Faculty of Business and Economics
Department of Transport and Regional Economics

Customized airline offer management: solving the assortment problem through multi-dimensional segmentation

Tailor-made luchtvaartaanbiedingen: het assortimentsprobleem oplossen door multidimensionale segmentatie

Thesis submitted for the degree of Doctor of Applied Economics (*Doctor in de toegepaste economische wetenschappen*) at the University of Antwerp to be defended by Daniel SCHUBERT

Supervisors: Prof. Dr. Christa Sys and Prof. Dr. Rosário Macário

Antwerp, 2025

The PhD researcher and supervisors declare that the PhD research was conducted according to the principles of scientific integrity, as mentioned in the general PhD regulations and charter for PhD researchers of UAntwerp and the integrity charter for PhD researchers and supervisors affiliated with the University of Antwerp.

Doctoral committee members

Supervisor: Prof. Dr. Christa Sys

Supervisor: Prof. Dr. Rosário Macário

Chair: Prof. Dr. Thierry Vanelslander

Jury: Prof. Dr. Amalia Polydoropoulou

Jury: Prof. Dr. Sascha Albers

Jury: Dr. Florian Martin

To my wife, Katharina

To my son, Jacob

Executive summary

This dissertation develops a novel solution for customized airline offer management with the aim to combine viability, usability, and feasibility. The solution is tested on real data from a major network airline. It expands the existing academic literature and practical applications as it suggests a cost-effective and understandable way for airlines to significantly improve the prediction accuracy of customer choice models without the need for complex models, using existing data and simple forecasts. The research shows that high-dimensional and data-driven segmentation, potentially aided by machine learning to solve data sparsity, can be combined with the traceability of discrete choice models.

Introduction

Airlines serve customers with different travel purposes and preferences, ranging from business travelers seeking a same day return flight between business cities to families with three children searching for their summer vacation or expats temporary relocating to a different part of the world for half a year. In response, airlines have customized their products, offering customers a wide selection of ancillary services on top of the mere aircraft seat. This development has been spurred by the increased competition due to the arrival of low-cost carriers (LCCs) in the early 2000s. The most popular ancillaries include baggage, seat reservation, flexibility, lounge access, increased legroom, and many more. Offering increasingly more ancillaries has enhanced the relevance of designing a customer journey that is both convenient and gives customers the option to customize the product to their specific needs. So far, airlines have solved this problem by creating *branded fares*, i.e. a static pre-selection of bundles recognizable to customers, which are aimed at solving for the most typical customer needs. Typically, airlines show three to four of these branded fares in step 1 of the customer journey. In addition, airlines offer their customers to purchase additional ancillaries *a la carte* on top of these branded fares in step 2 of the customer journey.

Due to the static pre-selection, some of the branded fares include ancillaries that are most likely not relevant to a particular customer making a search. For example, a business traveler browsing for a same day return flight is most likely not interested in check-in baggage. However, they might be interested in fast-track through security or other less common ancillaries, which they only find in step 2 of the customer journey in a long list of ancillaries. This is not the most convenient customer experience. Can airlines not be smarter than that, and offer a customized bundle option that is tailored to the specific search in step 1 of the customer journey?

This is the research objective this dissertation aims to address. It is built on the hypothesis that any search for an airline product implicitly and explicitly provides information to the airlines that they can use to respond with a targeted, customized bundle in addition to the static branded fares.

Problem statement

Due to its relevance, this research objective has received ample attention both from academic and industrial scholars. Solving the customized offer management problem requires solving the various subproblems of **segmentation** (how can customers be grouped into distinct segments that behave differently from each other?), **bundling** (how to create a bundle in real-time when a customer makes a search?), **assortment** (which of all possible products to display to a customer in step 1 of the customer journey?), and **pricing** (how to price the flight, the ancillaries, or the entire bundle?).

So far, existing literature and practical applications can be split into one of two categories. Either, they use discrete choice models catered to a single-digit number of segments. Most often, they distinguish the two classical segments of “Business” or “Leisure” travel, using the length of stay as the only criterion for segmentation These discrete choice models are simple to comprehend. However, they fail to capture the whole spectrum of customer preferences, and their prediction accuracy suffers from ignoring relevant information, such as when a customer searches or which channel the search comes through. Alternatively, they use machine learning models that can handle a much higher number of features and hence build significantly more segments, up to infinite segmentation where a customer search is not mapped to exactly one segment at all. These models are capable of catering to customer needs in a much broader way. However, they are typically difficult to understand as many complex machine learning models suffer from their black box character. When airline users cannot relate to how a model comes up with its segmentation, then the airline typically struggles with achieving adoption.

This is the research gap this dissertation aims to fill. It develops and validates a novel solution that delivers on both higher prediction accuracy than existing discrete choice models, and usability due to a simple, transparent, and understandable logic to link searches to segments. It does so by segmenting customer searches in a data-driven way into thousands to millions of segments, based on data included in the customer searches and hence available to airlines at no additional cost.

Methodology

This dissertation starts with outlining a conceptual new Offer Management System (OMS) that holistically solves the problems of segmentation, bundling, assortment, and pricing. The proposed OMS is built on two fundamental hypotheses that airline customers can be segmented into distinct segments that exhibit different purchase behavior, and that this segmentation can improve the prediction accuracy of customer choice probabilities for customer searches in the future.

Both feasibility and viability of the proposed OMS are validated with real airline data as the author had access to hundreds of millions of booked airline coupons between September 2018 and September 2023. To test the two fundamental hypotheses, the validation is split into two parts and uses both inductive and deductive research. First, the inductive research makes empirical observations in the data, then generalizes these to find patterns, and develops a theory based on this generalization. The deductive research tests this new theory on new data and confirms its validity with new observations.

Key findings and results

This dissertation proves the proposed OMS can predict offers relevant to a specific search with significantly higher accuracy than existing discrete choice approaches. The high statistical significance is established with three different error metrics and for five different time periods. Specifically, the results show that – in this order of importance – sales channel (e.g., airline's own website, or travel agency), customer loyalty status (different membership tiers, or no member), whether a flight is scheduled to run overnight or not, and booking weekday all have explanatory power which product, i.e. branded fare plus ancillaries, a customer purchases. The dissertation confirms that including these features increases the prediction accuracy of future customer choices with a statistical significance of 99.9%. Hence, it highly significantly helps airlines identify which customized bundle most likely resembles the specific customer's preferences.

At the same time, the segmentation, despite its high dimensionality with thousands of segments, is arguably much more understandable than complex machine learning models. This is because every search is mapped to precisely one customer segment based on the sales channel, customer loyalty status, whether the flight runs overnight or not, and the weekday of the booking. Moreover, experiments on different time periods confirm that changes in customer behavior can be captured with monthly model retraining. Lastly, the proposed OMS does not require customers to reveal their personal identity, hence is in full compliance with data privacy regulation.

From a methodological perspective, the dissertation tests four prediction models with increasing complexity. This has yielded an unexpected result. 99.9% of the improved prediction accuracy can be achieved with simple forecasts that assume the observed behavior in the training period continues for the future. More complex prediction models, using a novel application of the well-studied matrix factorization algorithm that has been proven extremely powerful in several applications including recommendation engines at Netflix and Amazon, only improved prediction accuracy for the last 0.1% of customer searches. These are the ones that map into infrequent segments that represent an infrequent combination of features like a rare sales channel, a rare loyalty status and an overnight flight.

Conclusions and implications

The results show that airlines do not need complex machine learning models to improve the prediction accuracy of which specific product a specific customer will likely purchase. The results do suggest however that airlines should use information they have already available in a customer search. Because the data is already available, this is a cost-effective way for airlines to significantly improve the prediction accuracy of their customer choice models with 99.9%

confidence. The proposed OMS is data-driven, can respond to searches in real-time, and is designed in modules for gradual embedding into existing processes, workflows, and system. Also, it is built in a way that it both works with existing revenue/offer management systems as well as innovations like continuous pricing and new distribution capabilities that are strategic priorities for airlines.

The findings recommend airlines to test the developed OMS and show customized bundle options next to the static branded fares in step 1 of the customer journey. These tests will answer whether the improved prediction accuracy leads to improved customer or business outcomes such as higher revenue or higher search-to-book conversion, and whether the solution is indeed understood and hence adopted. In addition to these online tests, future research can test the proposed OMS more holistically, test the prediction accuracy against complex machine learning models, and test generalization to other airlines or other sectors.

Final remarks

This dissertation adds to the academic literature on customized offer management. It is novel as it combines discrete choice and machine learning into a new offer management system that enhances transparency and improves prediction accuracy to help airlines show their customers a customized bundle in step 1 of the customer journey. As such, it would combine the convenience and recognizability of branded fares with the customizability and flexibility of *a la carte* ancillaries. The research highlights that using data that is available to airlines can achieve 99.9% of the benefits, whereas complex machine learning only improves the prediction accuracy of the last 0.1% of customer searches.

Samenvatting

In dit proefschrift wordt een nieuwe oplossing ontwikkeld voor het beheer van op maat gemaakte prijsoffertes van luchtvaartmaatschappijen, met als doel de haalbaarheid, bruikbaarheid en uitvoerbaarheid ervan te combineren. De oplossing wordt getest met *real data* van een grote netwerk luchtvaartmaatschappij. Het proefschrift breidt de bestaande academische literatuur en praktische toepassingen uit. Het stelt een kosteneffectieve en begrijpelijke methode voor luchtvaartmaatschappijen voor. Hiermee kunnen zij de voorspellingsnauwkeurigheid van klantkeuzemodellen aanzienlijk verbeteren. Dit gebeurt zonder de noodzaak van complexe modellen. De methode gebruikt bestaande gegevens en eenvoudige voorspellingen. Het onderzoek toont aan dat hoogdimensionaleen datagestuurde segmentatie, mogelijk geholpen door machine learning om dataschaarste op te lossen, gecombineerd kan worden met de traceerbaarheid van discrete keuzemodellen.

