 1984

Kanafani A. K. (1983), *Transportation Demand Analysis*, McGraw-Hill, New York NY, 1983

Lee S. (2000), *Modeling Passenger Disutilities in Airline Revenue Management Simulation*, S.M. Thesis, Massachusetts Institute of Technology, Cambridge MA, 2000

Mehndiratta S. R. (1996), *Time of the Day Effects in Inter-City Business Travel*, Ph.D Thesis, University of California Berkeley, Berkeley CA, 1996

Morrison S.A. and Winston C. (1985), *An Econometric Analysis of the Demand for Intercity Passenger Transportation,* Research in Transport Economics, Vol.2, 1985

Morrison S.A. and Winston C. (1986), *The Economic Effects of Airline Deregulation*, The Brookings Institution, Washington D.C., 1986

Prossaloglou K. E., Koppelman F. S. (1995), *Air Carrier Demand: An Analysis of Market Share Determinants*, Transportation, Vol. 22, No. 4, 1995

Prossaloglou K. E., Koppelman F. S. (1999), *The Choice of Carrier, Flight and Fare Class*, Journal of Air Transport Management, Vol. 5, No. 4, 1999

Train K. (2003), *Discrete Choice Models with Simulation*, Cambridge University Press, Cambridge MA, 2003

Walker J., Ben-Akiva M.E. (2002), *Generalized Random Utility Model*, Mathematical Social Sciences, Vol. 43, No. 3, 2002

Wilson J. L. (1995), *The Value of Revenue Management Innovation in a Competitive Airline Industry*, S.M. Thesis, Massachusetts Institute of Technology, Cambridge MA, 1995

Zickus, J. S. (1998), *Forecasting for Airline Network Revenue Management; Revenue and Competitive Impacts*, S.M. Thesis, Massachusetts Institute of Technology, Cambridge MA, 1998