Strategy in Seat Inventory Control

An empirical research at KLM on improving initial steering strategies

Public version

July 2011

W.M. Lohmann
Master thesis
University of Twente
Industrial Engineering & Management
Production & Logistic Management

KLM
Supervisor: ir. M. van de Pol

University of Twente
Supervisor 1: Dr. P.C. Schuur
Supervisor 2: Dr. M.C. van der Heijden

UNIVERSITY OF TWENTE.

Glossary

A

Aggressive strategy: a steering strategy that protects many seats in an attempt to force buy-ups and simultaneously ensure sufficient seats remain for high-yielding passengers.
ATD, availability turndown: a passenger who is denied a seat because all inventory is sold.

B

Bid price: the minimum price a passenger has to pay to get a seat on a leg.
Bucket: cluster of O&Ds that have approximately the same customer contribution on a leg.

C

Comparable challenges: a performance measurement method based on the concept of comparing units that face the same 'challenge' w.r.t. selling inventory.
Customer contribution: measure to determine the contribution of a product to the network. Calculated as the product fare reduced by the bid prices of all legs used by the product.

D

DAVN, displacement adjusted virtual nesting: RM optimization technique used by KLM.
Defensive strategy: a steering strategy that offers many cheap seats early in the booking window in order to prevent offering high-yielding late-booking passengers these seats.
Demand load factor: the relative size of demand in excess of capacity.
Dependent demand: demand for a fare class depending on the availability of other classes.
Dilution: the phenomenon that a seat is sold below its potential value.
Diversion risk: risk of passengers switching airlines when denied their preferred fare.

E

EMSR, expected marginal seat revenue: heuristic used by KLM to calculate protection levels.

F

Flight KPIs: indicators used by KLM to identify missed opportunities in inventory steering.

I

Initial steering strategy: steering strategy set by analysts six months before departure.
Inventory management: the process of selling seats in a manner that aims at minimizing the loss of potential revenue caused by spillage and spoilage.

L

LCA, lower class available: Indicator used in this research for yieldable share. LCA indicates if a lower class ticket was available when a passenger bought his ticket.
Leg: two cities between which a plane operates without landing in an intermediate city.
Local flow: passengers using only one leg (taking a direct flight).

M

Maximum demand forecast: forecasted demand with a reservation price equal to or higher than the class fare, i.e., the maximum number of passengers that might buy a fare.
MNL-model, multinomial logit model: model used to predict the likelihood of a passenger selecting a particular leg from a set of alternatives.

O **O&D, origin and destination:** the two cities between which a passenger actually travels.
Opportunity cost: performance measurement method that associates a certain penalty with each passenger that is turned down.

P **Performance monitor:** performance measurement method that compares the relative performance of different units on utilization, high-yield loss (measured by ATDs) and buy-ups (measured by RTDs).
PNV, passenger net value: forecast of the fare passengers pay for a certain product.
PoS, point of sale: geographical location where products are sold.
Priceable demand: demand consisting of passengers who buy the cheapest fare available, regardless of the associated conditions, i.e., passengers who cannot be segmented.
Protecting seats: the process of denying low fare passengers in the anticipation that their seats may be used for high fare passengers later on in the booking window.

Q **QSI, quality of service index:** indicator that represents the relative attractiveness of an airline's offerings on a certain O&D (complete description in appendix 1).

R **Recapture:** the phenomenon that denied passengers choose an alternative flight from the same airline.
Reservation price: the maximum price a passenger is willing to pay for a seat.
Revenue management: the process of optimizing revenues by asking different prices to different customers.
RFP, restriction free pricing: the pricing-situation where fares (prices) do not have enough conditions (restrictions) to generate the desired segmentation.
RTD, rate turndown: a passenger who is denied a seat as a result of protection.

S **Seat inventory control:** term used in literature for inventory management.
Segmentation: the process of distinguishing passengers with different reservation prices.
Steering strategy: the strategy w.r.t. the availability of fare classes throughout the booking window, i.e., the degree of protection that is used for a leg.
Spillage: the phenomenon that seats are sold for less than their potential value.
Spoilage: the phenomenon that seats remain unsold in the presence of demand.

T **TSFS, total seats for sale:** the amount of seats as determined by the overbooking policy.

U **Unconstrained demand:** an estimate of demand in case of infinite inventory.

Y **Yieldable demand:** demand consisting of passengers successfully segmented by fare conditions who buy their preferred fare regardless of the availability of cheaper fares.

Management summary

In the airline industry it is long recognized that different passengers are willing to pay different prices for the same seat on a plane. Substantial revenue can be captured by exploiting these differences. Airlines continuously try to serve the passengers willing to pay the highest price and quote them exactly that price: a process called revenue management.

Research motivation

At KLM, revenue management controls are calculated by an automated system. The heuristics used by the system rely on the assumption that demand is segmented, i.e., that KLM is able to force passengers to buy at the maximum price they are willing to pay. However, as this assumption has become increasingly invalid – passengers search for the cheapest ticket available – the system results (called steering controls) are not optimal. Analysts therefore adjust the steering strategy manually. Currently, no structured process exists to help analysts choose the best suited steering strategy in each situation. Specifically, to help analysts set the *initial steering strategy* six months before departure.

Research goal

This research aims to gain insight in the performance of different strategies, ultimately helping analysts to make informed decisions. The difference between strategies reviewed constitutes only their *initial aggressiveness*: the degree to which cheap tickets are made unavailable early in the booking window to force passengers to buy expensive tickets and ensure enough seats remain for high-yielding passengers who book close to departure.

Research method

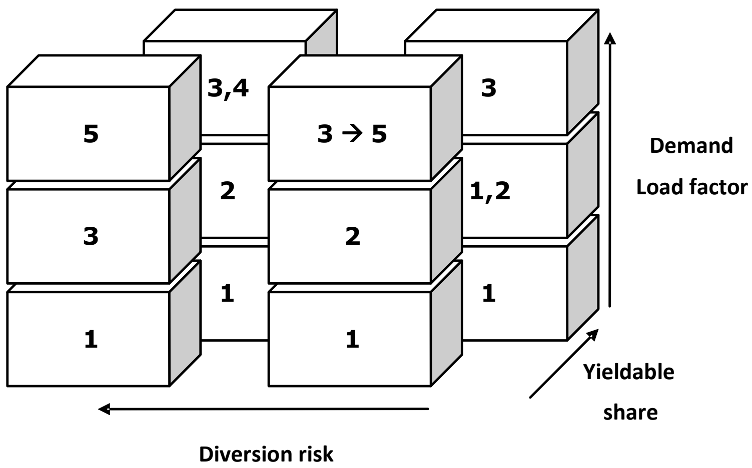
Empirical data of North Atlantic KLM flights departed in 2010 is analyzed for this research in order to study the relation between strategy and revenue management performance. The flights are scored on three indicators: the relative size of demand in excess of capacity (demand load factor), the risk of passengers diverting to other airlines (diversion risk) and the share of passengers successfully segmented by fare restrictions ('yieldable' share). Flights scoring equal on all three indicators are considered to face the same 'situation', implying that any differences in performance can be attributed to the steering strategy used.

Five different initial steering strategies are defined, the difference between them based on their aggressiveness. Performance is judged using a measurement approach developed for this research. The method measures for all flights the utilization, the number of low yield passengers forced to buy up (called rate turndowns - RTDs) and the number of high yield passengers who cannot be served (called availability turndowns - ATDs). All are measured in passengers as a percentage of capacity. Each steering strategy applied to the same situation is expected to result in a different balance between the three indicators, but the differences will not be random: more aggressive steering results in lower (or equal) utilization, fewer ATDs and more RTDs. The most desirable balance is determined by quantifying a marginal improvement on each indicator. The following figure illustrates this concept.

Strategy	Utilization	ATDs	RTDs
1. Very defensive			
2. Slightly defensive			
3. Neutral			
4. Slightly aggressive			
5. Very aggressive			

The value of each trade-off is calculated as the RTD gain + ATD gain - Utilization loss.

Results
For each situation the average indicator scores of all North Atlantic KLM flights from the dataset that were steered with the same strategy are calculated. Comparisons are made between the strategies to calculate the 'value' of switching from one strategy to another. The following figure shows the result of this analysis in the form of a strategy framework.



The numbers in this figure indicate for each situation which strategy should be applied in order to achieve the best performance (1 being the most defensive strategy). In 9 situations one strategy clearly outperforms all others, while two situations show two strategies that appear to perform equally well. In the remaining situation, (low diversion risk, low yieldable share, high demand load factor), strategy 3 and 4 seem both appropriate, while strategy 5 may also be used for days with exceptionally high demand load factors.

A sensitivity analysis is used to verify the robustness of the performance measurement method. Analyses are performed to assess the effect of incorrectly estimating ATDs and RTDs (utilization data is known to be accurate). Both analyses show no deviations from the original results when underestimating by 10% or overestimating up to 25%.

Conclusions
The strategy framework shows that relatively defensive initial strategies perform better in general. The rational is that most revenues can be captured by quoting high prices in the months closest to departure. To ensure all seats are used, a defensive initial strategy should be applied. As reported in scientific literature, increasingly defensive strategies perform better when the demand load factor decreases and when more passengers are segmented by fare restrictions. Remarkably, less risk on passengers diverting to other airlines proves to be an incentive for defensive initial strategies as well. A reasonable explanation is that flights with little diversion risk can profit even more from quoting high fares close to departure, as passengers have no other option but to buy a ticket. This generates much revenue, even though fewer seats are sold. To compensate, the initial strategy should be more defensive.

Recommendations
To improve inventory management performance at KLM, analysts are advised to apply the strategies suggested by this research. Currently, 60% of all flights is steered with the most appropriate strategy, which may be increased close to 100%. A step-by-step guide to help analysts assess each flight individually is provided at page 63. Also, future research (by KLM as well as by revenue management scientists) may benefit from the performance measurement method developed. The method turns out to be suitable, robust and insightful and is well suited for any form of empirical research on revenue management.

Table of contents

Introduction and Research design

Challenges in revenue optimization at KLM passenger transportation

1. Introduction

This thesis describes a master research conducted at KLM, a large Dutch airline carrier that was founded in 1919. The three core activities performed by KLM today are passenger transport, cargo transport and aircraft maintenance. Within passenger transportation, the pricing & revenue management department is responsible for optimizing revenues by selling seats in the most profitable manner. The department is split up in three subdivisions based on a geographical structure: one subdivision focuses on Europe, one subdivision on the North Atlantic and India and the third subdivision manages the remainder of the world. This research is conducted at the *North Atlantic subdivision*, which is organized as a joint venture: both costs and revenues are split equally between KLM, Air France, Alitalia and Delta Airlines. The team that works within this joint venture is comprised of analysts and managers from all four airlines. Its task is, put in the simplest way, to generate as much revenue as possible, given a predetermined flight schedule in the North Atlantic.

For decades already, maximizing revenues in the airline industry is no longer achieved by simply selling all available seats at a fixed price. Instead, today's airlines recognize that some passengers are willing to pay more than others for the same seat on a flight. To exploit such differences, many different prices are created, tailored towards the value different passengers assign to a seat on a flight. The trick is now to get the passengers willing to pay the highest price on board, and have them pay exactly that price (their *reservation price*). This process of optimizing revenues by asking different prices to different customers is called revenue management.

A key process in revenue management is to *segment* demand. That is, to be able to distinguish passengers willing to pay high fares from those who only want to pay low fares. Only when this is done successfully, airlines can prevent offering low fares to the high-paying segment. Traditionally in revenue management, the arrival time of passengers is used for this segmentation: passengers requesting tickets close to departure are expected to have higher reservation prices than passengers arriving months before departure. Therefore, it makes sense for an airline to let fares increase over time. Critical in this process is to determine when fares should increase. If fares increase too slowly, many cheap seats will be sold leaving insufficient seats for the high-paying demand segment: seats are sold for less than their potential value. This effect is called *spillage.* The risk on spillage can be reduced by increasing fares earlier, i.e., by denying passengers requesting a low fare their seat in the anticipation that the seats can be sold at a higher fare later on. This process is called *protecting* seats. However, if this results in fares increasing too early (seats are over-protected), some seats may remain unsold and potential revenue is lost. This effect is called *spoilage.* The revenue management department strives to protect exactly the right number of seats by increasing fares at the right pace. The formal definition of revenue management used at KLM is 'to sell seats in a manner that minimizes loss of potential revenue caused by spillage and spoilage'.