Inleiding

Luchtvaartmaatschappijen bedienen klanten met verschillende reisdoelen en voorkeuren, gaande van zakenreizigers die op dezelfde dag heen en terug willen vliegen tussen zakensteden tot gezinnen met drie kinderen op zoek naar hun zomervakantie of expats die tijdelijk voor een half jaar naar een ander deel van de wereld verhuizen. Als reactie hierop hebben luchtvaartmaatschappijen hun producten aangepast en bieden ze hun klanten een breed scala aan aanvullende diensten naast het boeken van de stoel in het vliegtuig. Deze ontwikkeling werd gestimuleerd door de toegenomen concurrentie als gevolg van de komst van low-cost carriers (LCC's) in het begin van de jaren 2000. De populairste aanvullende diensten zijn bagagebehandeling, stoelreservering (plus meer beenruimte), flexibiliteit, en loungetoegang. Het aanbieden van steeds meer extra opties (bagage, reisverzekering, stoelkeuze) heeft de relevantie vergroot van het ontwerpen van een klantreis die zowel handig is als klanten de mogelijkheid geeft om het product aan te passen aan hun specifieke behoeften. Tot nu toe hebben luchtvaartmaatschappijen dit probleem opgelost door merktarieven te creëren, d.w.z. een statische voorselectie van bundels die herkenbaar zijn voor klanten en die gericht zijn op het oplossen van de meest typische behoeften van klanten. Gewoonlijk tonen luchtvaartmaatschappijen drie tot vier van deze merktarieven in stap 1 van het klanttraject. Daarnaast bieden luchtvaartmaatschappijen hun klanten in stap 2 van het klanttraject de mogelijkheid om bovenop deze merktarieven *à la carte* extra accessoires aan te schaffen.

Door de statische voorselectie bevatten sommige merktarieven bijkomende kosten die waarschijnlijk niet relevant zijn voor een specifieke klant die zijn/haar zoekopdracht uitvoert. Bijvoorbeeld, een zakenreiziger die zoekt naar een retourvlucht op dezelfde dag is waarschijnlijk niet geïnteresseerd in ruimbagage. Ze kunnen echter wel geïnteresseerd zijn in een snelle doorgang bij de beveiliging of andere - minder gebruikelijke - bijkomstigheden die ze pas in stap 2 van het klanttraject en in een lange keuzelijst terugvinden . Dit is niet de meest handige klantervaring. Kunnen luchtvaartmaatschappijen niet slimmer zijn en een op maat gemaakte bundeloptie aanbieden die is afgestemd op de specifieke zoekopdracht in stap 1 van het klanttraject?

Dit is het onderzoeksdoel van dit proefschrift. Het is gebaseerd op de hypothese dat elke zoekopdracht naar een luchtvaartproduct impliciet en expliciet informatie verschaft aan de luchtvaartmaatschappijen zodat zij deze informatie dan kunnen gebruiken om te reageren met een klantgerichte, op maat gemaakte bundel in aanvulling op de statische merktarieven.

Probleemstelling

Vanwege de relevantie heeft deze probleemstelling veel aandacht gekregen van zowel academische als industriële wetenschappers. Het oplossen van het beheerprobleem van aangepaste aanbiedingen vereist het oplossen van de verschillende deelproblemen van

- segmentatie (hoe kunnen klanten worden gegroepeerd in verschillende segmenten die zich verschillend van elkaar gedragen?),
- bundeling (hoe creëer je een bundel in realtime wanneer een klant een zoekopdracht uitvoert?),
- assortiment (welke van alle mogelijke producten moet een klant zien te krijgen in stap 1 van het klanttraject?), en
- prijsstelling (hoe bepaal je de prijs van de vlucht, de nevenproducten of de hele bundel?).

Tot nu toe kunnen de bestaande literatuur en de praktische toepassingen in twee categorieën worden onderverdeeld. Soms gebruiken onderzoekers discrete keuzemodellen die gericht zijn op een ééncijferig aantal segmenten. Meestal maken ze een onderscheid tussen de twee klassieke segmenten van “Zakenreizen” of “Vrijetijdsreizen”, waarbij de verblijfsduur het enige criterium voor segmentering is. Ze slagen er echter niet in om het hele spectrum van klantenvoorkeuren vast te leggen en hun voorspelnauwkeurigheid lijdt onder het negeren van relevante informatie, zoals het tijdstip waarop de klant zoekt of via welk kanaal de zoekopdracht komt. Als alternatief gebruiken ze machine learning modellen die een veel hoger aantal kenmerken aankunnen en daardoor aanzienlijk meer segmenten kunnen opbouwen Dit kan gaan tot continue segmentatie waarbij een zoekopdracht van een klant helemaal niet in kaart wordt gebracht in discrete segmenten. Deze modellen kunnen op een veel bredere manier inspelen op de behoeften van de klant. Ze zijn echter meestal moeilijk te begrijpen omdat veel complexe machine learning modellen lijden onder het black box-karakter. Als werknemers van een luchtvaartmaatschappij niet kunnen begrijpen hoe een model tot segmentatie komt, heeft de luchtvaartmaatschappij het meestal moeilijk om het te gebruiken.

Dit is de onderzoeksleemte die dit proefschrift wil invullen. Het ontwikkelt en valideert een nieuwe oplossing die enerzijds een hogere voorspellingsnauwkeurigheid biedt dan bestaande discrete keuzemodellen en anderzijdsbruikbaarder is voor medewerkers van luchtvaartmaatschappijen en dit dankzij een eenvoudige en begrijpelijke logica om zoekopdrachten te koppelen aan segmenten. Dit gebeurt door zoekopdrachten van klanten op een datagestuurde manier te segmenteren in duizenden tot miljoenen segmenten Dit gebeurt op basis van gegevens die in de zoekopdrachten van klanten zijn opgenomen en dus zonder extra kosten beschikbaar zijn voor luchtvaartmaatschappijen.

Methodologie

Dit proefschrift begint met het schetsen van een conceptueel nieuw Offer Management System (OMS) dat de problemen van segmentatie, bundeling, assortiment en prijsstelling holistisch oplost. Het voorgestelde OMS is gebaseerd op twee fundamentele hypotheses, namelijk dat luchtvaartklanten kunnen worden gesegmenteerd in verschillende segmenten die verschillend aankoopgedrag vertonen en dat deze segmentatie de nauwkeurigheid van de voorspelling van de waarschijnlijkheid van klantkeuzes voor toekomstige zoekopdrachten van klanten kan verbeteren.

Zowel de haalbaarheid als de levensvatbaarheid van het voorgestelde OMS worden gevalideerd met bestaande luchtvaartmaatschappijgegevens aangezien de auteur toegang had tot honderden miljoenen geboekte luchtvaartcoupons tussen september 2018 en september 2023. Om de twee fundamentele hypothesen te testen, is de validatie opgesplitst in twee delen en wordt zowel inductief als deductief onderzoek gebruikt. Eerst doet het inductieve onderzoek empirische waarnemingen in de data, generaliseert deze vervolgens om patronen te vinden en ontwikkelt een theorie op basis van deze generalisatie Het deductieve onderzoek test deze nieuwe theorie op nieuwe data en bevestigt de geldigheid ervan met nieuwe waarnemingen.

Belangrijkste bevindingen en resultaten

Dit proefschrift bewijst dat het voorgestelde OMS aanbiedingen kan voorspellen die relevant zijn voor een specifieke zoekopdracht met een significant hogere nauwkeurigheid dan bestaande discrete keuzebenaderingen. De hoge statistische significantie wordt aangetoond met drie verschillende foutmaten en voor vijf verschillende tijdsperioden. Specifiek tonen de resultaten aan dat - in volgorde van belang – verklarend werken voor het product waarvoor de klant kiest: het verkoopkanaal (bijv. de eigen website van de luchtvaartmaatschappij of een reisbureau), de loyaliteitsstatus van de klant (verschillende niveaus van lidmaatschap of geen lidmaatschap), of een vlucht al dan niet 's nachts wordt uitgevoerd en de boekingsweekdag, Het proefschrift bevestigt dat het opnemen van deze kenmerken de nauwkeurigheid van de voorspelling van toekomstige klantkeuzes verhoogt met een statistische significantie van 99,9%. Het helpt luchtvaartmaatschappijen dus in hoge mate om te bepalen welke aangepaste bundel het meest waarschijnlijk overeenkomt met de voorkeuren van de specifieke klant.

Tegelijkertijd is de segmentatie, ondanks de hoge dimensionaliteit met duizenden segmenten, aantoonbaar veel begrijpelijker voor medewerkers van luchtvaartmaatschappijen dan continue segmentatie van complexe machine learning modellen. Dit komt omdat elke zoekopdracht wordt gekoppeld aan exact één klantsegment op basis van het verkoopkanaal, de loyaliteitsstatus van de klant, of de vlucht al dan niet 's nachts wordt uitgevoerd en de weekdag van de boeking. Bovendien bevestigen experimenten met verschillende tijdsperioden dat veranderingen in klantgedrag kunnen worden vastgelegd met een maandelijkse hertraining van het model. Tot slot vereist het voorgestelde OMS niet dat klanten hun persoonlijke identiteit bekendmaken waardoor het volledig in overeenstemming is met de GDPR regelgeving.

Vanuit methodologisch perspectief test het proefschrift vier voorspellingsmodellen met toenemende complexiteit. Dit leverde een onverwacht resultaat op. 99,9% van de verbeterde voorspellingsnauwkeurigheid kan worden bereikt met eenvoudige voorspellingen die ervan uitgaan dat het waargenomen gedrag in de trainingsperiode zich voortzet in de toekomst. Complexere voorspellingsmodellen, gebaseerd op een nieuwe toepassing van het bekende matrixfactorisatie-algoritme, verbeterden de nauwkeurigheid alleen bij de laatste 0,1% van klantzoekopdrachten. Dit algoritme is krachtig en wordt gebruikt in aanbevelingssystemen zoals die van Netflix en Amazon. Dit zijn de segmenten die in kaart worden gebracht in infrequente segmenten die een infrequente combinatie van kenmerken vertegenwoordigen, zoals een zeldzaam verkoopkanaal, een zeldzame loyaliteitsstatus en een overnachting.