Several factors complicate the work of the pricing & revenue management department. Most importantly, the department doesn't know exactly from every passenger how much he wants to pay and when he arrives. Both when a passenger arrives earlier than expected and when his reservation price is underestimated, the passenger will find a cheaper ticket available than the most expensive ticket he is willing to buy. The difference between the available fare and the passenger's reservation price represents potential revenue that is lost. Obviously, KLM would like passengers to buy at their reservation price, even when cheaper tickets are available. Therefore, other characteristics than just time are used to distinguish demand segments as well, most notably *fare conditions*. High-paying passengers are more interested in flexibility than low-paying passengers, so flexibility can be used to distinguish both. Inflexible conditions are attached to low fares to prevent high-paying passenger from buying them. The degree to which passengers are actually separated by such conditions greatly influences the results of revenue management.

It's now clear that KLM offers several fares with different conditions attached at the same time, rather than one fare that increases over time. Changes are made both in the level of the fares and the number of seats that are available for each fare. At KLM these two tasks are performed by separate departments. The 'pricing department' sets the fares for different trips, while the 'inventory department' allocates seats for different flights. For convenience, throughout the remainder of this thesis, the same terms are used as employed by KLM. A 'trip' (the two cities between which a passenger actually wants to travel) is called an Origin-Destination (*O&D*). A 'flight' (two cities between which a plane operates without landing in an intermediate city) is called a *leg*. Therefore, an O&D (for instance Brussels-New York) may consist of several legs (for instance Brussels-Amsterdam and Amsterdam-New York).

The inventory department strives to accept and reject passengers in such a way that total revenue (for the entire network) is maximized. In fact, the department determines the degree of protection for every leg. This is a highly complex process, as it involves a large amount of variables, options and decisions. For an airline the size of KLM, optimizing this process is too difficult to be performed manually. Therefore, a revenue management system based on sophisticated heuristics is used to calculate controls on a daily basis. These controls are said to 'steer' a leg: they determine which passengers will be accepted on a leg and which will be denied. The system is designed to yield (near) optimal results, but it relies heavily on the assumption that demand segmentation is perfect: passengers will buy a fare that matches their valuation for a seat, even if cheaper seats are available at the same time. However, this may not always be the case: some passengers search for the cheapest fares regardless of their conditions. In this situation the system no longer provides optimal results, which is one of the reasons why manual adjustments by analysts are a key element in the practice of revenue management. Depending on their prediction of passenger behaviour, analysts change the number of seats that are protected for each demand segment. There is, however, no formally defined approach to this process. Each analyst has his/her own preferences regarding interventions and assumptions about passenger behaviour. As a result, different strategies may be applied to the same situation, possibly leading to sub-optimal performances.

The steering of legs could be improved by formally structuring the process of determining the right steering strategy. Unfortunately, insufficient knowledge is available on the performance of different strategies. First of all, since passenger behaviour differs between legs and changes as time develops, no strategy is always superior. Second, the performance of an approach is very difficult to measure: it's not possible to know exactly what passengers would have done, if more or fewer fares would have been available. Still, more knowledge on the performance of different strategies would be highly valuable, as it could enable KLM to better tail its steering strategies towards optimizing revenue. This thesis aims to contribute to that knowledge by identifying effective strategies, ultimately helping analysts to choose the strategy that maximizes revenue.

2. Research design

This chapter first clarifies the motivation of this research that was already mentioned briefly in the introduction. The main challenge in revenue management at KLM is stated, followed by the goal of this thesis. Next, research questions are defined along with a research scope. This chapter finalizes by discussing the research method and the outline of this thesis.

2.1 Research motivation

Exploiting the different reservation prices passengers have for the same seat on a leg is only possible when these passengers can segmented. Currently, both a passenger's arrival time and his preferences w.r.t. fare conditions are used to estimate his valuation for a seat and segment him accordingly. The success of this segmentation largely determines the success of revenue management. Unfortunately, imperfect segmentation is quite likely to be the case in today's airline industry. Passengers may buy the cheapest fare available at the moment they make a booking, regardless of the associated conditions. This behaviour has two important implications that cannot be handled by the revenue management system. First, it implies that demand for a certain fare level may consist of passengers who have that fare as their reservation price and some passengers with higher reservation prices. As demand forecasts are based on historic demand, they will be affected as well: rather than indicate the number of people that have a fare as their reservation price, the forecasts now reflect the number of people that will request a fare! As the model doesn't recognize this distinction, the amount of passengers willing to pay high fares is underestimated. Second, if both time and conditions fail to segment passengers, we're back at the original problem: KLM cannot always prevent offering cheap seats to high-yielding passengers.

So how can these situations best be handled? The first problem might be corrected by using an *aggressive strategy*: restricting access to cheap fares early in the booking window. This should force passengers to buy at their actual reservation prices, rather than the lower fare they are also willing to buy. However, if bookings come in (too) slowly, many cheap seats will still be available late in the booking window. Possibly even cheaper seats have to be made available, to prevent leaving with empty seats. Unfortunately, this will increase the second problem: late high-yielding passengers buying down (spillage). Figure 1 shows this situation.

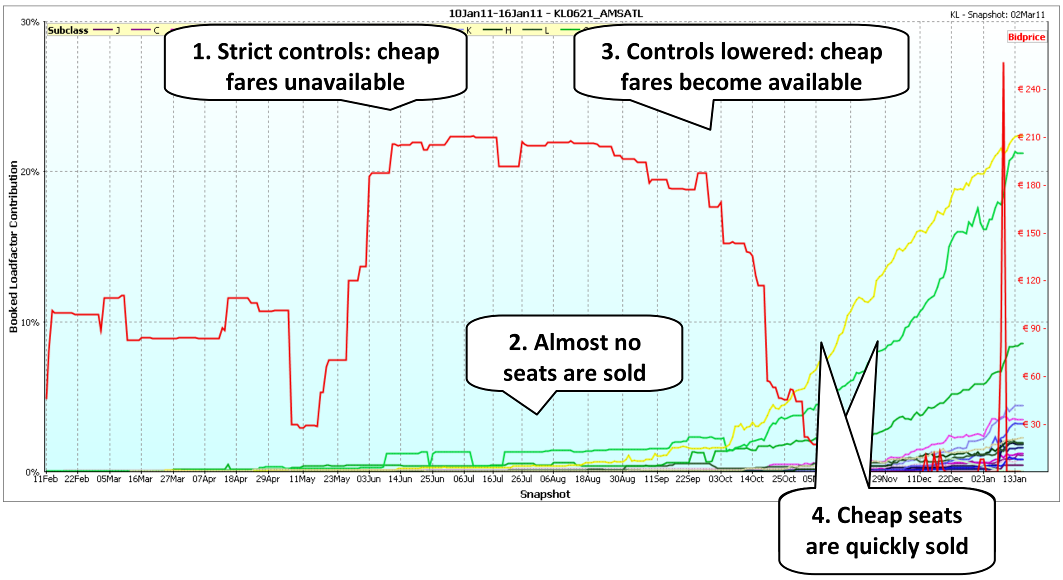


Figure 1: Example booking curves for too aggressive controls (week 2, 2011. Amsterdam – Atlanta).
See next page for an explanation.

In this figure, the red line represents the *bid price*: the minimum price a passenger has to pay to get a seat on a leg. The line gives an indication of the availability of cheap fares. The other lines are the booking curves (cumulative accepted passengers over time) per fare. The yellow and green lines represent the cheapest fares.

Protecting too many seats can thus lead to spoilage in case insufficient requests come in for the higher fares and to spillage when cheap fares are made available to late-booking passengers. On the other hand, a *defensive strategy* would be to increase access to cheap seats early in the booking window, filling up the plane quickly. Subsequently, cheap seats can be made unavailable to late booking high-yielding passengers, without risking leaving with empty seats. However, if too many cheap seats are sold to early-booking passengers, insufficient seats remain for the high-yielding passengers. Figure 2 shows this situation.

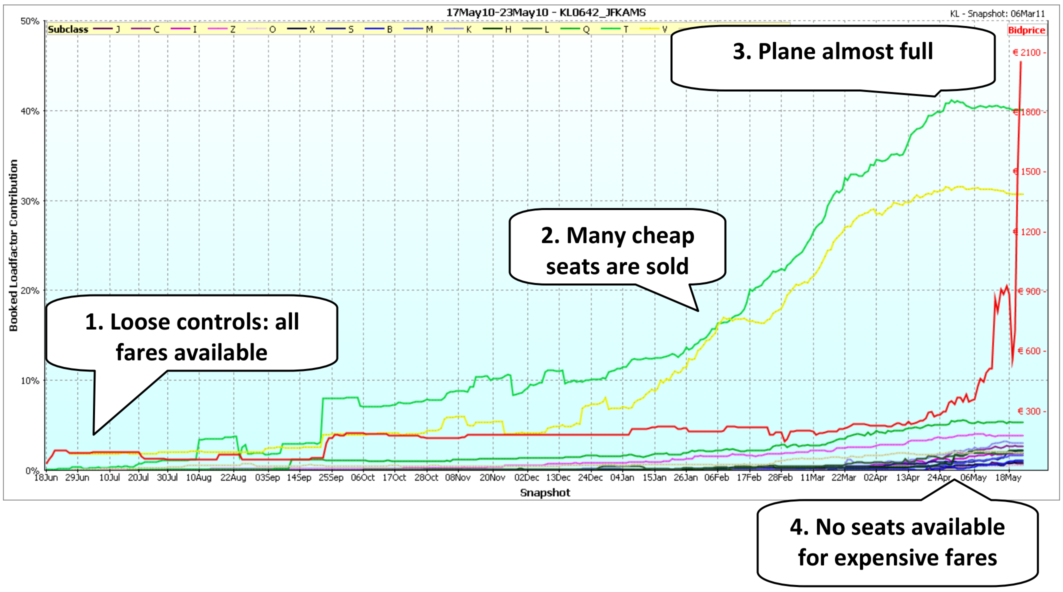


Figure 2: Example booking curves for too defensive controls (week 20, 2010. New York – Amsterdam).

Which steering strategy yields the most revenue depends on the number of potential passengers and their behaviour. How many early-booking passengers are inclined to buy-up for instance? Would restricting access for them yield enough to compensate for the risk of having to increase access again, resulting in late booking passengers buying down? Or are late-booking passengers not inclined to do this anyway? Since this passenger behaviour may differ between legs and change as time develops, there is not one strategy that is always superior. Therefore, it's the task of inventory analysts to predict passenger behaviour and choose a fitting strategy. The problem is that there is no structured process to help analysts choose the right strategy for every leg. Consequently, legs facing the same situation may be steered with different strategies, some achieving a better performance than others.

Problem statement: KLM lacks a structured process to help inventory analysts choose the best steering strategy for every leg w.r.t. maximizing revenue.

In order to assess each flight individually and apply the strategy that is expected to perform best, more knowledge is required on the performance of different strategies in different situations. This thesis aims to contribute to that knowledge, ultimately helping inventory analysts to make informed and structured decisions when choosing a steering strategy.

Research goal: to gain insight in the performance of different steering strategies in different situations in order to help inventory analysts with selecting the best suited strategy.

2.2 Research question

Recall, we established that the goal of inventory management at KLM is to optimize revenue by allocating seats to passengers (steering). A revenue management system is used to help achieve this, but, as demand segmentation is not perfect, it provides sub-optimal results. In order to improve steering performance the strategy used is adjusted by analysts based on their predictions of passenger behaviour. Currently, no structured process exists to help analysts choose the best suited *initial steering strategy*, i.e., the steering strategy analysts set six months before departure. In order to develop such a process, more knowledge is required on the performance of different strategies in different situations. The goal of this thesis is to provide this knowledge. Consequently, the main research question is:

What initial steering strategy in inventory management at KLM yields the best results?

2.2.1 Causal model

First of all, it's important to notice that a strategy affects the inventory management performance. That is, it affects how well the seat allocation balances increasing the share of high-yielding passengers and avoiding empty seats. This performance subsequently affects revenue. However, many more constructs affect revenue of which the most important are fare levels, capacity and demand. Figure 3 shows these relationships.

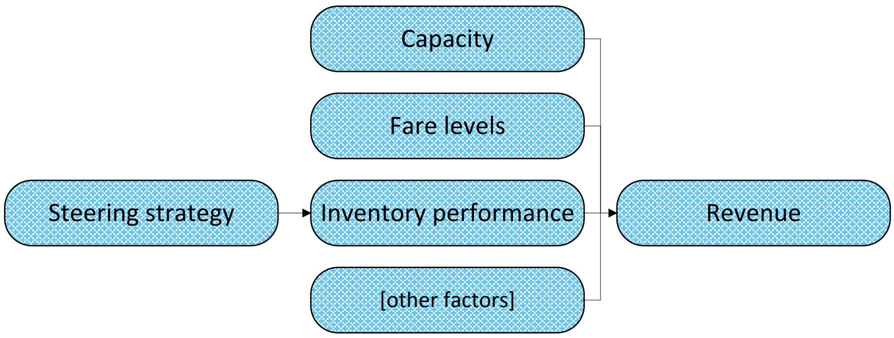


Figure 3: causal model relating strategy to revenue.

Of all these constructs, only inventory management (IM) performance can be affected by the inventory department. Constructs like demand and fare levels are beyond their control. Optimizing revenue is for this department equal to optimizing IM performance! Hence, only the direct relationship between strategy and IM performance is the focus of this research. As mentioned before, this relationship is not likely to be the same in all situations. Under certain circumstances, strategy 1 may lead to the best IM performance, while in other cases strategy 2 results in the best performance. The 'situation' does not directly affect performance, but it affects the *relation* between strategy and IM performance. Figure 4 shows the causal model that now emerges, which is central to this research.

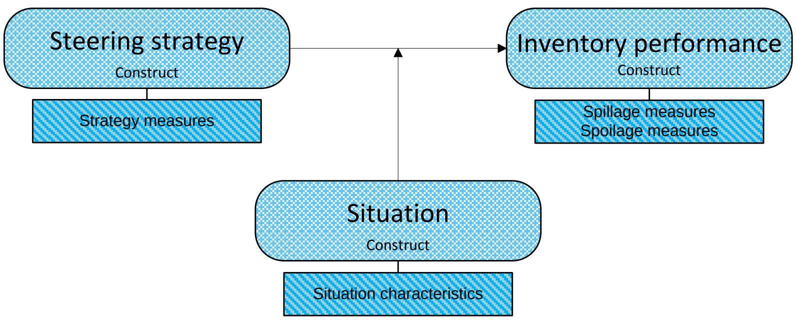


Figure 4: Detailed causal model between strategy and IM performance.

2.2.2 Sub questions

The primary objective of this thesis is to gain insight in the performance of different steering strategies in different situations. The causal model presented in section 2.2.1 is developed and tested for this purpose. A structured approach is used as many questions have to be answered before actually judging the performance of different strategies. This section briefly discusses the sub questions that are formulated for this research.

Sub question 1. How is inventory management currently handled at KLM?

An intensive study of the inventory management systems used at KLM is necessary to understand the calculations and assumptions underlying it. Additional attention is given to the opportunities analysts have for adjusting/overruling the system. Answering this sub question results in an in-depth understanding of current inventory steering practices at KLM.

Sub question 2. What influential situations are mentioned in literature and what type of strategy is suggested for each of them?

Existing knowledge on strategy and performance and the conditions that may influence their relation is the foundation for this research. An exhaustive literature review provides an overview of current knowledge. Most importantly, the conditions identified as influential are used to define different 'situations'. Combining the results of previous research allows to predict what type of strategy (aggressive or defensive) will be best in each situation. In fact, answering this sub question provides a framework for different situations and an hypothesis of the answer to the main research question.

Sub question 3. What measures can be developed to distinguish different situations and strategies and to judge performance?

First, measures are derived which are necessary both to position a flight in the framework and to determine the strategy that has been applied. Subsequently, quantitative measures are developed in order to judge the performance of different strategies. Measures from literature and existing KLM practices are compared to identify and further develop the most suited measure for all these constructs. Especially the measure of inventory management performance is a valuable result from this research in itself.

Sub question 4. For each situation, which initial steering strategy results in the best performance?

Results from different strategies are compared and analyzed using the newly derived measures in order to determine the best performing strategy in each situation. In addition, a sensitivity analysis is performed to verify the robustness of the measurement method. The result is an overview of strategies and the conditions under which they achieve the best results. This overview provides the knowledge KLM needs to make better decisions in inventory steering.

Sub question 5. How can KLM best use the results from this research?

Having established what initial steering strategies are best suited for every situation, KLM still needs guidance on how to profit from this knowledge. This final question aims to demonstrate both how inventory analysts can implement the results of this research in their daily routine and how future research efforts by KLM may benefit from methods developed by this thesis.

2.3 Research method

This thesis judges the performance of strategies that have been applied by inventory analysts at KLM to North Atlantic legs. Data on a large number of legs is used to make comparisons and find correlations. This type of empirical research is called a cross-sectional design (de Vaus, 2009). The data is collected from existing KLM databases: since the forecasting modules used by KLM need much historical data in order to provide accurate results, all historic data is stored. This vast amount of data can be used to make suitable comparisons that allow for valid claims about the performance of different strategies.

2.4 Research scope

First, this research aims to help analysts choose the best strategy when they start steering a new leg. Later in the booking window of a leg they may decide to adjust this strategy. Such adjustments are based on the pattern of incoming bookings, which cannot be predicted for an individual leg. Therefore, this research focuses on the strategy best suited to start with, six months before departure, which is called the 'initial strategy'.

Second, this research is performed at the inventory department of the North Atlantic subdivision (abbreviation AMS/RW). Only legs steered by this department are subject to data-analyses, because legs from other departments are expected to behave differently. This implies that results will not (necessarily) apply to these departments. Also, the focus on inventory management means that actions of the pricing department of AMS-RW are outside the scope of this research.

Finally, revenue management encompasses both overbooking control and seat inventory control. Overbooking aims to compensate for cancellations and no-shows (passengers who do not turn up at the airport for their flight) by selling more tickets than the number of seats available. The focus of this thesis is on seat inventory control which refers to revenue management as described in this chapter: selling seats in a manner that minimizes loss of potential revenue caused by spillage and spoilage.
Note: 'seat inventory control' is a term used in literature, while 'inventory management' is used at KLM. As overbooking is excluded, both terms have the same meaning in this thesis.

2.5 Research deliverables

KLM analysts adjust the initial steering strategy determined by a revenue management system. Their adjustments are based on predictions of passenger behaviour and personal experience. KLM needs a structural process to assess each leg individually that allows analysts to choose the most appropriate initial steering strategy. Inventory management performance and ultimately revenue will increase as more legs are steered effectively.

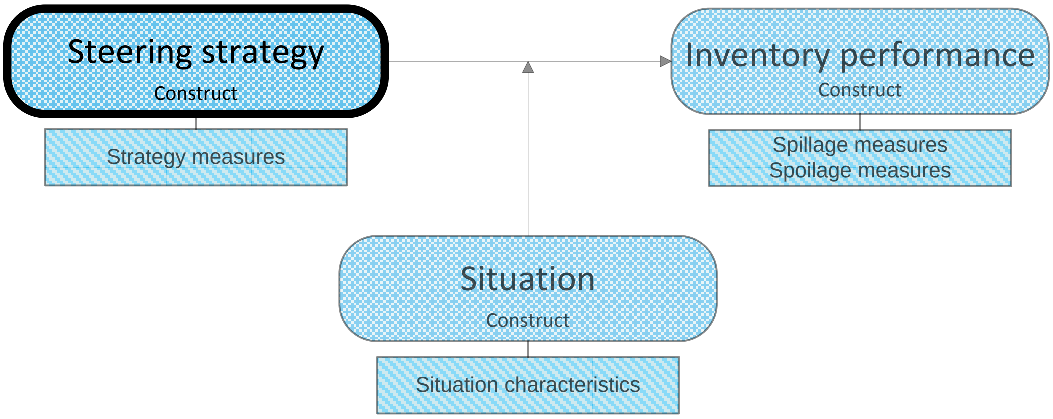
This research will benefit KLM by identifying the most effective initial steering strategies in different situations. A stepwise guide provided in this thesis helps analysts to assess legs and advises which actions to take. Legs steered using the guide will achieve a better balance between spillage and spoilage, contributing to the objective of maximizing revenue.

Secondly, this research develops a new method of performance measurement for empirical research on revenue management. A research road map is provided in this thesis that can easily be used by researchers to adapt the method developed for this research. Both KLM and RM scientists may improve their performance measurement by using the road map.

2.6 Thesis outline

The remaining chapters of this thesis address the sub questions that have been derived in section 2.2. First, chapter 3 provides an elaborate description of the inventory management systems currently used at KLM. Both the logic behind inventory management and the concept of steering by analysts will be discussed [sub question 1]. Then, chapter 4 reviews literature on the performance of different strategies in different situations. A framework for situations is derived based on the conditions that influence the relation between strategy and performance [sub question 2]. Chapter 5 develops measures to position a flight in the framework and to determine the strategy that has been applied. Also, a literature review on performance measurement is presented. Chapter 6 describes an enhanced method of performance measurement that is developed specifically for this research [sub question 3]. Then, chapter 7 discusses the results of empirical data analyses on historical KLM legs [sub question 4]. Chapter 8 shows how KLM can benefit from this research, both by improving inventory steering and by further research using the methods developed for this thesis [sub question 5]. Finally, a short summary, conclusions and recommendations are presented.

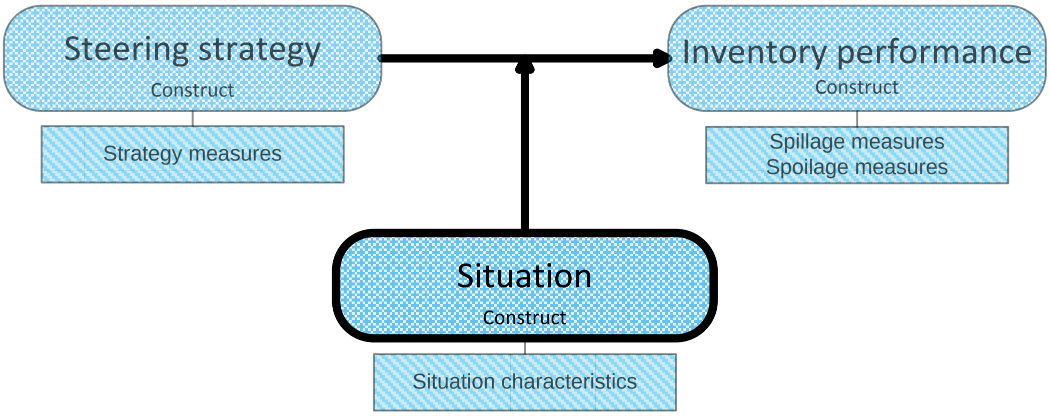
Current Steering Strategy



What steering strategies are currently employed at KLM?

This part is not available in public version

Scientific Literature



What inventory steering strategies are proposed in literature?

4. Literature review

This thesis aims to find well-performing initial steering strategies in revenue management, while recognizing that the traditional demand segmentation assumption is not entirely valid. The assumption, in fact, has become increasingly invalid since the introduction of low-cost airlines which forced large carriers to remove restrictions on their fares. As this situation has developed over the past decade, few authors have already studied the subject. This chapter summarizes their findings in order to see which effects are most relevant. Moreover, the literature review should provide an understanding of which conditions influence the relation between strategy and performance and if they imply more or less protection.

As a starting point, a research article by Weatherford and Ratliff (2010) is used, which reviews revenue management literature that incorporates *dependent demand*. 'Dependent demand' in their article means demand for a certain class depends on the availability of other classes. For example, J-class demand might be higher when class J+1 is closed compared to when that class is opened. This is exactly the situation central to this research (even though I have called this 'imperfect segmentation'. In fact, more terms for this situation exist, like 'demand leakage' or 'imperfect fencing'). Three important additional difficulties are mentioned by Weatherford and Ratliff: forecasts become conditional on the revenue management controls in place, competition affects the upsell rates and *restriction free pricing* (RFP) leads to more *dilution* risks. Restriction free pricing means that fares (prices) do not have enough conditions (restrictions) to generate the necessary segmentation, while dilution refers to seats being sold below their potential value. Again, these problems are easily translated to the challenges described in previous chapters. Conditional forecasts arise when requested fares are forecasted (rather than reservation prices), which appears to be more easy in practical situations. Competition, in my view, would be a condition that affects the relationship between RM controls and performance. Finally, increasing dilution risks refer to passengers arriving earlier than expected who find lower fares than their reservation price are available.

The articles described by Weatherford and Ratliff can be divided in two broad categories: articles providing tools to better forecast passenger behaviour and articles providing methods to better optimize RM controls. Articles regarding optimization usually adopt one or more tools of the former category to improve the forecasts used. Both categories are described as they are obviously intertwined. However, new optimization methods should provide more insight for this research than forecasting tools.

4.1 Forecasting tools for passenger behaviour

Demand forecasting with imperfect segmentation is extremely difficult: passengers may buy-down, choose another flight of the same airline or switch to a competitor. Several methods are developed to solve this problem, each providing different results: some calculate upsell rates, some maximum demand forecasts and some preferred fare demands. This section discusses three different methods: FRAT5 curves, logit demand models and Q-forecasting.