Conclusies en implicaties

De resultaten laten zien dat luchtvaartmaatschappijen geen complexe machine learning-modellen nodig hebben om de nauwkeurigheid van de voorspelling te verbeteren van een specifiek product dat een specifieke klant waarschijnlijk zal kopen. De resultaten suggereren echter wel dat luchtvaartmaatschappijen gebruik moeten maken van informatie die ze al beschikbaar hebben bij het zoeken naar een klant. Omdat de gegevens al beschikbaar zijn, is dit een kosteneffectieve manier voor luchtvaartmaatschappijen om de voorspellingsnauwkeurigheid van hun klantkeuzemodellen aanzienlijk te verbeteren met een betrouwbaarheid van 99,9%. Het voorgestelde OMS is datagestuurd, kan in *realtime* reageren op zoekopdrachten en is ontworpen in modules voor geleidelijke toepassing in bestaande processen, workflows en systemen. Het is ook zo gebouwd dat het zowel werkt met bestaande systemen voor inkomsten- en aanbodbeheer als met innovaties zoals continue prijsstelling en nieuwe distributiemogelijkheden die strategische prioriteiten zijn voor luchtvaartmaatschappijen.

De bevindingen nodigen de luchtvaartmaatschappijen aan om het ontwikkelde OMS te testen en aangepaste bundelopties te tonen naast de statische merkprijzen in stap 1 van het klanttraject. Deze tests geven antwoord op de vraag of de verbeterde nauwkeurigheid van de voorspellingen leidt tot betere klant- of bedrijfsresultaten, zoals een hogere omzet of een hogere conversie van zoekopdrachten naar boekingen en of de oplossing inderdaad wordt begrepen en dus geaccepteerd door medewerkers van luchtvaartmaatschappijen. Naast deze online tests kan toekomstig onderzoek het voorgestelde OMS meer holistisch testen, de voorspellingsnauwkeurigheid toetsen aan complexe machine learning modellen met continue segmentatie en generalisatie testen naar andere luchtvaartmaatschappijen of andere sectoren.

Slotopmerkingen

Dit proefschrift draagt bij aan de academische literatuur over het beheer van gepersonaliseerde aanbiedingen. Het is vernieuwend omdat het discrete keuzemodellen en machine learning combineert in een nieuw aanbodbeheersysteem dat zowel transparant is als de nauwkeurigheid van voorspellingen verbetert. Dit helpt luchtvaartmaatschappijen om hun klanten een op maat gemaakte bundel aan te bieden in stap 1 van het klanttraject. Zo combineert het de gebruiksvriendelijkheid en herkenbaarheid van merktarieven met de aanpasbaarheid en flexibiliteit van à la carte accessoires. Het onderzoek benadrukt dat het gebruik van beschikbare data bij luchtvaartmaatschappijen 99,9% van de voordelen kan opleveren terwijl complexe machine learning alleen de nauwkeurigheid van de voorspelling van de laatste 0,1% van de zoekopdrachten van klanten verbetert.

Table of contents

List of tables

List of figures

List of abbreviations

ACDP	Ancillary choice dynamic program
AI	Artificial intelligence
ANN	Artificial neural networks
API	Application programming interface
ATPCO	Airline Tariff Publishing Company
B2B	Business-to-business
B2C	Business-to-consumer
BF	Branded fare
CB	Content-based (filtering)
CF	Collaborative filtering
CNN	Convolutional neural networks
CPU	Computer processing units
DC	Discrete choice (analysis)
EMSR	Expected marginal seat revenue
GAN	Generative adversarial networks
GDPR	General Data Protection Regulation
GDS	Global distribution system
GenAI	Generative artificial intelligence
GPU	Graphics processing unit
JS(D)	Jensen-Shannon (divergence)
IATA	International Air Transport Association
IIA	Independence of irrelevant alternatives

KL	Kullback-Leibler (divergence)
LCC	Low-cost carrier
LSTM	Long-short-term-memory
MECE	Mutually exclusive, collectively exhaustive
MF	Matrix factorization
ML	Machine learning
MNL	Multinomial logit (model)
NDC	New distribution capability
OBL	Optimal booking limit
OI	Optimizer increment
OM	Offer management
OMS	Offer management system
OR	Operations research
O-D	Origin-destination
PCA	Principal component analysis
PE	Prediction error
PODS	Passenger origin destination simulator
RL	Reinforcement learning
RM	Revenue management
RMS	Revenue management system
RNN	Recurrent neural networks
RP	Revealed preferences
RPK	Revenue passenger kilometer
RQ	Research question

SL	Supervised learning
SP	Stated preferences
SVM	Support vector machine
UL	Unsupervised learning
WAPPE	Weighted average percentage point error
WTP	Willingness to pay
XGB	Extreme gradient boosting (algorithm)

1 Introduction

Airlines serve many different customers and face the challenge of designing product and pricing strategies matching these diverse customer preferences. The business challenge of managing and optimizing offers tailored to the specific customer making a search – referred to in this text as the **customized offer management** problem – requires airlines to balance customer preferences and profitability. As such, customized offer management has become increasingly relevant to airlines as an instrument to maximize outcomes for their customers and their business.

This dissertation starts with an overview of airline revenue management, pricing, distribution, ancillary services, bundling, assortment, and segmentation practices. It puts recent developments towards customized offer management in historical context. Based on that, the dissertation shall contribute towards the academic literature by conceptualizing and validating a solution to predict customer choice for airline products based on parameters of the particular customer search. It uses high-dimensional segmentation and machine learning to improve the prediction accuracy of discrete choice models. With that, it presents a step forward towards the airline industry's long-term goal of customized offer management to combine both the simplicity of pre-select bundles, commonly termed branded fares, as well as the flexibility and customizability of *a la carte* ancillaries.

Chapter 1 describes the research problem, provides the context of airline offer management, and derives the dissertation's research objective. The chapter is structured as follows. Section 1.1 outlines the research problem and sets it into context of the recent Covid disruption and the rise of (Gen)AI facilitating usage of large amounts of data more cheaply and accessibly. Section 1.2 provides context on airline offer management with its various subproblems as well as definitions of the terminology used. Section 1.3 synthesizes the context into the research objective and questions. Section 1.4 outlines the structure of the remaining dissertation.

1.1 Research problem

Airlines are confronted with the challenge to design product and pricing strategies in an environment characterized by three characteristics. First, customers differ with respect to both the services they demand and their willingness to pay (WTP) for these services. Second, the airline product is customizable with the base product – the mere aircraft seat – to be complemented with a multitude of ancillaries. These include seat reservation, baggage, flexibility, refundability, fast track through airport security, lounge access, inflight entertainment (wifi, etc.), ground transportation to/from the airport, pollution compensation, and potentially many more. Third, airlines' retail-oriented business model results in many individual customer searches requiring real-time responses with a high level of automation.

Naturally, airlines aim to maximize customer and business outcomes of every individual search. These can be customer satisfaction, search-to-book conversion, seat load factor, revenue, or profitability. To solve the offer management problem, it can be simplified by breaking it down into the subproblems of segmentation, bundling, pricing, and assortment. **Segmentation** aims to identify distinct groups of customers with similar behavior. The assumption is that customers in a segment have similar preferences, whereas the distinct segments represent distinct and different preferences (Teichert et al., 2008). **Bundling** refers to a product bundle, i.e. an unbreakable entity of base product plus ancillaries with one single price tag. Goal is to identify and create the bundle that most closely resembles individual customer preferences. Various studies demonstrate that bundling, under certain conditions, can increase sellers' profits (e.g., Stigler, 1963; Schmalensee, 1984). In addition, offering unique bundles to every customer avoids potential conflicts with price discrimination laws (Adams & Yellen, 1976). Further, bundling enables more robust prediction of customers' WTP since WTP estimation for bundles exhibits smaller variance than for single ancillaries (Bakos & Brynjolfsson, 1999). This makes price discrimination strategies more powerful and creates a link between bundling and **pricing**. Optimal pricing necessitates an understanding of the customer's individual WTP. Many studies confirm the profitability of successful airline revenue management strategies (see Belobaba et al., 2016 for an overview). Finally, **assortment** optimizes the selection of all possible bundles to be displayed to the customer.

This dissertation introduces the term **"customized offer management"** to describe the challenges of segmentation, bundling, pricing, and assortment, specifically applied to the airline context. The approach focuses on leveraging both explicit and implicit customer information from search behavior to develop bespoke product and pricing strategies for each individual search. This compares to airlines primarily relying on rule-based customer segmentation and pricing algorithms that categorize customers into a limited number of segments, often just two, with the most common criterion being travel purpose Business vs. Leisure. Most academic literature on offer management follows this practice and has developed models with single digits of customer segments. These rule-based systems typically decompose the customer buying journey into two steps (Figure 1).