<u>FRAT5 curves</u>
A relatively simple approach to estimating upsell described by Belobaba and Hopperstad (2004) is by using a 'FRAT5 curve'. This is a negative exponential curve plotted on the booking window that gives the FRAT5 value for each point in time prior to departure. This value relates to the ratio between two fares at which 50% of lowest fare passengers are

expected to sell up when the lowest fare closes (hence the name Fare RATio 50 percent). For example, if the curve includes the point [80,1.5], this implies that 80 days in the booking window, 50% of the passengers choosing the lowest fare are expected to sell up to a fare that is 1.5 times as expensive if the low fare would be made unavailable. This value can subsequently be used to calculate upsell percentages for different fare ratios. By choosing another curve (estimating a different 50% upsell ratio), exogenous effects such as competition can be taken into account (Weatherford and Ratliff, 2010).

<u>Logit Demand models: the MNL-model</u>

A second approach uses 'logit demand models', a class of models from which the multinomial logit model (*MNL-model*) is the simplest. The aim of these models is to 'predict the likelihood of selecting a particular flight within a set of alternatives' (Gallego and Ratliff, 2009). To calculate the probability of a flight being selected, its 'attractiveness' is compared to the attractiveness of the available alternatives. To measure attractiveness, the MNL-model actually models passenger behaviour by defining parameters for attributes of a product (such as price). As the parameters reflect the importance passengers attach to the attribute, the result of the function is the required measure for attractiveness. An alternative approach is used to define how attractive the alternative of not flying/switching airlines is (for example, this measure could be based on market share). Table 5 is an example taken from Weatherford and Ratliff (2010) (OA = other airline, DNF = do not fly).

Alternative	**Price**	**Other**	**Utility**	**exp$^{U(tility)}$**	**Probability**
Leg 1, high fare	900	225	-3.9375	0.02	4.3 %
Leg 1, low fare	600	200	-2.5	0.08	18.2 %
Leg 2, high fare	900	125	-4.1875	0.02	3.4 %
Leg 2, low fare	600	100	-2.75	0.06	14.2 %
OA/DNF				0.27	60.0 %

Formula's: Utility = $\sum \beta_n * X_{in}$ Probability = $\exp^U / (\sum \exp^U)$
Parameters: β price = -0.005, β other = 0.0025, market share OA = 0.6

Table 5: Multinomial logit demand model example.

For example the utility (attractiveness) of the high fare of flight 1 is calculated as -0.005*900 + 0.0025*225 = -3.9375. The attractiveness of OA/DNF is based on the ratio between the market share of other airlines and the airline considered (60%/40%). This ratio multiplied by the sum of the attractiveness of all alternatives offered by the airline is the utility for OA/DNF: 60/40 * (0.02+0.08+0.02+0.06) = 0.27.

MNL-models can be used to calculate upsell rates, as is shown by Weatherford and Ratliff (2010). This is achieved by recalculating the probabilities of choosing a flight after an alternative has been made unavailable. In the same example shown in table 5, the low fare of leg 1 is closed, which implies that the other alternatives (even though their utility doesn't change) are *relatively* more attractive. For instance, the probability of selecting high fare from flight 1 was previously calculated as (0.02/0.45 = 4.3) and becomes (0.02/0.37 = 5.3) in the new situation. Of the 18.2 % of passengers that would have chosen low fare when it was available, 1 percentage point moves to high fare when low fare is closed. The upsell rate from low to high fare is therefore 1/18.2 = 5.5 %.

Ratliff (2008) demonstrates that MNL-models can also be used to acquire forecasts of first choice demands. Using the same logic as Weatherford and Ratliff he first calculates recapture rates. This rate is the percentage of passengers that choose an alternative of the same airline when their preferred fare is closed (and is thus '*recaptured*'). Upsell viewed in this way, is 'a special case of recapture'. Then spill is calculated as the share of passengers that cannot choose their preferred fare since it's unavailable (hence, this is *not* spillage as defined for this research). The idea that is now applied is that of 'demand mass balance'. That is, observed bookings in a class arise from first choice customers for that class being accepted and first choice customers for other classes being recaptured. First choice demand, however, consists of the same first choice accepted passengers and spilled passengers. Hence, the observed bookings reduced by the recaptures should equal the first choice demand reduced by the spill. Figure 8 illustrates this idea.

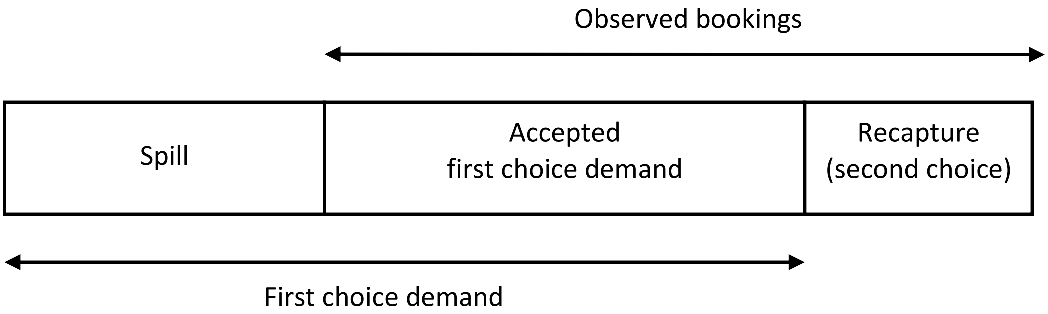


Figure 8: The concept of 'demand mass balance'.

The demand mass balance provides an equation linking recapture, spill and first choice demand (three unknown variables). Also, spill depends on first choice demand (via class availability) and recapture depends on spill (via the recapture rate). This leaves three equations with three unknown variables, which can be solved quite easily.

<u>Q-forecasting</u>
Besides FRAT5 curves and logit demand models, a third approach that can be used is Q-forecasting. This method, described by Belobaba and Hopperstad (2004), first transforms the observed historical bookings in all classes to an equivalent lowest class demand (which is usually called the Q-class). For example, with a sell-up rate of 10% to T-class and 10 observed bookings in T-class, the equivalent Q class demand would be 100. A maximum demand forecast per class can now be obtained by multiplying the sell-up rate with the total Q-class demand. If, for instance, the equivalent Q-class demand for all historic bookings would be 300, the T-class forecast (with again the 10% upsell rate) would be 30. Because the sell-up rate indicates the total number of people willing to buy up (including those willing to buy up to an even higher fare), Q-forecasting results in *maximum demand forecasts*. For any class the forecast reflects the expected number of passengers with a reservation price at the class fare level or a higher reservation price.

4.2 Optimization methods including dependent demand

This section discusses four optimization methods in chronologic order of their development: fare adjusted EMSR-B, yieldable/priceable demand models, DAVN-MR and choice-based RM models.

<u>Fare adjusted EMSR-B</u>
The first method incorporating some form of dependent demand was designed by Belobaba and Weatherford (1996). This method uses 'regular' demand forecasts (based on historical bookings) and upsell rates. Most tools mentioned in the previous section had not been

developed in 1996, so upsell rates were derived ‘either through customer surveys or statistical analysis of booking patterns’. The method, which is known as fare adjusted EMSR-B, incorporates upsell rates in the traditional EMSR-B heuristic (a slightly more aggressive variant to KLM’s EMSR-A heuristic). In short, EMSR-B protects seats for a high fare as long as the expected revenue of reserving that seat (chance of selling the seat multiplied by the high fare) is higher than the low fare. The new method corrects the chance of selling the seat via the upsell rate, which increases the expected revenue of protecting the seat. Therefore, this model will result in more protection than unadjusted EMSR-B. Also, the model recognizes different strategies need to be applied for different demand load factors (the demand as a percentage of capacity). Logically, the larger the (positive) difference between demand and capacity, the more protection should be used.

Unfortunately, some crucial but unrealistic assumptions have to be fulfilled for the method to perform well. The most important critique is that the method corrects forecasts because they are biased by imperfect segmentation, but protects as if segmentation could be perfect in the current situation (this is later called an ‘inherent contradiction’ by Gallego and Ratliff (2009)). Moreover, the model assumes passengers arrive exactly in order of low-to-high reservation price, while real arrivals are only on average from low-to-high. Finally, the model assumes passengers will always request a fare that is one class below their reservation price and subsequently can only be forced to buy up one class. As this is not realistic, the model underestimates both the true upsell potential and the dilution risk of passengers arriving early. This latest effect is resolved by Gallego and Ratliff (2009), who suggest applying EMSR-B with maximum demand inputs. In fact, this is the same model as developed by Belobaba and Weatherford, except that demand forecasts are generated using Q-forecasting. As this accounts for passengers buying up several fares, the method results in more protection than regular fare adjusted EMSR-B.

<u>Models distinguishing yieldable and priceable demand</u>

Boyd and Kallesen (2004) suggest demand may actually be hybrid: some passengers buy at their reservation price and some buy the cheapest fare available. The former category is called *yieldable demand*, the later *priceable*. This reflects the situation in which demand segmentation based on conditions is partly successful. The model Boyd and Kallesen use to determine their protection levels is not specified, but the conclusion is simple: the larger the share of priceable passengers, the higher protection levels should be. This is intuitive given their assumptions: maximum demand forecasts are used stochastically and passengers arrive in low-to-high order. Revenue is lost when fewer passengers for a certain class arrive than expected and the higher class passenger that arrives next is priceable. With a large share of priceable passengers, there is a high chance of the first arriving passenger of a class being priceable, which implies a high chance of lost revenues. Even though this is logical, the model has some serious drawbacks: first, whenever revenue management is used, demand exceeds capacity, especially for low fares. The situation of fewer low-fare passengers arriving than their booking limit is exceptional. Moreover, also this model uses the invalid assumption of strictly low-to-high arrivals. The assumption implies cheap tickets will always be unavailable when high fare passengers arrive since low fare passengers already bought those tickets. In this situation, the risk of dilution (caused by high fare passengers who buy a cheap ticket) is underestimated. The concept of dividing demand in priceable and yieldable passengers, however, is a useful insight that is further studied by Fiig et al. (2010).

<u>Displacement adjusted virtual nesting – marginal revenue (DAVN-MR)</u>

Fiig et al. introduce DAVN-MR, a method that aims at adjusting demand and fare forecasts by recognizing that demand is dependent. With these adjustments, regular RM optimization techniques can be used, such as DAVN in this example. The method, like Boyd and Kallesen (2004), assumes passengers either buy at their reservation price (yieldable passengers) or buy down as much as possible (priceable passengers). DAVN-MR requires accurate forecasts of this partitioning in yieldable and priceable passengers for each fare class. The adjusted demand forecast consists of the yieldable passengers and buy-up passengers that are calculated with Q-forecasting. That is, the adjusted demand includes passengers buying a ticket regardless of class-availabilities and all passengers that can be forced to buy op to the class. Consequently, the adjusted demand is a maximum demand forecast.

The adjusted fare is calculated as the weighted average of the yieldable fare (which is the actual fare) and the priceable fare. The priceable fare is calculated assuming 100% priceable passengers (i.e., all passengers buy the cheapest fare). Now, assume class g is the lowest open class with demand D_g and fare F_g, while class a is the highest class. If class h is opened the airline will earn $(\sum_{i=a}^{h} D_i) * F_h$ (blue area figure 9) instead of $(\sum_{i=a}^{g} D_i) * F_g$ (orange area figure 9). The difference – the incremental revenue – is thus $D_h * F_h$ (the new passengers buying the lowest fare) reduced by $(\sum_{i=a}^{g} D_i) * (F_g - F_h)$ (the existing passengers buying a cheaper fare). The adjusted fare can now be calculated as the incremental revenue per new passenger.

Adjusted fare:

$$\frac{(D_h * F_h) - (\sum_{i=a}^{g} D_i) * (F_g - F_h)}{D_h}$$

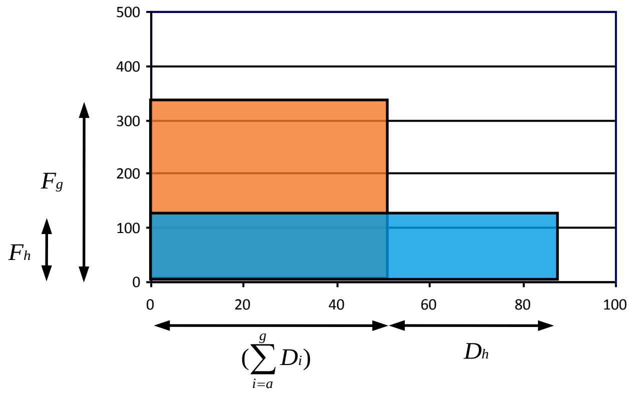


Figure 9: Priceable fare calculation.

Using the adjusted fare and demand forecasts, Fiig et al. (2010) calculate booking limits with regular DAVN. The resulting protection is higher compared to both regular fare adjusted EMSR-B and fare adjusted EMSR-B with maximum demand inputs. This is a logical result, since DAVN-MR uses maximum demand *and* corrects for 'maximum' buydown behaviour.

Also, Fiig et al. distinguish between a monopoly market and a competitive market. In the latter situation, passengers may choose, when their preferred fare is unavailable, between buying up (recapture) and switching airlines (diversion). The additional risk of losing customers to other airlines (called *diversion risk*) implies less protection compared to the monopoly case. The study concludes that 'the sell-up potential in a competitive market is lower than in the monopoly case, even with the same passenger choice characteristics'.

A disadvantage of this method is, as mentioned before, that it requires accurate forecasts of the ratio between yieldable and priceable passengers. Also, the final optimization again relies on the assumption that passengers arrive in low-to-high-order. Finally, the differentiation of passengers based on their behaviour is quite simple: either passengers buy at their reservation price or they buy-down as much as possible. In practice passengers may behave differently, for instance, buy below their reservation price but above the cheapest fare available. This situation is better attended by models including *customer choice* such as the models described below.

Choice-based EMSR

Gallego and Ratliff (2009), in the same paper as referred to earlier, introduce the first model that incorporates customer choice. That is, when passengers arrive, rather than buying their preferred fare (adjusted EMSR-B, EMSR-B with maximum demand) or choosing between the cheapest fare and their reservation price (Boyd and Kallesen, DAVN-MR), they select a flight from the set of available alternatives. This selection behaviour is modeled with an MNL-model. The heuristic Gallego and Ratliff present first calculates the maximum high fare demand when the set of possible alternatives consists only of the highest fare and switching airlines/not flying and then protects deterministically for all this demand. The method recognizes that maximum demand will not be achieved when lower fares are (partly) opened, because arriving passengers will find a different set of available alternatives. Protection is reduced by the heuristic accordingly. This model is the most realistic with regard to modeling buy-down behaviour. Not all or no passengers are inclined to buy down, but a certain percentage buys down based on the attractiveness of the high fare relative to low fares. In fact, this means fare conditions (in MNL modeled as attributes) segment passengers to some degree, a realistic assumption. However, this complicates the model significantly, which forced the authors to make another assumption: arriving passengers are 'homogenous customers with time-independent choice behaviour'. This implies that whenever a passenger arrives, he/she is equally likely to choose a high class fare, a medium fare or to refrain from flying (passengers cannot be separated *at all* by time of arrival). The result is that, with the same class-availability, dilution risk is equal across the booking window: an unrealistic situation. For example, business passengers who tend to arrive late are much less inclined to choose for not flying. Hence, more revenues are diluted by offering them a cheap fare, since they would have been recaptured in higher classes had the fare been unavailable. Besides this, the model is designed for single-legs and is not yet suited for an O&D-network.

The protection generated by the model cannot easily be compared to other methods, as protection depends on the chosen parameters of the MNL-model. Though, on average, since the model incorporates buy-up and buy-down across several classes (like DAVN-MR), more seats are protected compared to fare adjusted EMSR-B. The model also recognizes differences in the ratio between demand and capacity (called the *demand load factor*). If the demand in excess of capacity increases, fewer cheap seats need to be sold to fill the plane, which means protection increases.

Gallego and Ratliff also provide an adjustment of their heuristic to incorporate mixed arrivals. That is, passengers choosing a medium fare and passengers choosing a low fare – when faced with the same set of alternatives – arrive randomly, rather than consecutive. Gallego and Ratliff find that this situation needs less protection, because early arriving high-yielding customers 'steal' seats from the lowest bucket. With the same protection levels, low-yielding customers would be denied seats in order to protect for customers that have already arrived. However, these results refer to mixed arrivals of passengers who intend to buy their preferred product, regardless of the availability of cheaper products. This is not very insightful, as the real dilution risk arises from *priceable* passengers arriving mixed.

Choice based RM

A research that also incorporates customer choice behaviour is van Ryzin and Vulcano (2008). They assume several customers types, defined by the order in which they prefer to buy products. One of the products, again, being the switch airlines/do not fly option. For each of the customer types a demand distribution is generated based on simulations. Then, starting with a set of initial protection levels, a simulation-based algorithm optimizes protection (i.e., compares revenues generated by simulations with different protection levels until no more improvements can be found). Even though the authors simplify by assuming continuous demand and protection levels, the algorithm converges to (local) optima. The method may (not yet) be practical for large airlines, but the paper does provide insight in the consequences of different passenger behaviour assumptions. The authors present several small examples - representing different situations - in order to get ‘a qualitative understanding of the differences in the capacity control decisions that result’. Table 6 gives an indication of the forecasts that are needed as input to this model. As can be seen, this approach to forecasting passenger behaviour can account for time-dependent choice behaviour: a large advantage over the approach designed by Gallego and Ratliff.

Passenger type	**Order of arrival**	**Preference order**	**Mean demand**
Price-sensitive	1st	[flight 1-LF, flight 2-LF, OA]	40
Buy-up	2nd	[flight 1-LF, flight 2-LF, flight 1-HF, flight 2-HF, OA]	50
Price-insensitive	3rd	[flight 2-HF, flight 1-HF, OA]	20

Table 6: Example required input model van Ryzin and Vulcano (2008).
LF = low fare, HF = high fare, OA = other airline.

Three interesting examples are briefly discussed. The first (simplest) example involves one leg with two fare classes. First two passenger types arrive randomly mixed: passengers only willing to pay the low fare and passengers choosing the cheapest available fare (either high or low), i.e., priceable passengers. Then, a third passenger type arrives, always choosing the high fare. Figure 10 provides a graphical illustration of this situation. The best strategy is to sell no low fare seats at all: all priceable passengers will be forced to pay the high fare, while all low fare passengers are rejected. The resulting load factor of this strategy is significantly less than 100%. Of course, if the low fare passengers would have arrived earlier, they could have been offered seats to fill capacity. However, they cannot be segmented as they arrive mixed with the priceable passengers, which implies it’s better to protect all seats.

A second example considers two parallel flights, each having two subclasses. Three passenger types are defined, all concerning yieldable passengers. The first arriving type is willing to buy the cheapest fare on both flights. After this type is denied or accepted, the remaining two types arrive, each of them willing to buy the highest fare on one of the flights. The best strategy protects more seats than regular EMSR-B (no need for the adjusted version since there are no priceable passengers). The rationale is that the ‘penalty’ for denying a low-fare passenger is less than the low fare, because the passenger is likely to switch to the other flight. This provides an interesting insight: it makes more sense to deny a passenger his first choice when his second choice is with the same airline, rather than with a competitor.

The third example involves four parallel flights with two classes and three passenger types each: passengers always buying low fare, passengers willing to buy low or high fare depending on availability and customers always buying high fare (apparently successfully segmented by conditions), as shown in figure 10. The passenger types arrive in low-to-high order. Interestingly, even though the flights all have a different mix of passenger types, maximum demand unadjusted EMSR-B controls appear optimal. This difference with the first example is due to better segmentation: we know when the priceable passengers arrive, so we can force them to buy-up *and* offer cheap tickets to the low fare passengers.

The paper concludes that customer choice based optimization might be extremely advantageous, if - and only if - passenger behaviour can be modeled correctly. That is, if model inputs such as the example in table 6 can be accurately forecasted.

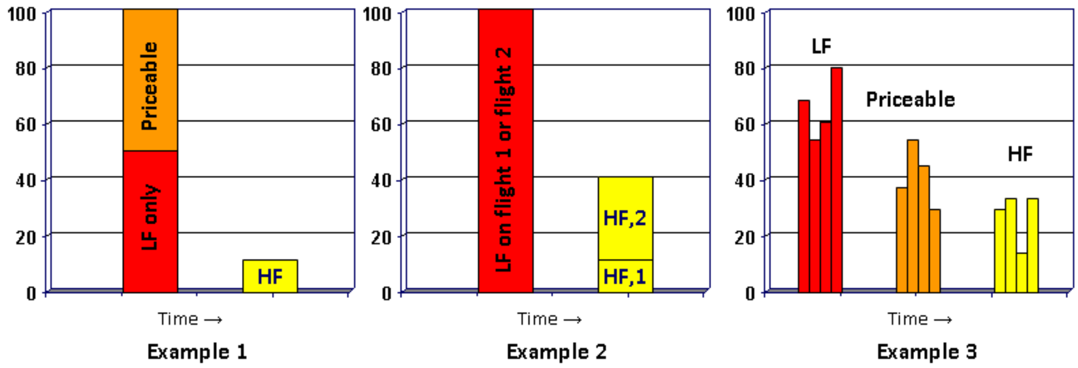


Figure 10: Arrival sequences of passenger types in three examples taken from van Ryzin and Vulcano (2008).

4.3 Conditions influencing the optimal degree of protection

The studies reviewed in this chapter provide more insight in the relation between strategy and RM performance. Several conditions have been identified that influence the optimal degree of protection. This section provides an overview of these conditions and develops a framework that can be used to define different situations.

First, the effect of mixed arrival patterns as studied by Gallego and Ratliff (2009) is omitted from the overview as the result (significant decreases in protection) contradicts all other studies. Most likely, this is because the homogenous passenger behaviour assumption is simply too unrealistic. Second, also the effect of priceable passengers arriving mixed with low fare passengers, as studied by van Ryzin and Vulcano (2008), will not be incorporated in this thesis. This is primarily because the research design employed is not suited to do so: for flights that already departed it's not possible to distinguish low fare passengers from priceable passengers who bought a low fare ticket. Also, the effect of mixed arrivals has received significantly less attention in literature, most likely because its impact is relatively small. Nevertheless, future efforts might be undertaken, for example in a case-study, to gain insight in the effect of mixed arrivals.

All other studies reviewed describe the effect of any of three factors: the share of yieldable passengers, the demand load factor and the risk of diversion. Recall the following definitions of these three factors:

- Yieldable passengers	Passengers who purchase a product because of the associated conditions, regardless of the availability of cheaper tickets.
- Demand load factor	The ratio between total demand in case of infinite inventory and the capacity (actual inventory level).
- Diversion risk	The risk of passengers switching airlines – rather than buying up – when denied their preferred fare.

As all these factors influence the relation between strategy and performance, flights can only be compared when they are equal w.r.t these factors. In this situation, the difference in RM performance can only be caused by a different strategy rather than a different situation. For this research this implies legs are considered to be in the same situation when they face comparable shares of yieldable passengers, comparable demand load factors and comparable diversion risk. Figure 11 shows a 3-dimensional coordinate system which can now be used as a framework for defining situations.

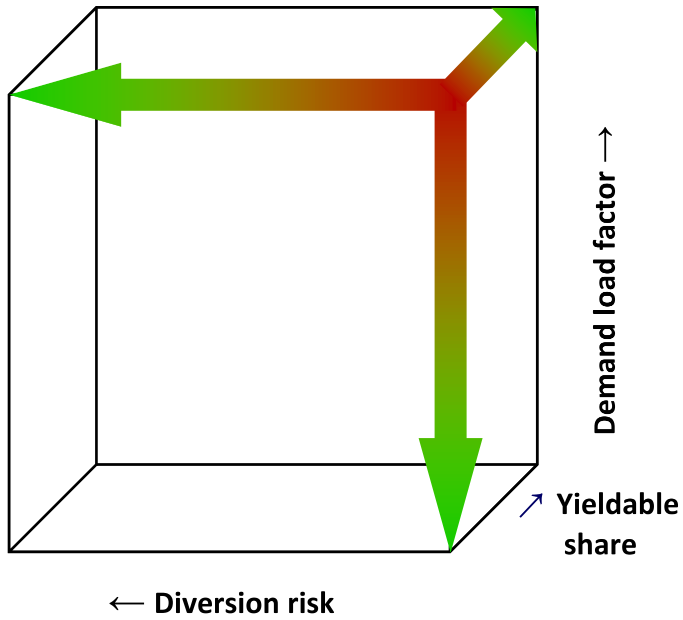


Figure 11: Framework of influential conditions and the expected changes in best suited strategies. The red side of the arrows refers to aggressive strategies expected to perform well, the green side to defensive strategies.