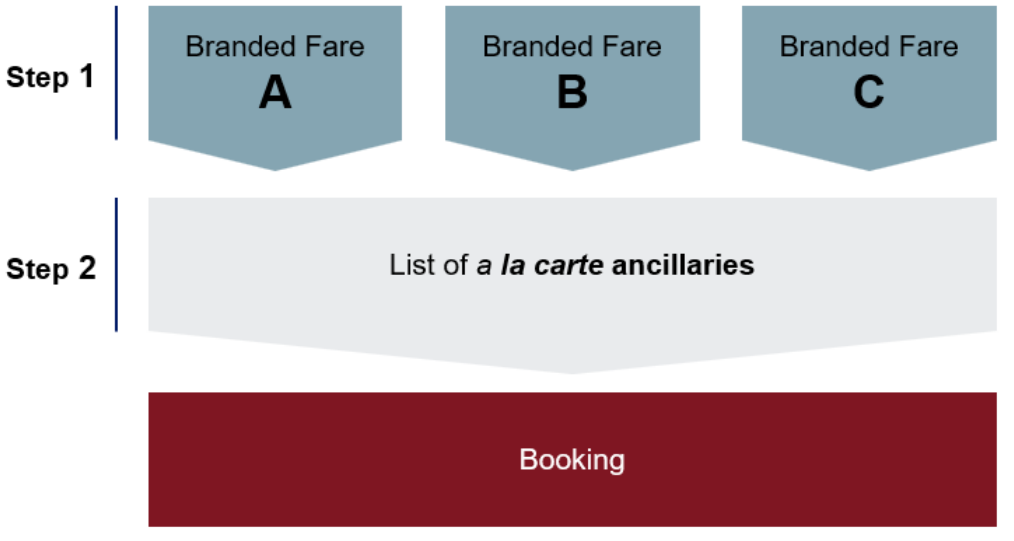

Figure 1: *Typical airline customer journey. Customers select amongst 3-4 branded fares in step 1, and then select from a list of a la carte ancillaries in step 2 before proceeding to a booking.*

In the first step, customers are presented three to four pre-set bundle options, so-called *branded fares*. These branded fares are products designed to offer customers a convenient and recognizable purchase experience and to meet the assumed needs of the biggest customer segments. Within an airline, branded fares are typically the same, independent of what customers search for, why they search, and other characteristics of the search. Whilst recognizability is the advantage, this standardization leads to offering some customers branded fares, which include ancillaries that are irrelevant for their search. An intuitive example is offering checked-in baggage to a customer searching for a same-day return flight between cities that are major business centers. For this search, fast-track through security or a taxi to/from the airport seem the more adequate offers, but are typically not part of airline branded fares. In the second step, customers are presented with a list of *a la carte* ancillaries to select from. With these, customers can truly customize their product. However, going through all these options is both time-consuming and might overwhelm customers.

Furthermore, given the sheer size of combinations of features that come with each customer search, hard-programmed rule-based approaches seem unlikely to produce optimal results. Optimality is not only difficult to achieve due to inherent analytical complexity, but also due to organizational, institutional and industry structures grown over decades. In particular, the multi-channel offering and, hence, interfaces with multiple distribution systems, has caused most legacy airlines to develop their offer management in gradual steps. However, new industry trends like new distribution capability (NDC) and One Order as well as direct distribution via airlines' own websites have begun to open the door to more revolutionary new frameworks of dynamic offer generation and customized real-time offer management strategies. Even more, these trends necessitate airlines to respond to each customer search with instant and customized offers and prices (Mahendru et al., 2024). Moreover, recent advancements in data processing and (Generative) Artificial Intelligence make data availability, processing, storage, and manipulation cheaper and more easily accessible. At the same time, airlines need to fit these models into existing business processes and systems. Most importantly, airline users like Revenue Management Analysts need to understand and trust the models as well as learn how to use them. This challenges many AI applications, whose black box character is at odds with usability (Vinod, 2020).

Solving the offer management problem is of paramount importance to the economic viability of airlines. An estimated benefit of 7 USD revenue increase per passenger (IATA, 2023; McKinsey, 2019) has inspired the International Air Transport Association (IATA) to structure Modern Airline Retailing as a program and drive it across the industry (IATA, 2025). Many leading airlines started implementation of airline retailing, including British Airways, Lufthansa Group, Turkish Airlines, Air India, Emirates, Air France – KLM, American Airlines, Singapore Airlines, United, and many more (IATA, 2025). An additional revenue of 7 USD per passenger sums up to 40 billion USD across the airline industry by 2030. This makes a significant difference for airlines, who, even before the Covid pandemic, struggled with profitability. IATA & McKinsey (2022) estimated the pre-pandemic yearly total economic loss of airlines at 18 billion USD. The Covid-19 pandemic disrupted the airline industry with unprecedented declines in air travel demand and flight operations as well as hundreds of billions of economic impact. In 2020 and 2021, global airline traffic measured by revenue passenger kilometers (RPK) fell by 66% and 58%, respectively, compared to 2019 levels (IATA & McKinsey, 2022). In revenue terms, 2020 set airlines back to 2004 levels with a 55% decline from pre-Covid 2019 (McKinsey, 2022). At the peak of the crisis, airlines reported reductions in capacity ranging from 80% to 95% during March and April 2020 (Avionics International, 2020). The economic losses for airlines worldwide were estimated at $168-175 billion in 2020 and $104 billion in 2021 (IATA & McKinsey, 2022; McKinsey, 2022). Next to emphasizing the paramount importance of airlines to optimize expected profitability out of every search, the pandemic also highlighted that customer preferences can rapidly change. This poses an additional challenge to airlines when designing their customized offer management strategies of the future.

In general, innovation projects to improve airline performance are spurred by either new value generation or cost reduction purposes, or both. Customized offer management should deliver on both these objectives. **New value generation** focuses on three strategic drivers. First, more relevant offers lead to both increased search-to-book conversion and higher customer WTP. Customized offers help airlines to differentiate against their competition and avoid the commodity trap. Second, higher customer satisfaction results in higher customer loyalty. Third, more robust WTP estimation helps airlines understand and capitalize on every customer search. **Cost reduction** follows from automation if any advancement becomes implemented at scale. This seems particularly conceivable if an advancement addresses the offer management problem as holistically as possible, i.e. comprises as many of the offer management subproblems as possible.

Abstracting from the airline world and assuming a more general view on the transportation market, similar challenges can also be found in other transportation companies, for which the three characteristics of differentiated customer needs, customizable product, and a high number of searches are fulfilled. These could be providers of high-speed or overnight trains, ferries or auto trains, and potentially even freight carriers like shipping lines or cargo train operators. Consequently, the audience of offer management research expands from airlines as primary research addressees to transportation providers in general.

1.2 Context: airline offer management

This section introduces the context of offer management with its various subproblems and defines relevant terminology. It starts with revenue management, pricing, and distribution in Section 1.2.1. Next, ancillary services are discussed in Section 1.2.2 before covering customer segmentation (Section 1.2.3) as well as bundling and assortment (Section 1.2.4). Finally, Section 1.2.5 puts these subproblems back together to define the customized airline offer management problem. Throughout the section, both groundbreaking academic research as well as existing applications to solve the airline offer management problem are covered.

1.2.1 Revenue management, pricing, and distribution

Revenue management (RM) is a strategic discipline to optimize revenue through the effective management of pricing and inventory. There are several prerequisites to deploy RM profitably, all of which are fulfilled in the airline industry. Inventory is perishable as unsold seats at departure present lost revenue opportunities. There is temporal variability in demand (seasonality). Customers differ in their trip purpose, required needs for services, and willingness to pay (WTP). Lead times vary as well. Some customers book long-time ahead, others shortly prior to departure. Further, the large amount of customer searches requires airlines to respond in real-time and with high level of automation. For these reasons, airlines were among the first to develop modern RM systems, which then also spread to other industries like hotels, rental cars and retail. Many studies confirm the profitability of successful airline RM strategies (see Belobaba et al., 2016, for an overview).

Historically, RM goes back to Littlewood (1972) from the British Overseas Airways Corporation who proposed that airlines maximize revenues instead of passenger occupancy on a flight for the perishable seat inventory. A common definition of the primary goal of RM is to maximize overall revenue by selling the right seats to the right customers at the right time (Freiberger, 2024). RM involves analyzing historical data and forecasting future demand to dynamically adjust prices based on customer WTP, market conditions, available capacity, and competitive actions.

Following this definition, **pricing** is a subset of RM problems along with others like capacity steering and demand forecasting. Optimal pricing matches individual customer WTP. To exploit individual WTP, airlines segment their customers and estimate price-response curves and elasticities for each segment. Restrictive fare rules play a crucial role in this segmentation. A common example is that lower-priced fare classes are reserved for customers willing to meet specific conditions, such as weekend stays, to deter business travelers from booking lower fares (Belobaba, 2011). The assumption of “perfect” fare restrictions led to the assumption of independent demand in each fare class in many early RM models. These include the expected marginal seat revenue (EMSR) models (Belobaba, 1987) that have been widely adopted in the airline industry (Van Ryzin & McGill, 2000) as well as the optimal booking limits (OBL) model (Curry, 1990; Brumelle & McGill, 1993). Also, these models assumed demand to arrive strictly in increasing order of WTP although this was relaxed with demand arrivals modeled as Markov decision process solved by dynamic programming (Lautenbacher & Stidham, 1999).

The effectiveness of differential pricing based on fare restrictions was significantly reduced when low-cost carriers (LCCs) entered the market with simplified and almost unrestricted fare structures (Belobaba, 2011). Competition with LCCs forced legacy airlines to adapt, weakening their ability to segment demand with fare restrictions and invalidating the assumption of independent demand between fare classes. Continued reliance on traditional RM systems with unrestricted fares led to a downward spiral in airfares (Cooper et al., 2006), necessitating adaptations in RM systems. An overview of adaptions is provided by Strauss et al. (2018).

Airline **distribution** relies on both direct and indirect channels. Direct channels, mostly airline websites, are under direct airline control. This is not the case for indirect channels like Global Distribution Systems (GDSs), which processed almost 50% of global flight bookings in 2015 (Taubmann, 2016). These indirect channels limit airlines' influence how their services are priced and offered. The power dynamics in airline distribution and their effects on industry evolution were studied by Albers et al. (2024). Further, technological limitations restrict the number of price points, and corresponding airline booking classes, to twenty-six, and airfares can only be updated in fixed intervals through organizations like the Airline Tariff Publishing Company (ATPCO). Most importantly, GDSs and flight comparison websites rank airline itineraries almost exclusively based on price and schedule. This severely limits airlines' flexibility to differentiate themselves through different service offerings.

Motivated by advancements in other industries, airlines began to explore opportunities to overcome the rigidity of twenty-six pre-defined price points (Golrezaei et al., 2014; Gallego et al., 2016). Wittman & Belobaba (2019) developed a definitional framework for two distinct approaches to dynamic pricing. First, dynamic price adjustments amend pre-defined published fares based on estimated WTP, trip purpose and competitor fares (Fiig et al., 2016). Second, continuous pricing leaves price points entirely behind, instead choosing prices from a continuous and unlimited set. However, such differential pricing practices could raise regulatory implications (Wittman, 2018).