For each situation comparisons will be made to identify the strategy that resulted in the best RM performance (i.e., the strategy that best balanced increasing yield and avoiding empty seats). Based on the literature review conducted in this chapter, we can predict if the best strategy will be more aggressive or more defensive when changing one of the three conditions. For example, with diversion risk and demand load factor equal, the situation with many yieldable passengers will require a less aggressive approach than the situation with few yieldable passengers. The same logic can be applied to the other two conditions. The resulting predictions are illustrated in figure 11: the arrows point in the directions where increasingly defensive strategies are expected to perform best. Table 7 (next page) provides an overview for all three factors of their effect on the optimal degree of protection.

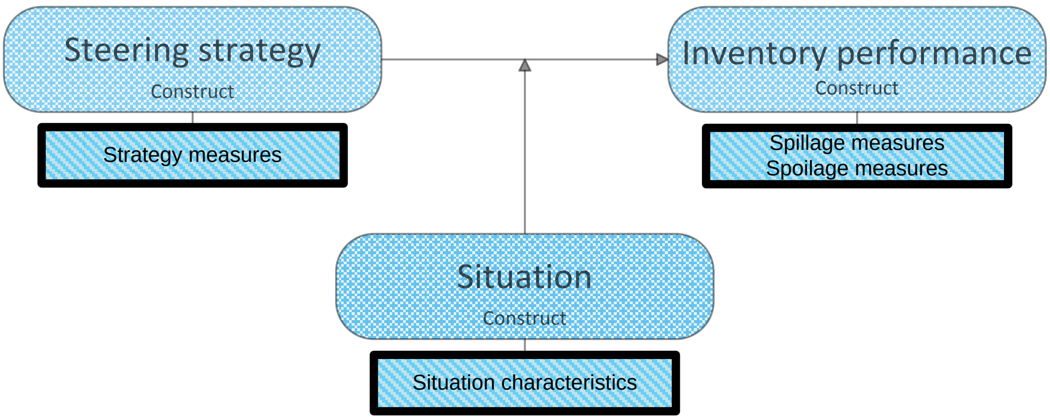
Also note that the situation which requires the most aggressive approach is not the same as the most desirable situation. Many yieldable passengers allow for a more defensive approach, because the reduced spillage risk doesn’t need to be balanced with spoilage risk. Yet, revenues are expected to be higher compared to the situation with many priceable passengers, which has to be steered more aggressively.

The empirical part of this research serves to validate the above shown model. Also, it should provide more insight in the exact degree of protection that is most appropriate in all situations. Finally, the results allow to estimate the relative importance of the conditions. Based on the knowledge gained from the literature review, we cannot predict if, for example, an increase in both demand load factor and diversion risk, requires more aggressive or more defensive steering.

Situation	Effect on the degree of protection	Studied by:
Many yieldable passengers.	This is in fact the 'old' situation in which segmentation based on conditions still performs well: passengers arriving early will not cause much spillage due to buy downs. There is no need to accept more risk on empty seats by increasing protection.	Virtually all independent demand research, Belobaba and Weatherford (1996), Boyd and Kallesen (2004).
Many priceable passengers.	Because priceable passengers are not segmented by conditions, it's very costly to have cheap tickets available when they arrive. Less availability (more protection) is a good solution as long as the decreasing risk on spillage outweighs the additional risk on empty seats.	Belobaba and Weatherford (1996), Boyd and Kallesen (2004), van Ryzin and Vulcano (2008), Gallego and Ratliff (2009), Fiig et al. (2010).
Few (or even no) demand in excess of capacity.	If demand does not exceed capacity, there is no need to protect seats for high-yielding passengers. Only minor protection may be used to avoid buy-downs.	Belobaba and Weatherford (1996), Gallego and Ratliff (2009).
Demand greatly exceeds capacity.	More passengers have to be refused in this situation, simply due to insufficient seats. Obviously, airlines prefer to deny the lowest yielding passengers and increase protection.	Virtually all RM research, Belobaba and Weatherford (1996).
Passengers inclined to switch airlines (high diversion risk).	Denying passengers their first choice ticket is useful for two reasons: because other passengers will buy more expensive tickets or in order to force the same passengers to buy up. However, if they are not inclined to buy up, but switch airlines instead, there is no point in protecting for this reason.	van Ryzin and Vulcano (2008), Fiig et al. (2010).
Passengers are likely to be recaptured on another flight of the same airline (low diversion risk).	Denying passengers who prefer to stay with the same airline (on another flight) implies less opportunity costs than denying passengers inclined to switch airlines. This can be balanced with less risk on denying a high fare passenger (increase protection).	van Ryzin and Vulcano (2008), Ratliff (2008).

Table 7: Conditions influencing the relationship between strategy and performance.
Overview developed based on literature discussed in this chapter.

Measurement



How can the constructs best be measured?

5. Deriving measures

This chapter identifies measures for all constructs in the causal model that is central to this research. The purpose of these measures is to enable quantitative comparisons between different strategies in the same situation, to assess which performed best. First, section 5.1 derives measures for the initial steering strategy, aimed to assess the degree of protection used to steer a leg. Then, section 5.2 identifies measures for all conditions derived in chapter 4 (demand load factor, yieldable share and diversion risk). These measures are used to position a leg in the situational framework. Finally, a measure for inventory management performance is developed. As this is both the most complex and the most important measure, literature is reviewed to identify and further develop a suitable measure. Interested readers can find a summary of the literature review in section 5.3. Readers may also continue with chapter 6, which discusses the measure actually used by this research.

5.1 Measuring steering strategy

This section is not available in public version

5.2 Measuring the flight situation

This section is not available in public version

5.3 Measuring the performance of a steering strategy

One of the most important questions in revenue management is if revenues actually increase by applying it, and if so, how much they increase. This is not an easy-to-measure issue as revenue management is applied in dynamic, complex real-life businesses. Still, it's highly important for all companies applying revenue management to know the impacts of the decisions and actions they take. Also for this thesis, a suitable measure for revenue management performance is a prerequisite for valuable results. This section discusses a literature review on this subject that is very insightful for the interested reader. Other readers may continue with chapter 6, which discusses only the measure used in this thesis.

We'll first briefly introduce KLM's flight key performance indicators (*KPIs*), which are currently used to assess performance. These are four indicators (spillage, spoilage, buy-up, overbooking) that indicate if an opportunity to improve performance has been missed. For example, many empty seats on a leg, while passengers were refused close to departure, indicates opportunities to reduce spoilage were not seized. However, the indicators are booleans that do not provide much foundation for quantitative comparisons. For example, what is the value of a missed spillage-opportunity, compared to a buy-up-opportunity? Also, the indicators cannot be combined to derive one overall score for the steering performance of a flight. Hence, different, more insightful measures have to be identified for this research. Due to both the complexity and importance of valid performance measurement, a scientific literature review is used to identify and compare different methods of measurement.

5.3.1 Literature review

As a starting point, we'll use an overview on revenue management (RM) literature composed by Chiang et al. in 2007. Seven notable studies are mentioned that together capture the scientific efforts undertaken on the issue of performance measurement. The difficulty of measuring RM performance is most sharply explained by Rannou and Melli (2003). They claim that both actual revenues and comparisons with competition are of no use 'as they do not allow the isolation of the actual impact of revenue management itself'. It's not difficult to think of a situation, they argue, where such indicators 'decrease due to negative market conditions, whereas the revenue management system has actually helped to generate significant additional revenues'. The logic behind this problem is clear: when several or even all inputs change, which of them caused the output to change? Other possible explanations for improvements (or losses) can be divided in two categories: event-based effects (like competition) or trend-based effects (like economic development) (Anderson and Blair, 2004; Lieberman and Raskin, 2005).

The studies pointed out by Chiang et al. use different approaches to overcome this problem. In general, four different techniques are studied, which are categorized by both Lieberman and Raskin (2005) and Chiang et al.(2007):

1. Compare differences between treatment and control groups in an experiment.
2. Compare performance before and after introduction of revenue management.
3. Calculate the percentage of the theoretical maximum revenue that is achieved.
4. Use simulation studies to estimate the effect of different policies.

Chiang et al. simply call these four approaches 'methods' and Lieberman and Raskin mention them as 'ways' to estimate incremental revenue. It's however of crucial importance to notice that only the third approach is actually a *measure*. The first two and the fourth approach are *research design*s that still require a measure of performance (for example actual revenue).

Lieberman and Raskin (2005) note that the above presented list of measurement options is ranked 'in the order of their distance from empirical measurement'. In a sense, they argue, this also approximates the desirability of using them. Indeed, a randomized experiment would provide the most valid results due to its ability to equate groups on all variables but one (the 'independent variable'). However, such an approach would be highly costly and time-consuming. As a consequence, no large-scale revenue management studies have been performed, at the time of writing, using a randomized experiment. Studies applying any of the remaining three measurement alternatives are discussed below.

Jain and Bowman (2004) use comparisons between datasets before and after the implementation of a revenue management system. They argue the difference between both is caused by a seasonal trend, a market trend and by revenue management performance. Both the effect of the market trend and the seasonal trend are removed by comparing with data sets from the same period exactly one year earlier (so both prior to adopting the new system). This method assumes trends are stable over time and external events (competitors' actions) have no influence. Lieberman and Raskin (2005) describe a method - also based on pre and post system implementation comparisons - that aims to remove other influences in a different manner. They search for equal situations (same season, same inventory availability and same market response to the company's position) to make the comparisons. In such a situation, the company faced the same 'challenge' w.r.t optimizing the sale of its inventory, hence the name '*comparable challenges*'. This approach should better capture the influence of trends. The authors admit, however, that for each event-based effect (promotions, competitors' actions) a regression discontinuity estimate is required. For many such events, the effort required for their approach may not outweigh its benefits. Also, they demonstrate their approach for renting storage facilities and make no claims of its applicability in the airline industry.

Calculations based on a theoretical maximum are implemented by Rannou and Melli (2003). The maximum revenue that can possibly be achieved is calculated after departure: all unconstrained demand is ranked from high to low value and the revenues associated with them are added until capacity is reached. This approach, obviously, relies heavily on the accuracy of the unconstrained demand estimates. Unfortunately, it’s difficult to estimate this accuracy (Lieberman and Raskin, 2005). Moreover, the method itself is rather simple and does not take into account passengers who requested tickets below their reservation price. A study by Anderson and Blair (2004) focuses on estimating maximum potential revenue as well. Their method associates a certain *opportunity cost* with all passenger requests that are turned down. For instance, when utilization is high, turning down a high value customer implies the airline has lost the opportunity of replacing the cheapest passenger with that high value passenger, which ‘costs’ the difference between the high and low fare. As all opportunity costs arise from turning passengers down, it is assumed that no opportunities are lost if no passengers are denied. However, it may be the case that passengers who are accepted would have been willing to pay more themselves! Also, the method explicitly assumes prices increase monotonically over time, while this is not always true. A third disadvantage is that accurate information on the number of passengers turned down is not easily available, especially in the airline industry.

Simulation studies are numerous in revenue management literature. The most well-known are a series of studies using the Passenger Origin-Destination Simulator (PODS). PODS is a simulator developed by Boeing and enhanced by MIT-researchers that is able to simulate a large-scale airline environment. Seven airlines provide their reservation data to PODS, allowing the simulation parameters to be increasingly calibrated towards reality. Eguchi and Belobaba (2004) studied the Japanese airline market using PODS. Although many more PODS-studies exist, this example suffices to explain the approach taken. First, Eguchi and Belobaba modified the PODS model to represent the Japanese airline industry (which is characterized by many group bookings). Then, for both the revenue management approach proposed in the study and the current approach of a Japanese airline, 20 trials are performed from which the overall revenues are averaged. The difference in average revenue is attributed to the different approach, allowing the authors to conclude that - in this case - EMSR-B seat inventory control performs better than current practices in Japan.

Other, not PODS-related, simulation studies take an equivalent approach. For example, Weatherford (2002) uses simulations to estimate revenue management performance when many different fares exist in the same booking class. He assumes customer arrivals for different booking classes can be represented with Poisson distributions. Using a large number of trials (5000) he shows that the revenue difference between his Leg Bid Price decision rule and regular EMSR-B is statistically significant. However, as with PODS, these results depend on the assumptions underlying the simulator. Different studies use different assumptions with regard to arrival times, forecast uncertainty, passenger behaviour etc., but any assumption will be a simplification of reality. The uncertainty about the validity of these assumptions is the largest disadvantage of simulation studies. An advantage is that, with small samples, the maximum revenue achievable can be calculated (in some studies simply by complete enumeration). This means simulation allows for both a relative measure between different approaches and an absolute measure of the performance of a single approach.