The urgency to improve pricing capabilities intensified during the Covid-19 pandemic and its drastic impact on air travel demand. When demand changes rapidly, airlines can no longer afford slow response times and long times to market resulting from the traditional distribution landscape. Compared to other transport domains like air cargo, ocean carriers, freight forwarding, and logistics, as well as most other sectors, passenger airline pricing practices can be viewed as advanced, though not at the forefront anymore (McKinsey, 2020a).

In summary, RM plays a critical role in maximizing airline revenue through sophisticated pricing strategies that respond dynamically to market conditions and available capacity. As airlines navigate a rapidly changing landscape post-pandemic, they must continue to innovate their RM practices by leveraging technology and data analytics to enhance their pricing, offering, and distribution strategies. This will be pivotal for airlines aiming to remain competitive in an increasingly complex market environment.

1.2.2 Ancillary services

Deregulation of air travel, initially started in the United States in 1978 before reaching nearly all parts of the world by the early 2000s (Belobaba et al., 2016), not only enabled competition and increased price pressure on airlines, but also spurred the development of the “no-frills” business model of LCCs. This caused the shift away from earlier all-inclusive airfares to today’s disentanglement of all kinds of ancillaries from the base fare for the mere aircraft seat. While many legacy airlines still position themselves as “premium” by including some services in their base fares, they too have had to follow the trend and reduced the services included in fares, offering them as add-on to customers instead.

Ancillary revenues can be defined as any airline revenue generated by activities and services beyond the simple transportation of customers from A to B (IdeaWorks, 2019). Ancillary revenues have grown over the last decades. They are estimated at over 100 billion USD worldwide for 2023, accounting for 15% of total airline revenue and reaching the previous pre-Covid all-time high from 2019 (IdeaWorks, 2019; CarTrawler, 2023). This growth has helped airlines counter the long-term trend of decreasing airfares and maintain profitability. LCCs have continued to drive innovation in ancillary revenue, but legacy carriers have also increased their focus on it. Spirit Airlines made over 56% of their revenue from ancillaries in 2023 and Jet2.com achieved a new global record of 95.83 USD ancillary revenue per passenger in 2023. (IdeaWorks, 2024). Delta Air Lines earned over 8 billion USD ancillary revenue in 2022, a 37% increase from the previous year (Relay 42, 2024). Whilst initially most relevant on short-haul routes – in line with LCCs primarily operating short-haul – ancillaries are of increasing importance on long-haul as well. This has inspired research on customer WTP for ancillaries on long-haul flights (Chiambaretto, 2021).

Ancillary services can be grouped into various categories. There are airline-own services, such as check-in baggage, advanced seat reservation, increased legroom, priority boarding. etc. Typically, baggage fees are the most relevant, composing 60% of total LCC ancillary revenues (IdeaWorks & CarTrawler, 2018). Other ancillaries concern ground facilities like lounge access or fast track through security. A third group includes third party services, for which airlines receive a commission. Examples are ground transportation, rental cars or hotel sales.

While practically all airlines rely on quantitative models to price their seats, most employ static pricing for ancillary services. However, since ancillary sales mostly happen via direct distribution channels, airlines enjoy more pricing flexibility for them than for the seat itself. Therefore, airlines tend to use them as test cases for dynamic and continuous pricing (IdeaWorks & CarTrawler, 2018).

An example is US LCC Spirit Airlines, which analyzed the impact of dynamic baggage fees, varied based on search request, travel date, route and time of purchase (CAPA Centre for Aviation, 2019). In addition, Shukla et al. (2019) described the dynamic pricing model developed by Deepair solutions, providing machine learning (ML) based pricing recommendations specific to each customer interaction aimed to optimize expected revenue per customer. They conclude ML can outperform human rule-based approaches and achieve 36% higher conversion from offers to ancillary bookings and 10% higher revenue per offer.

Bockelie & Belobaba (2017) distinguish two customer types, sequential and simultaneous consumers. *Sequential consumers* choose their flight first and evaluate ancillaries afterwards. This resembles the approach of most flight search engines and GDSs. *Simultaneous consumers*, on the contrary, select flight and ancillaries at the same time based on a combined overall WTP. For simultaneous consumers, ancillary pricing directly affects passengers' choice for one airline over another. For that reason, seat and ancillary pricing need to be performed together to maximize overall revenue (Bockelie, 2019; Lu, 2019). Particular care needs to be taken with those ancillaries coming at variable and/or opportunity costs for airlines. Several approaches were proposed for joint seat and ancillary pricing. The optimizer increment (OI) model increases fare values in the RM optimizer by the expected ancillary revenue contribution of an incremental passenger (Hao, 2014). However, Bockelie (2019) showed OI only yields optimal results under limited conditions, and instead proposed the ancillary choice dynamic program (ACDP), which is an extension to the dynamic program in Talluri & van Ryzin (2004). Both OI and ACDP represent assortment optimization methods, selecting which combinations of pre-defined sets of fares and ancillary prices to display, rather than optimization methods for flights and ancillary prices themselves. On the other hand, Ødegaard & Wilson (2016) optimize flight and ancillary pricing, but ignore current reliance of airline distribution systems on pre-set flight and ancillary price points.

In summary, the literature highlights the growing importance of ancillary revenue for airlines, the various categories of ancillary services, and the need for sophisticated pricing strategies that optimize seat and ancillary revenue together. As the industry continues to evolve, airlines will need to further innovate and customize their ancillary offerings to meet the changing needs and preferences of their customers.

1.2.3 Customer segmentation

Airlines segment their customers to tailor their services, marketing efforts, and pricing models to meet the diverse needs of their passengers. Historically, customer segmentation in the airline industry relied on demographic factors such as age, gender, income, and business vs. leisure travel purpose (Belobaba et al., 2016; Camilleri, 2018; Avram, 2019). Goal was to understand how these variables help airlines recognize customer needs and preferences. An example is Skift (2017) comparing travel behavior of different generations. They report Millenials travel 35 days per year, compared to 29 days for Gen Z, 26 days for Gen X and 27 days for Baby boomers.

Over the last two decades, and especially with the rise of LCCs, airlines shifted to more nuanced segmentation techniques. Behavioral segmentation considering factors like booking patterns, travel frequency, and price sensitivity has gained traction. It aims to help airlines understand which passengers are willing to pay more for flexibility or additional services. Teichert et al. (2008) identified that frequent flyers often exhibit different purchasing behaviors compared to leisure travelers. Their research is based on surveying 5800 frequent flyers and concluded five market segments: efficiency/punctuality; comfort; price; price/performance; and catch all/flexibility. In their study, each segment chose their product based on different evaluations of variables like punctuality, flexibility, schedule, catering, price, and product. Wittmer & Hinnen (2016) analyzed psychographic segmentation focusing on criteria like trip motivation, destination, length of flight, length of stay, travel class, frequency of flying, cultural passenger background, airline preference, loyalty program membership, seat preference within a compartment, and environmental considerations. Main challenge with many of these criteria is they are more difficult to measure than demographic segmentation variables.

Over the last years, further trends have encouraged airlines to adopt a more dynamic approach to customer segmentation. Digitalization, environmental awareness, and lifestyle changes have increased complexity of consumer behavior. An example is the trend of “bleisure” travel as a mix of business and leisure travel (Wittmer & Hinnen, 2016). Moreover, leaps in data analytics have enabled airlines to collect and process vast amounts of customer data more efficiently, allowing for more effective personalized marketing campaigns and offerings tailored to specific customer segments. Enhanced segmentation also helps airlines capitalize on the increasing importance of ancillary revenues. If airlines understand which segments are likely to purchase which ancillary, then they can significantly improve their profitability. In particular, airlines can optimize their offering and enhance customer loyalty by identifying underserved segments (Avram, 2019).

In summary, customer segmentation in the airline industry has evolved from simple demographic to more advanced behavioral analyses that consider a range of factors influencing passenger choices. This shift towards a more nuanced modeling of consumer preferences reflects broader trends and emphasizes the relevance of customization to improve customer satisfaction and increase revenue.

1.2.4 Bundling and assortment

Bundling refers to creating a product bundle, i.e. an unbreakable entity of base product plus ancillaries with one single price tag. Early studies into the economics of bundling were conducted by Stigler (1963), demonstrating that bundling can increase profits under certain conditions, and Adams & Yellen (1976). Several types of bundling can be distinguished (see Kobayashi, 2005, for an overview). *Mixed bundling* describes bundles sold alongside the individual products, whereas in *pure bundling*, some products are only available as bundle component, but not for individual purchase. Schmalensee (1984) showed both mixed and pure bundling could be more profitable than *unbundled (“a la carte”)* sales. Bundling was found to be particularly effective when items with negatively correlated WTP are bundled together. Bundling is especially popular when marginal costs are low and inventories unlimited (Bakos & Brynjolfsson, 1999), with information goods and software products as examples.

Later studies explored the psychological aspects of bundling, relaxing earlier assumptions of rational consumers and additive WTP. These psychological concepts influential in consumer decision-making include anchoring (Yadav, 1994) and the center-stage effect, i.e. consumers’ tendency towards choosing the option presented in the middle (Raghubir & Valenzuela, 2006; Rodway et al., 2012). Methodologically, bundling enables more robust prediction of customers’ WTP with smaller variance than estimating WTP for single ancillaries (Bakos & Brynjolfsson, 1999). This can enhance the effectiveness of price discrimination strategies. Offering unique bundles to every customer can further help avoid potential conflicts with price discrimination laws (Adams & Yellen, 1976).

Optimal product pricing also depends on alternative products available for purchase. For that reason, **assortment** optimization is closely related to bundling. Assortment optimization aims to optimize the selection which bundles to display to customers. It originated in the retail industry with an optimization of products on store shelves (Kök et al., 2008). In the context of bundling, joint optimization of pricing and assortment becomes crucial. This led to the development of integrated models, such as the product planning model by Ferreira & Wu (2011). Optimal bundling, pricing and assortment were further researched by Bulut et al. (2009) and Gürler et al. (2009).