Finally, Chiang et al. (2007) mention a research by Blair and Anderson (2002) which doesn't neatly fit into any of the four categories. The approach called 'performance monitor' aims to compare different units – legs in the airline industry – in the same time period, as an experiment would do. Units that are subject to similar conditions (like seasonality, competition, etc.) are compared to each other. The units are not evaluated based on revenue, but on their relative performance w.r.t utilization, high yield loss and low yield availability. For instance, if two legs achieved the same utilization, but one of them had much more low yield availability, that leg was steered inferior to the other. The rationale is that the other leg, under the same conditions, refused more low yield passengers and still managed to sell equally many seats. In a sense, this approach resembles the comparable challenges method in that in compares the results of different decisions taken when faced with the same situation/challenge. (In fact, if conditions remain approximately equal over the course of a year, performance monitor can also be used to compare the same leg over time, rather than different legs at the same moment). Different to comparable challenges, the measure used by performance monitor is not actual revenue, but the balance achieved between spillage and spoilage. This balance is a better measure of performance as it is not biased by price differences. Another advantage of performance monitor is that it requires no assumptions on what levels of, for instance, utilization are strong or poor, as it uses relative benchmarks. A disadvantage is that, unlike a randomized experiment, legs can only be equated on conditions known to influence performance. There is no guarantee that these explain all differences between legs.

Summarizing the studies discussed, two fundamentally different categories of approaches can be distinguished. One category that attempts to measure the steering performance of a single flight and a category that makes comparisons to see which approach performed superior. Simulation studies might be categorized in both. Within comparison-studies the largest differences are based on the type of comparisons that is used. Some methods compare the same flight over time (pre- and post-test), some compare different flights at the same time (experiment) and some compare different instances of the same flight at the same time (simulation). Within each of these types, the methods may use different measures, such as revenue or the balance between spillage and spoilage indicators. These insights allow for a more complete categorization of performance measurement approaches than provided by Chiang et al. (2007). Figure 14 provides an overview based on this categorization, including details and disadvantages of each method. For convenience, the flight KPIs currently used by KLM are included as well.

5.3.2 Choosing the best measure

As figure 14 details, not one of the measures is without disadvantages. The most suited measure for any research effort depends on the situation at hand. For this research it's important to note that part of the value of the research stems from the fact that it is empirical. The goal is to determine which of the approaches applied by KLM actually performs best, not which should be the best in light of a set of assumptions. Simulation-based measures are therefore ruled out. Also, comparing the same flight over time will not help the purpose, as the same analyst is assigned to a flight for several years. Most likely, the same strategy is applied each time. Third, as mentioned before, due to its high consumption of resources, an experiment is not a realistic option.

This leaves the performance monitor method from Blair and Anderson (2002), backward (after departure) unconstraining as described by Rannou and Melli (2003) and the opportunity cost approach by Anderson and Blair (2004). Performance monitor seems much more appropriate than the other two options: Rannou and Melli rely too much on unconstraining accuracy, fail to mention how buy-up should be incorporated and still require a framework for making comparisons. Opportunity cost on the other hand ignores the realistic situation of a decreasing bid price and cannot incorporate buy-up behaviour either. (Third, opportunity cost requires difficult to acquire information on passenger request turndowns, but the same can be said for performance monitor).

In fact, performance monitor is the only comparison-based method that isn't based on revenue. It proposes a new measure (rather than a new research design) that allows to judge the actual steering performance between different legs. The next chapter first explains the concept of performance monitor more elaborately. Next, an adjustment is proposed that allows for more quantifiable indicators and therefore better judgment.

	Relative measurement *When comparing, which flight was better?*		Absolute measurement *How good was a single flight steered?*	
Category	**Same flight over time**	**Different flights, same time period**	**Same flight, same time**	**Theoretical maximum**
General disadvantage	Influence of time-based trends, no absolute measure	No absolute measure	Not empirical	Measurement accuracy, no comparisons
Method	**Pre-test vs. post-test**	**Randomized experiment**	**PODS**	**Backward unconstraining**
Author(s)	*Jain and Bowman (2004)*	*No examples*	*Eguchi and Belobaba (2004), many other authors*	*Rannou and Melli (2003)*
Measure	Revenue	Revenue	Revenue	Percentage revenue captured
Disadvantage	Ignores many event-based influences, can't rule out alternative explanations	High costs, time-consuming	Not empirical, relies on validity of model-assumptions	Relies on accuracy unconstraining, doesn't capture sell-up potential
Method	**Performance monitor**	**Performance monitor**	**Other simulation studies**	**Opportunity cost**
Author(s)	*Blair and Anderson (2002)*	*Blair and Anderson (2002)*	*Weatherford (2002)*	*Anderson and Blair (2004)*
Measure	Balance spillage and spoilage	Balance spillage and spoilage	Revenue	'Lost' revenue
Disadvantage	Fairness of comparisons: can't completely rule out other explanations	Fairness of comparisons: can't completely rule out other explanations	Not empirical, relies on validity of model-assumptions	Requires a lot of accurate data on turndowns and fares. Doesn't account for decreasing bid price and sell-up potential
Method	**Comparable challenges**			**Flight KPIs**
Author(s)	*Lieberman and Raskin (2005)*			*KLM*
Measure	Many measures, incl. revenue			Missed opportunities
Disadvantage	Uncertainty about equivalence of 'challenges', can't account for events			Does not provide a scale to measure performance

Figure 14: Performance measurement approaches overview.

6. Enhanced performance monitor

Performance measurement in this research is based on performance monitor, a method developed by Blair and Anderson (2002). This chapter introduces the original method and discusses ‘enhanced performance monitor’, an adjustment developed for this thesis.

6.1 The original performance monitor

Blair and Anderson argue that both revenue and ‘superior methods such as revenue per unit (RPU) fail to capture the entire picture, understand the opportunity cost and address accountability’. Therefore, performance monitor uses comparisons to assess the *value of differences between units* solely caused by RM, thereby ‘specifically looking to quantify the financial impacts of spoilage and a lack of yielding (discount limitation)’.

Performance monitor uses ‘turndown information’: data on passengers who are denied a ticket. This information turns out to be highly valuable in performance measurement, but data on turndowns may be difficult to acquire in the airline industry. This issue is discussed at the end of this section. Performance monitor uses three variables to judge performance: utilization, rate turndowns (*RTD*) and availability turndowns (*ATD*). An RTD is a passenger who is denied a certain product because seats are protected, an ATD is a denial because no seats are left. Obviously, RTDs involve low-yield passengers, while ATDs usually involve high-yield passengers. The utilization provides additional information to ‘interpret’ the appropriateness of both turndowns. For example, many RTDs in case of high utilization would imply sound policy, as the denied seats are filled with (most likely) higher yielding passengers. RTDs associated with low utilization means passengers were denied, hence forced to buy up, but not all of them actually did. The appropriateness of this trade-off depends on the situation. Many ATDs in case of high utilization reflect too little protection, while many ATDs and low utilization implies more serious mistakes (as it makes no sense at all to deny high yielding passengers to depart shortly after with an empty plane).

The method is based on relative benchmarks: units are grouped together and performance is judged within a single group. A group might involve similar units, the same unit at different points in time, or any combination deemed relevant to gain insight in performance. The original performance monitor by Blair and Anderson measures for all legs the three indicators utilization, ATD and RTD as a percentage of the capacity (for example, if 100 passengers are denied a product on a 400 seat plane for protective reasons, RTD would be 25%). Then, for all legs the indicators are set as either ‘high’ or ‘low’, reflecting respectively an above or below average score of the particular leg. The ‘quality grid’ shown in figure 15 shows the possible combinations of indicators in order of preference.

Utilisation		*ATD*	*RTD*	*Comments*	*Score*
High	Top Tier	Low	High	Yielded nicely	8
High	Top Tier	Low	Low	Could have yielded more	7
High	Mid Tier	High	High	Could have definitely yielded	6
High	Mid Tier	High	Low	Turned away high value	5
Low	Mid Tier	Low	Low	Indifferent — too much fleet?	4
Low	Bottom Tier	Low	High	Yielded — cost rentals	3
Low	Bottom Tier	High	Low	Turned away business	2
Low	Bottom Tier	High	High	Very poor	1

Better

Figure 15: Quality grid. Source: Blair and Anderson (2002).

Unfortunately, Blair and Anderson fail to mention why this is the preferred order. While some combinations are clearly better than others (the same utilization is definitely more desirable when achieved with many RTDs and few ATDs compared to vice versa), others are not unambiguous. For example, turning down many passengers, thereby forcing buy-ups while simultaneously ensuring no high yield passengers are denied, might be advantageous, even when utilization is reduced (score 3). In general the order in figure 15 seems logical, but there is no guarantee that, for instance, all scores 7 are superior to all scores 3. The next section discusses an enhanced version of performance monitor that resolves this issue.

6.2 Development of enhanced performance monitor

As admitted by the authors, the quality grid is designed with broad applicability in mind, allowing researchers to compare many different groups and gradually gain insight in performance. For instance, a single leg might be compared in a group with legs serving the same region, a group serving the same continent as well as in a group composed of the same leg in different time periods. For this purpose the general order suffices. However, for this research, a more accurate measure is needed, as the value of the research results largely depends on their reliability. Such a measure can be derived from the performance monitor concept, when taking advantage of a crucial difference between this research and the general performance monitor approach: rather than judging a single leg in many different groups, this research only judges legs that are steered under similar conditions. Assuming that indeed all relevant conditions are equal between legs, the differences that arise can only be attributed to different steering. In this situation, a ranking of all combinations, like the one provided in the quality grid, is no longer necessary. For instance, how could it be possible that turning down many passengers for protection still implies many high yield passengers have no seat left (score 6), while accepting many low-yield requests does not result in insufficient seats (score 7) ? Clearly, score 6 and 7 will not be achieved by applying different strategies to the same situation (as is the case for 4 and 8, 4 and 6, etc.). In fact, in general, it’s expected that more RTDs will leave both utilization and ATDs at most equal, possibly reduced. Also, since the difference between the five strategies identified in section 5.1 is based only on the degree of protection, we can expect strategy 2 to have more RTDs than strategy 1, strategy 3 more than strategy 2, etc. Combining these observations, figure 16 shows the expected change in indicators when applying the five strategies to similar legs.

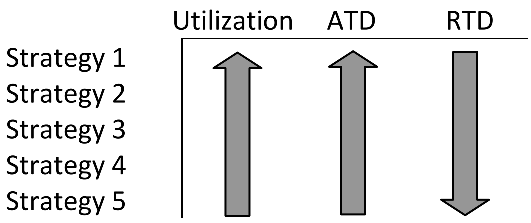


Figure 16: Change-directions performance indicators.

Like in the original performance monitor, it’s expected that every strategy results in a different combination of indicator scores (that is, a different balance between spillage (many ATDs, few RTDs) and spoilage (low utilization). However, the change in balance is not random, but in a *predictable direction*: RTDs will increase as strategies become more aggressive, possibly – but not necessarily – at the expense of ATDs, utilization or both. Whenever an improvement in RTDs (less spillage) results in a reduction in utilization (more spoilage), the change may be considered a trade-off. While the value of a balance is difficult to judge, the desirability of this type of trade-offs may be estimated.

A small example of two legs steered with two different strategies illustrates this concept. Both legs have only three fare classes: high, medium and low fare. Estimates of the costs of loss of utilization and the gain of fewer ATDs and more RTDs are based on the opportunity cost approach from Anderson and Blair (2004). They argue that when utilization decreases after applying a more aggressive strategy, apparently passengers were forced to buy up, but declined to do so. In a three-fare situation, most likely these passengers are the low fare passengers, implying that every lost passenger 'costs' the low fare. In a similar way, the passengers that were successfully forced to buy up are assumed to pay just one higher fare, say medium instead of low fare. The difference between the fares is the gain of the strategy change. Finally, a decrease in ATDs means more low fare passenger are displaced by high fare passengers, which yields the difference between those fares. The result is three simple calculations that determine the value of changing from strategy j to strategy j+1.

▪ U_j = utilization strategy j	▪ Δ U = $U_{j+1} - U_j$	▪ LF = low fare
▪ ATD_j = ATDs strategy j	▪ Δ ATD = $ATD_{j+1} - ATD_j$	▪ MF = medium fare
▪ RTD_j = RTDs strategy j	▪ Δ RTD = $RTD_{j+1} - RTD_j$	▪ HF = high fare

1. Utilization: the 'lost' passengers would have been willing to pay low fare.
   *Value: Δ U*LF*

2. ATDs: previously denied high fare passengers displace low fare passengers.
   Value: - Δ ATD(HF-LF)*

3. RTDs: Low fare passengers successfully forced to buy up pay medium fare.
   RTDs that do not force a buy up leave an additional seat available that either serves a high yield customer (fewer ATDs) or remains empty (less utilization). Successful RTDs are therefore estimated by (Δ RTD + Δ ATD + Δ U).
   Value: (Δ RTD + Δ ATD + Δ U)(MF-LF)*

Total value of the trade-off: the sum of all three calculations.
Value: (Δ RTD + Δ ATD + Δ U)(MF-LF) - Δ ATD*(HF-LF) + Δ U*LF*

Table 12 shows the data for the hypothetical example with two legs/strategies and three fare classes.

Strategy	Utilization	ATD	RTD	
Strategy 1	95 %	2 %	10 %	(Balance)
	Δ = - 3	Δ = - 1	Δ = + 15	Trade-off
Strategy 2	92 %	1 %	25 %	(Balance)

HF = 300, MF = 200, LF = 100

Table 12: Hypothetical example enhanced performance monitor.

In this example, an additional 15 % of the passengers who have a seat on the leg steered with strategy 1 are forced to buy up by strategy 2. Of these passengers, 1/15 are displaced by high-yielding passengers who had no seat in the former situation and 3/15 result in an empty seat. Apparently, the remaining 11/15 bought a more expensive ticket and is still seated in the leg steered with strategy 2. As utilization decreases, this change is considered a trade-off. The formula can be applied to see if this trade-off is desirable:

Value = 11% *(200-100) + 1% * (300-100) – 3% * (100) = 10 €/seat

This implies that by changing from strategy 1 to strategy 2, it's expected the airline will earn an additional 10 € per seat. Table 13 extends this example to an n-fare situation. Here, α_j is the lowest available fare in strategy j. γ_j is the fare difference between this lowest fare and the second lowest fare. Finally, β_j is the difference between the highest fare and again the lowest available fare α_j. Figure 17 provides a graphical illustration of these parameters.

In general, when changing from strategy j to strategy j+1, the value of the trade-off involved is estimated as follows: $\text{Value}_{j\rightarrow j+1} = (\Delta\,RTD + \Delta\,ATD + \Delta\,U)*\gamma_j - \Delta\,ATD*\beta_j + \Delta\,U*\alpha_j$.

Strategy	**Indicator values**			**Value trade-off**
Strategy j	U_j	ATD_j	RTD_j	(Balance)
	Δ U	Δ ATD	Δ RTD	Value tradeoff: (Δ RTD + Δ ATD + Δ U) * γ_1 - Δ ATD* β_1 + Δ U* α_1
Strategy j+1	U_{j+1}	ATD_{j+1}	RTD_{j+1}	(Balance)
	Δ U	Δ ATD	Δ RTD	Value trade-off (Δ RTD + Δ ATD + Δ U) * γ_2 - Δ ATD* β_2 + Δ U* α_2
Strategy j+2	U_{j+2}	ATD_{j+2}	RTD_{j+2}	(Balance)

Table 13: Outline enhanced performance monitor.

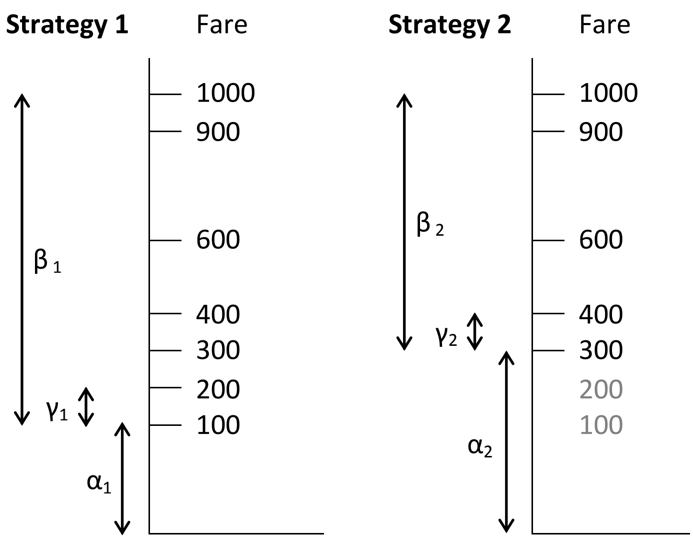


Figure 17: Hypothetical example parameters enhanced performance monitor.

The above presented enhanced performance monitor can be used to measure RM performance for this thesis. Rather than comparing two legs steered with two different strategies, two sets of legs are compared. Both sets are in the same group/situation, implying all legs are subject to the same conditions. The difference between the sets constitutes only the different strategy used to steer the legs. Within a set, one strategy is applied. In order to make comparisons, the set averages are used.

Section 6.3 estimates parameters that are necessary to apply enhanced performance monitor at KLM. Then, section 6.4 reflects on enhanced performance monitor by comparing it to current KLM practices and scientific literature. The final section of this chapter provides an overview of all measures used in this research, developed in chapters 5 and 6.

6.3 Estimating indicators and parameters

This section is not available in public version

6.4 Reflection enhanced performance monitor

The RM performance measure is a highly important part of this thesis. Incorrect or inaccurate measurement can invalidate the research results. Unfortunately, no perfect measure exists at the time of writing. The above presented enhanced performance monitor is the most suited empirical performance measure currently available. This section briefly recaptures why the method is superior to other methods.

<u>Performance monitor vs. other methods derived from literature</u>

Scientific studies on performance measurement have offered several methods that can be classified in two categories: one measuring the absolute performance of a single leg, the other comparing the relative performance of several legs to each other. Absolute measurement often relies on unconstraining: a process from which the accuracy cannot easily be estimated. Other disadvantages are the inability to capture sell-up potential and the amount of difficult-to-acquire data that the methods require.

Empirical relative measurement methods often involve a new research design that should validate using revenue as measure, rather than an actual new measure. Such methods cannot account for event-based influences and have difficulties ruling out alternative explanations. Only performance monitor (Blair and Anderson, 2002) offers an alternative to revenue, as it’s based on the balance achieved between spillage and spoilage. This measure is not biased by price-differences and therefore allows comparisons between different units.

<u>Performance monitor vs. current KLM flight KPIs</u>

The three indicators utilization, ATD and RTD in a sense reflect respectively the KLM flight KPIs spoilage, spillage and buy-up (overbooking being left out when judging inventory management performance). The difference is that, unlike KLM’s KPIs, the performance monitor indicators are relative, scalable and expressed in the same unities (customers as a percentage of capacity). Relative measurement assures good performance isn’t overshadowed by tough conditions or vice versa, scalable measures allow for more precise measurement and equal unities enable quantitative judgment of trade-offs between the indicators. In this way, performance monitor is a more advanced form of KLM’s flight KPIs.

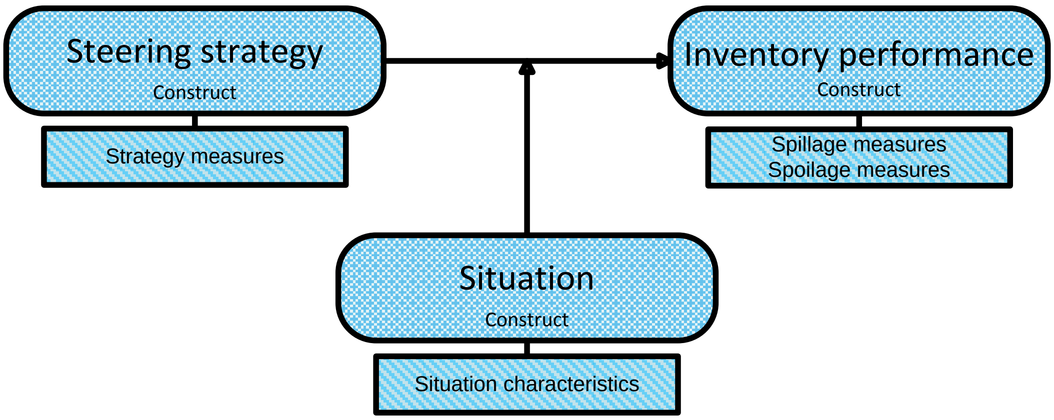
<u>Performance monitor original vs. enhanced performance monitor</u>

Final judgment in the original performance monitor is based on a general ranking of all combinations of indicators, each indicator scored as either ‘high’ or ‘low’. The ranking, however, is not necessarily optimal, as no estimates of the costs or gains of an indicator changing from high to low are made. The enhanced version does calculate such estimates, using the actual indicator values rather than a high/low classification. This extension is possible because only legs steered under the same conditions are compared: in this situation, all differences between legs can be attributed to difference in strategy. In fact, this approach combines performance monitor with the concept of comparable challenges (Lieberman and Raskin, 2005).

6.5 Overview of all measures (summary chapter 5 and chapter 6)

This section is not available in public version

Empirical research and Conclusions



Empirical data from KLM databases analyzed

7. Empirical testing

This chapter is not available in public version

8. Implementation guidelines and a future research roadmap

This chapter is not available in public version

9. Summary, conclusions and recommendations

This chapter is not available in public version

Reflection

This chapter is not available in public version

References

- Anderson, C.K. and Blair, M. (2004) Performance monitor: the opportunity costs of revenue management. *Journal of Revenue & Pricing Management*, **2**(4), pp.353–367.

- Belobaba, P. and Hopperstad, C. (2004) Algorithms for revenue management in unrestricted fare structures. *INFORMS Revenue Management Section Meeting; June 2004, Cambridge, Massachusetts*.

- Belobaba, P. and Weatherford, L. (1996) Comparing decision rules that incorporate customer diversion in perishable asset revenue management situations. *Decision Sciences,* **27**(2), pp.343–363.

- Blair, M. and Anderson, C. K. (2002) Performance Monitor. *Journal of Revenue and Pricing Management*, **1**(1), pp.57–66.

- Bootsma, P.D. (1997) *Airline Flight Schedule Development*. Ph.D. Thesis, University of Twente.

- Boyd, A. and Kallesen, R. (2004) The science of revenue management when passengers purchase the lowest available fare. *Journal of Revenue & Pricing Management*, **3**(2), pp.171–177.

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

- Desiraju, R. and Shugan, S.M. (1999) Strategic service pricing and yield management. *Journal of Marketing*, **63**(1), pp.44–56.

- Eguchi, T. and Belobaba, P.P. (2004) Modelling and simulation of impact of revenue management on Japan’s domestic market. *Journal of Revenue and Pricing Management,* **3**(2), pp.119–142.

- Fiig, T., Isler, K., Hopperstad, C. and Belobaba, P. (2010) Optimization of mixed fare structures: theory and applications. *Journal of Revenue and Pricing Management,* **9**(1-2), pp.152–170.

- Gallego, G., Li, L. and Ratliff, R. (2009) Choice-based EMSR methods for single-leg revenue management with demand dependencies. *Journal of Revenue and Pricing Management,* **8**(2), pp.207–240.

- Jain, S. and Bowman, H.B. (2005) Measuring the gain attributable to revenue management. *Journal of Revenue & Pricing Management*, **4**(1), pp.83–94.

- KLM, (2010) *[No title]*. Amstelveen, the Netherlands, KLM. Compilation of revenue management PowerPoint presentations and documents, issued between 2004 and 2010 with the purpose of employee training. Document for internal use only.

- Lieberman, W.H. and Raskin, M. (2005) Comparable challenges: a new approach to performance measurement. *Journal of Revenue & Pricing Management*, **4**(2), pp.174–184.

- Rannou, B. and Melli, D. (2003) Measuring the impact of revenue management. *Journal of Revenue & Pricing Management*, **2**(3), pp.261–270.

- Ratliff, R., Rao, B., Narayan, C. and Yellepeddi, K. (2008) A multi-flight recapture heuristic for estimating unconstrained demand from airline bookings. *Journal of Revenue and Pricing Management,* **7**(2), pp.153–171.

- van Ryzin, G. and Vulcano, G. (2008) Computing Virtual Nesting Controls for Network Revenue Management Under Customer Choice Behavior. *Manufacturing & Service Operations Management,* **10**(3), pp.448-467.

- Vaus, de D.A. (2009) *Research Design in Social Research*. Los Angeles: Sage.

- Vulcano, G., van Ryzin, G. and Chaar, W. (2010) Choice-based Revenue Management: An empirical study of Estimation and Optimization. *Manufacturing & Service Operations Management,* **12**(3), pp.371-392.

- Weatherford, L. (2002) Simulated revenue impact of a new revenue management strategy under the assumption of realistic fare data. *Journal of Revenue and Pricing Management,* **1**(1), pp.35–49.

- Weatherford, L.R. and Ratliff, R.M. (2010) Review of revenue management methods with dependent demands. *Journal of Revenue and Pricing Management,* **9**(4), pp.326–340.

Appendices

Appendices are not available in public version