To distinguish themselves from their competitors and to manage their ancillary service portfolio, airlines utilize both *branded fares (bundling)* as well as *a la carte pricing (unbundling).* Branded fares, i.e. offering customers pre-defined, static bundles, are easier for customers to comprehend than an overwhelming choice of *a la carte* ancillaries. On the contrary, *a la carte* sales provide more flexibility and enable customers to truly customize their product. **Customized bundling** can be viewed as a third strategy combining the best out of both branded fares and *a la carte* sales. The primary challenge for airlines is to identify which customized bundle most closely meets customer preferences for every individual search. If airlines succeed in this task, it might also help them move towards customized pricing at the same time, which arguably has become increasingly important with disruptions like Covid-19.

In conclusion, customized and integrated pricing, segmentation, bundling, and assortment present the long-term vision for leading airline groups as they adapt to evolving consumer preferences and market conditions.

1.2.5 Putting the pieces together: customized offer management

The challenge to maximize expected customer or business outcomes of every individual customer search through pricing, segmentation, bundling, and assortment is aggregated as **"customized offer management"** in this dissertation. Further to the definitions in the previous sections, **"Flight pricing"** refers to pricing of the aircraft seat, whereas **"ancillary pricing"** refers to pricing of ancillaries and bundles.

Customized offer management aims to use the information customers explicitly and implicitly provide in their request to develop product and pricing strategies bespoke to each individual search. More precisely, it optimizes the assortment which of all possible bundle combinations to be displayed to a particular customer search at which price. NDC, One Order, increasing direct distribution, and advancements in data processing enable airlines to receive, extract and process more information about every single search request, and finally to respond to each search with a set of customized offers. More information about a customer search also facilitates better estimation of search-specific WTP. Moreover, NDC paves the way for effective deployment of continuous pricing, eliminating the concept of fare classes in the long run. Putting it all together, NDC in the limit enables customized offer management offering airlines entirely new possibilities to capitalize on. It introduces key features known from modern e-commerce with customer benefits of more transparency, flexible prices, and attractive offers in a digital environment (Lufthansa, 2018).

Customized offers – consisting of the flight plus possibly one or more ancillary services (together defined as "product"), either as bundle or *a la carte* items – at customized pricing. This would enhance market relevance and customer satisfaction as well as airline profits through two levers. First, better offers increase customer WTP and help airlines escape the commodity trap. Second, WTP estimation would happen with greater precision due to an improved understanding of the particular customer making a search. Airlines have increasingly adopted dynamic pricing strategies that integrate bundled offerings tailored to specific customer segments. Also, they have progressively begun to use personalized bundling tactics, product suggestions, and dynamic pricing to offer ancillaries tailored to individual customer preferences and behaviors (Relay 42, 2024).

Due to its practical relevance, customized offer management has recently attracted scientific attention. Madireddy et al. (2017) and Vinod et al. (2018) proposed solutions how airlines could transition from branded fares to customized dynamic offer generation by segmenting customers using trip characteristics. Shukla et al. (2019) extend this with the inclusion of temporal, market-specific, journey-specific and price-related features. Another approach is the dynamic offer generation project as part of the Passenger Origin-Destination Simulator (PODS) at the Massachusetts Institute of Technology

(PODS, 2021). As described in Wang (2020), they developed new discrete choice optimization models and heuristics to price flight and ancillary offers. Moreover, they derived a framework to help airlines decide dynamically between pure bundling and unbundled ancillary sales. In addition, they showed how such models can be integrated with traditional airline RM systems. Building on the PODS project, Wang et al. (2023) suggested a Markov chain choice model for bundling and ancillary pricing. Fiig et al. (2018) outlined the unresolved scientific challenge of dynamic airline offers, but points towards potential enablers, such as NDC, availability of more shopping data and advancements in statistical data analysis and ML.

Following the research problem and the context of airline offer management, the research objective and questions are discussed next.

1.3 Research objective and questions

The goal of the dissertation is to help advance both academia and airlines in solving the customized offer management problem. Working backwards from this goal, this research aims to deliver on two main criteria.

First, customers should be presented **offers relevant to their specific search in a convenient way,** hence combining the simplicity of pre-select branded fares with the customization of *a la carte* ancillaries. Customers should be presented a customized bundle already in step 1 of the customer journey (Figure 2). This way, the customer will find what they are looking for already in step 1 and they will not have to browse through the *a la carte* ancillaries in step 2 to proceed to a booking.

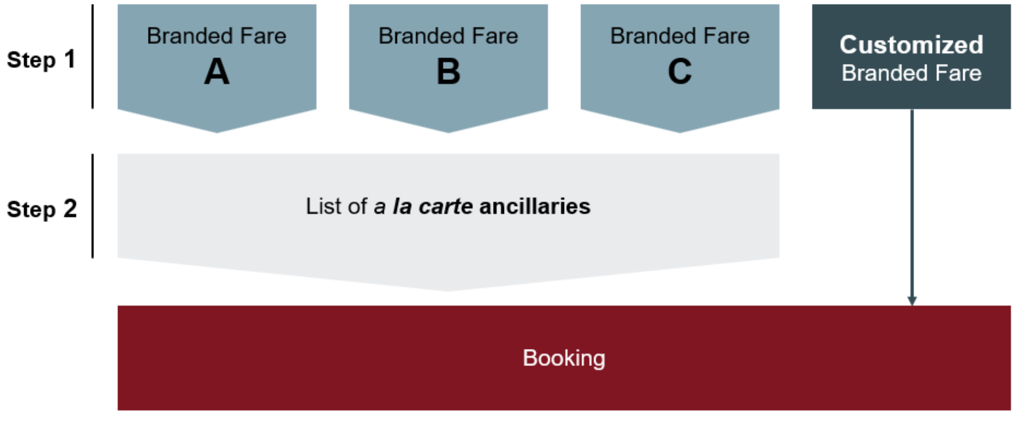

Figure 2: *Customer journey with an additional branded fare option displayed in step 1, which is a customized option relevant to the specific customer making a search.*

This requires solving the assortment problem; and more specifically, solving it in a way customized to the respective customer search. Given the heterogeneity of customers, this requires high-dimensional segmentation that goes beyond the traditional Business/Leisure or single-digit number of segments in existing discrete choice models. The segmentation is deemed effective if it helps airlines predict customer preferences with higher accuracy and if the model continuously learns and captures potential changes in customer behavior. Also, the proposed solution must be able to respond to a customer search in real-time. Finally, it should be applicable to all customer searches, independent whether the personal customer identity is declared or unknown to the airline.

Second, the proposed solution should be **usable to airlines.** To help change management, it needs to fit into existing airline processes and systems. In particular, the solution needs to be transparent, trusted, and understood by airline users. Hence, its logic for customized offer management must not be perceived as black box like some machine learning approaches, but need to be kept as simple, tractable, and robust as possible. For that reason, segments should be clearly identifiable and MECE (mutually exclusive, collectively exhaustive). Further, the proposed solution should present a cost-effective option for airlines to enhance their offer management. Next to a high level of automation, it should work with data readily available to airlines instead of requiring them to acquire external data. Also, the solution should be flexible and easily adapt to new settings such as new routes or new ancillary services. Lastly, the proposed solution needs to comply with data privacy regulation, i.e. work with anonymized data.

Putting both criteria together, this research aims to suggest and validate a solution for high-dimensional customer segmentation that improves the prediction accuracy of airline customer choice in a way understandable to airline users.

Concretely, the following research questions shall be answered:

1. Can airlines segment their customers into a large (thousands-millions) but finite number of MECE customer segments that exhibit significantly different choice behavior?
2. Can airlines use this segmentation to improve the prediction accuracy of future customer choices?
3. Can potential data sparsity problems be solved?
4. How much of the prediction accuracy improvements can be achieved in a disruptive event, such as the Covid-19 pandemic?
5. How can airlines practically implement the solution, cognizant of the trade-off between effectiveness and cost/complexity?

In Chapter 2, methods and existing solutions for the research problem will be presented. This includes an in-depth review of AI, ML and discrete choice models: their history and applicability to the customized offer management problem. With this additional context, the research questions will be refined, and the research contribution and novelty will be presented at the end of Chapter 2. Figure 3 visualizes this process of refining the research questions.

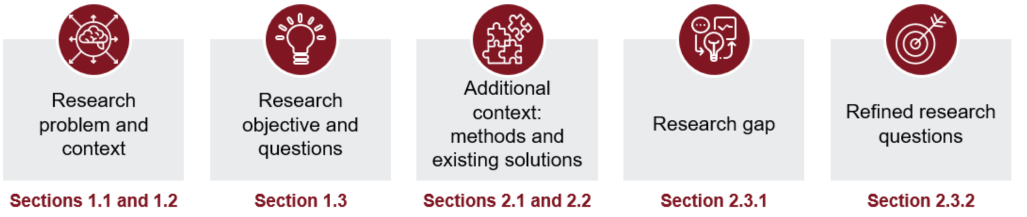

Figure 3: *Logical flow to refine the research questions.*[1]

1.4 Structure of the dissertation

The remainder of the dissertation is structured as follows and as shown in Figure 4.

Chapter 2 reviews existing methodological solutions to solve the research problem in the fields of Artificial intelligence (AI), Machine learning (ML), and Discrete choice analysis (DC). To understand advantages and disadvantages of these approaches, historical development, theoretical foundation as well as practical applications are discussed. Based on this additional context, the research gap is identified, the research questions are refined, and research contribution and novelty are presented.

Chapter 3 proposes a novel conceptual methodology to solve the research problem with a combination of machine learning and discrete choice analysis. It suggests a holistic offer management architecture and conceptualizes both high-dimensional segmentation and how to resolve data sparsity with a novel application of the established matrix factorization algorithm. Chapter 3 highlights the innovations of the proposed methodology and suggests how the conceptual architecture can be validated. It closes with a discussion of various practical

[1] In Section 2.3.2, the research questions from Section 1.3 are refined with additional context and gap from Sections 2.1, 2.2, and 2.3.1.

implementations. Chapter 3 is an evolution of a paper published in the Journal of Revenue and Pricing Management in 2021 (Schubert et al., 2021).

Chapter 4 tests the first fundamental hypothesis to validate the methodology from Chapter 3 with inductive research of 496 million real airline coupons between 2018 and 2023. Observations are generalized to patterns and the results are discussed. The chapter concludes the first fundamental hypothesis can be accepted: different customer segments behave significantly differently in their product choice behavior. Chapter 4 is based on a paper written and approved by the airline partner for publication.

Chapter 5 tests the second fundamental hypothesis to validate the methodology from Chapter 3 with deductive research. The 496 million real airline coupons are split into different time periods to run multiple experiments. The hypothesis is tested with three different prediction models of increasing complexity. Results are presented with three metrics, all of which support the conclusion that the second fundamental hypothesis can be accepted: choice probabilities of future customer searches can be predicted significantly more accurately when conducting high-dimensional segmentation with data readily available to airlines. Chapter 5 is based on a paper written and approved by the airline partner for publication.

Chapter 6 synthesizes and reflects on the findings. It answers the research questions and discusses generalization to other airlines, other time periods, and other sectors. The chapter closes with a discussion of implications on both academia and the airline and wider transportation industry.

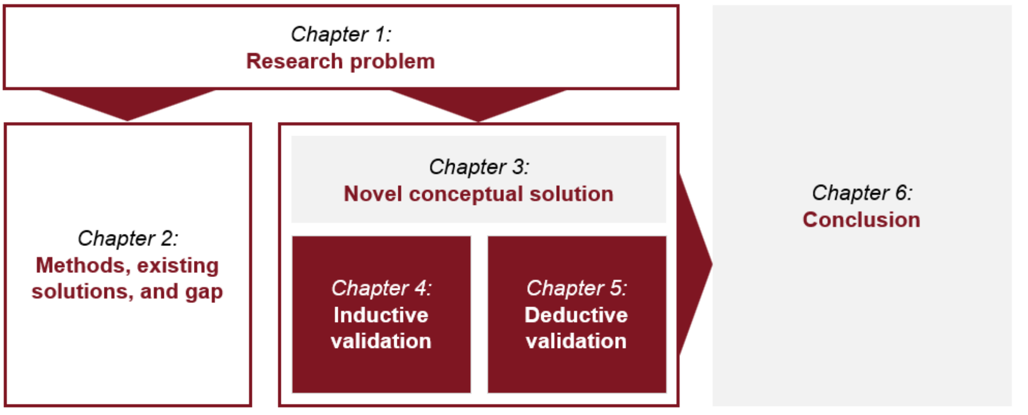

Figure 4: *Building the argument in the dissertation.*

The dissertation is a cumulative manuscript with Chapters 3-5 as sequentially published or written standalone papers. For that reason, there is some repetition of literature review and context setting when reading the full thesis.

Lists of tables, figures, and abbreviations are given in the beginning of the dissertation. References can be found at the end of this dissertation. An Appendix, if applicable, is included at the end of the respective chapter.

2 Methods, existing solutions, and gap

Having defined the research objective, Chapter 2 reviews methods and existing attempts to solve the customized airline offer management problem. Both machine learning and discrete choice models have been proven extremely powerful in solving various problems. Both were originally built to solve different problems but have more and more expanded into other areas as well. This renders a comparison especially valuable. Section 2.1 introduces artificial intelligence (AI) with its subfield machine learning (ML). Section 2.2 covers discrete choice (DC) models. In both sections, historical developments, theoretical foundations, and practical applications are reviewed, and the methods are evaluated against the research objective. Advantages and disadvantages with respect to the customized offer management problem are synthesized to the research gap, refined research questions, research strategy and confidential airline data used, as well as the audience and relevance of this dissertation in Section 2.3.

2.1 Artificial intelligence and Machine learning

In a broad sense, AI describes the ability of machines to perform cognitive functions associated with humans like interacting with the environment, perceiving, reasoning, and learning. Several technologies have enabled AI to solve practical business problems, such as computer vision; language processing and translation; robotics and autonomous vehicles; and machine learning (ML). Before introducing the most important classes of ML algorithms

(Section 2.1.2), this section provides historical context on the evolution of AI (Section 2.1.1). Data strategies and model selection are discussed in Section 2.1.3. The section closes with zooming in on the class of recommender systems (Section 2.1.4) and especially the matrix factorization algorithm (Section 2.1.5).

2.1.1 Enablers of AI (r)evolution

The evolution of AI has been enabled by three major developments, namely the advancement of algorithms, explosion of data available, and the exponential increase in computing storage and power. The first algorithm capable of learning independently was the perceptron algorithm developed by Rosenblatt (1958), which laid the groundwork for the later development of artificial neural networks (ANNs). One year later, Samuel (1959) coined the term "machine learning" as "a field of study that gives computer the ability to learn without being explicitly programmed". Model training advanced significantly with the introduction of backpropagation, allowing ANNs to optimize themselves without human intervention (Rumelhart et al., 1986; Werbos, 1990). The PageRank algorithm by Page et al. (1999) enabled the ranking of web pages. It was the initial prototype of Google's search engine and paved the way for increased consumption of the internet. Deep learning techniques were further popularized by Hinton et al. (2006) facilitating efficient training of these models.

In 1991, the European Organization for Nuclear Research (CERN) began opening the World Wide Web to the public. However, it was not until the early 2000s that broadband adoption enabled widespread internet usage, surpassing one billion users in 2006 (McKinsey, 2020b). The Web 2.0 era shifted the focus from passive content absorption to interactive content creation. This enabled business models like Facebook and YouTube, which debuted in 2004 and 2005, respectively. The introduction of the iPhone in 2007 fueled the smartphone revolution and massively propelled data generation. Global smartphone sales reached 300 million in 2010, representing a nearly 2.5 times increase from

2007. Already by 2014, the number of global mobile devices surpassed the number of humans (McKinsey, 2020b).

As early as 1965, Moore's law (Moore, 1965) recognized the exponential growth in chip power. In 1997, increased computing power contributed to IBM's Deep Blue defeat over world chess champion Garry Kasparov. The release of the world's first graphics processing unit (GPU) by Nvidia in 1999 fundamentally increased computing speed compared to computer processing units (CPUs). The ability to leverage GPUs for ML tasks was further facilitated by Nvidia's Compute Unified Device Architecture (CUDA) framework introduced in 2007. Another major advancement came with Google's introduction of upgraded tensor processing unit (TPU) technology in 2016-2017 (McKinsey, 2020b).

Boser et al. (1992) demonstrated how support vector machines (SVMs) can handle nonlinear problems through a technique known as the kernel trick, presenting a major step forward in natural language processing. Speech recognition emerged as the primary application for recurrent neural networks (RNNs) and long-short-term-memory (LSTM) models (Hochreiter & Schmidhuber, 1997). While word embeddings were already used in 2001, Google's word2vec model (Mikolov et al., 2013) enabled faster training and paved the way for business applications such as analyzing survey responses and recommending products to consumers. Vaswani (2017) and Google AI's introduction of bidirectional encoder representations from transformers (BERT) in 2018 further progressed natural language processing. The You Only Look Once (YOLO) algorithm by Redmon et al. (2016) enabled huge progress in real-time object detection, paving the way for applications like self-driving cars, airport luggage scans, and intelligent traffic signals.

Generative AI is a transformative subset of artificial intelligence focused on creating new content (text, images, audio, or video) based on learned patterns from existing data. This technology allows machines not only to learn from data but also to generate new information that resembles their training inputs. It has led to significant advancements solving problems across various industries. One landmark development in this area was Goodfellow et al. (2014) who introduced generative adversarial networks (GANs) in 2014. GANs consist of two neural networks competing against each other to produce increasingly realistic data. GANs have proven particularly effective in addressing challenges related to insufficient training data. OpenAI's generative pre-trained (GPT) model released in October 2022 has facilitated improved natural language generation capabilities that enable near-human quality content creation across applications such as chatbots and automated writing tools. Further applications include creative fields like art and music composition through tools like DALL-E and Jukedeck, which generate unique visual content and musical compositions based on user prompts. Generative AI has begun to reshape healthcare by aiding in drug discovery and personalized medicine as well as marketing, where it enhances content creation and customer engagement strategies. As organizations increasingly adopt generative AI tools such as ChatGPT and Google's Gemini (previously named Bard until February 2024) for diverse applications ranging from text generation to visual art creation, it is projected that generative AI could add between $2.6 trillion to $4.4 trillion worth of value to the global economy by optimizing workflows and enhancing creativity across various domains (Turing, 2024).

The introduction of Google’s MapReduce algorithm (Dean & Ghemawat, 2004) presented a breakthrough, which substantially increased the amount of data that can be processed. The launch of Amazon Web Services in 2004 enabled widespread access to powerful IT systems through cloud storage and computing solutions. With Microsoft Azure and Google Cloud Storage further cloud services entered the market in 2010. UC Berkeley’s Spark revolutionized real-time analytics capabilities by allowing seamless updates to big data (Zaharia et al., 2010). Google’s TensorFlow deep learning framework was open-sourced in 2015 and enabled developers to collaboratively build scalable deep learning models. Similarly, Facebook open-sourced PyTorch, granting researchers access to even more deep learning algorithms. Many resources offer open-source, built-in ML algorithms pre-programmed for implementations in Python or other programming languages. These developments have significantly accelerated innovation within AI. They foster an environment in which more models can be created and applied to an ever-growing array of problems. These problems include the transportation industry, where McKinsey (2020c) estimates hundreds of billions representing 1-8% of revenue potential from deploying AI across personalized offering as well as pricing and promotion. For passenger airlines, McKinsey (2024) estimates the potential net worth from implementation of improved retailing at 45 billion USD by 2030, equivalent to 2-3% of revenue or 15% of EBITDA.

In summary, the evolution of AI has been driven by advancements in algorithms, data availability, computing power that can process larger amounts of data, and more recently, generative AI reshaping industries and enhancing creative processes. As organizations continue to explore these innovations while navigating ethical considerations surrounding their implementation, generative AI stands at the forefront of transforming how humans and technology interact to solve increasingly complex problems.

2.1.2 Machine learning (ML)

Machine learning (ML) is generally defined as subset of AI (Alpaydin, 2020). ML algorithms detect patterns in datasets unable for human eyes to spot. This way, they learn to make predictions and give recommendations. In contrast to rule-based approaches, ML algorithms do not need to receive explicit programming instructions. Instead, they adapt to new data, enabling them to improve accuracy and effectiveness over time. Different types of analytics in increasing order of complexity are descriptive, predictive, and prescriptive analytics (McKinsey, 2020b). Descriptive analytics describes what happened and is employed widely across industries. Predictive analytics aims to anticipate what will happen in a probabilistic manner. Data-driven organizations use predictive analytics as crucial source of insight. Prescriptive analytics provides recommendations on how to achieve defined goals and is deployed by leading data and internet companies. ML focuses on predictive and prescriptive analytics.

ML models have been proven to be powerful in solving several real-world business problems. To name just a few, some well-documented ML applications include face recognition (Kaur et al., 2020), computer vision (Szeliski, 2010), self-driving cars (Eliot, 2017), machine translation (Bahdanau et al., 2014), neuroscience (Mnih et al., 2015), social network analysis (De et al., 2012), personalized marketing (Mari, 2019), and online retailing (Weber & Schütte, 2019). Designing customized offer management strategies for every individual airline search is methodologically similar to classification and recommendation tasks that have already been solved. ML algorithms are often categorized into supervised, unsupervised and reinforcement learning techniques, which will be subsequently discussed next.

First, **supervised learning (SL)** processes labeled input data to detect the (latent) relation between given inputs and given outputs. Goal is to accurately predict the label/output of data not seen before. In an iterative process, SL algorithms automatically generate identifying characteristics from the examples fed into them without requiring the programmer to input task-specific rules. SL can be viewed as algorithm to approximate complex and hidden functions, making it very powerful for regression and classification tasks of new incoming data (Figure 5). Well-known SL algorithms include linear and logistic regression (Cramer, 2002); naïve Bayes (Joyce, 2003); linear and quadratic discriminant analysis (Cohen et al., 2003); decision trees and random forests (Fawagreh et al., 2014); support vector machines (SVM; Boser et al., 1992); and many more.

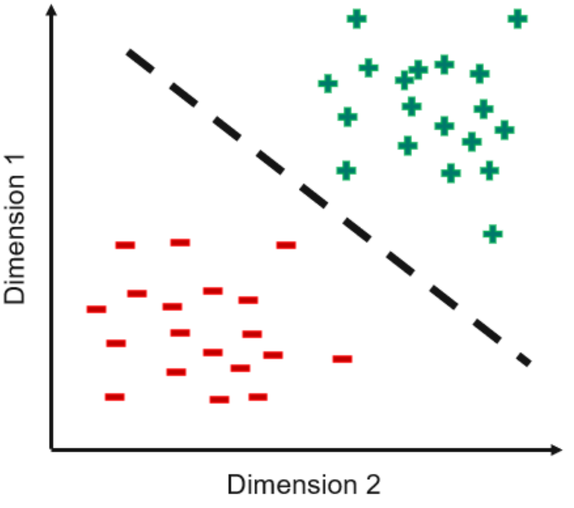

Figure 5: *Supervised learning is powerful for classification tasks, labeling new incoming data as either a red minus or a green plus, separated by the black dashed line.*[2]

Second, **unsupervised learning (UL)** aims to provide a compact representation of unstructured data by detecting patterns that humans cannot see. UL models infer structure from the data by identifying groups of data with similar behavior (Figure 6). As opposed to SL, UL does not require human supervision in the form of labeled data, but autonomously develops meaningful classifications and reveals latent relationships instead. Popular UL algorithms include K-means and hierarchical clustering (Nagpal et al., 2013); recommender systems (Section 2.2.1.4); dimensionality reduction with various algorithms, such as autoencoders (Hinton & Zemel, 1994) or principal component analysis (PCA; Shlens, 2014); and many more.

[2] The black dashed line marks the boundary between the two classification outcomes. New data to the left and below the line is classified as a red minus. New data to the right and above the line is classified as a green plus.

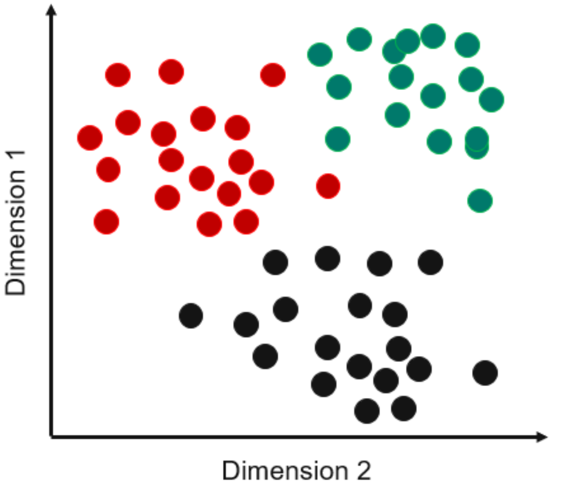

Figure 6: *Unsupervised learning groups data into clusters of similar behavior.*[3]

Third, **reinforcement learning (RL)** seeks to learn optimal behavior maximizing a reward function (Wang et al., 2016). At each state, the algorithm chooses amongst a set of potential actions and receives a (positive or negative) reward for the particular action chosen. Learning from experience, the algorithm is finally trained to follow an optimal policy of actions in each possible state. RL is applied in cases where training data is sparse and thus interacting with the environment presents the only strategy to learn. An important trade-off in RL is exploitation of known positive rewards vs. exploration of states not visited before. An interesting variant is inverse RL (Zhifei & Joo, 2012), which attempts to infer an unknown reward function from observing behavior. Inverse RL consequently helps detect those (unobservable) preferences most consistent with the (observable) choices people make.

Deep learning is often viewed as further subfield within ML (Deng & Yu, 2014; Nvidia, 2016). Deep learning techniques were particularly brought forward by LeCun (1998), Hinton et al. (2006), and Krizhevsky et al. (2012). Deep learning progressively extracts signals from raw inputs through multiple steps, so-called

[3] In this example, each data point, represented by dots, is grouped into precisely one cluster, either the green, red, or black.

layers, thereby learning more and more complex structures in the data. Artificial neural networks (ANNs) are the basis of deep learning algorithms. They mimic the functioning of neurons in human brains (Lake et al., 2017). Often, deep learning outperforms traditional ML approaches with respect to accuracy (McKinsey, 2020b), but requires much more data for training and introduces the risk of over-engineering for simpler applications. Popular deep learning algorithms and respective use cases are convolutional neural networks (CNNs) used to infer information from images (Albawi et al., 2017); recurrent neural networks (RNNs) applied to learning from time-series data, text generation or language translation (e.g., Mikolov et al., 2010); generative adversarial networks (GANs) capable of creating artificial data representative of actual data (Bengio et al., 2014; Goodfellow et al., 2014); and many more.

2.1.3 Data strategies and model selection

Setting up ML algorithms splits into three phases, namely initial model training, cross-validation, and model testing with data unseen before. For this purpose, the data available need to be split. For example, in the specific application to customized airline offer management, **initial model training** could happen on data on past booking behavior. With these alone, unsupervised learning could detect patterns between various searches and to perform dimensionality reduction. Supervised learning would additionally require information on which product was purchased and at which price. To estimate the search-specific WTP, price paid could be viewed as lower bound. Including data on searches not successfully converted to bookings could improve training. For **cross-validation**, a share of the training data would need to be eliminated from the set used for model training. Instead, the withheld data would be used to cross-validate the model. Cross-validation is particularly important to compare several algorithms for the same application and can therefore be used to identify those models with the proper trade-off between overfitting and underfitting (see Raschka, 2018, for definitions and an overview). Finally, **model testing** confronts the trained and cross-validated model with new data that have neither been used for training nor for cross-validation. Testing allows parameter fine-tuning and can also facilitate continuous model improvements through periodical parameter updates. Parameter updating could deploy exponential smoothing (Gardner Jr, 1985) to prioritize more recent observations over older ones. In the airline offer management use case it might be able to capture potential changes in customer preferences and trends.

Often, different models can be applied to the same business case. Deciding which algorithm to use carries important consequences and depends on the nature of the data available. Often, certain algorithms are precluded because of lacking adequate data. In this case, organizations could opt for systematic data collection in the future. The rise of ML has caused many organizations to rethink their data strategy, to pay more attention to what they could do with the data they have available, and how data from many different sources could be integrated into a common company-wide framework. With respect to the research objective at hand, model development, training and cross-validation initially would need to happen on past data. If the model looks promising and performs well in cross-validation, it might be possible to conduct live test runs for model testing. Such live test runs could happen in a well-defined subsample of distribution channels and/or routes. Furthermore, the test runs could be framed in a smart way to gain exactly the information and data desired. It is important to understand that many ML algorithms are likely to produce disappointing results in the real world at first. However, testing with actual data allows them to significantly improve over time and likely achieve high degrees of performance.

Further starting points for algorithm selection are similar business problems that have been approached with ML in the past, or an analysis which models fit best with company strategy or culture. Identifying the “best” model for a particular application means resolving the trade-off between model simplicity and explanatory power. Therefore, this research suggests starting with simpler approaches, assess their power, and then move on to more complex models. For the **recommendation of customized bundles,** the goal is to predict which of all possible bundle combinations best fits to the particular customer search. Learning happens based on specific characteristics, or features, included in the search. For this question, matrix factorization, an unsupervised recommender system technique, will be analyzed next.

2.1.4 Recommender systems

Recommender systems are crucial to the commercial success of platforms like YouTube, Amazon, Netflix, and are behind most of today's webpages. They are deployed to recommend movies, articles or items based on preferences of other consumers with similar attributes (Frias-Martinez et al., 2006; Frias-Martinez et al., 2009; Kim et al., 2010). To identify the relevant data necessary to make a recommendation, these algorithms detect cluster behavior in user data. Purchase or like probabilities are estimated and used to suggest users content relevant to their specific characteristics. Ricci et al. (2011) provide a comprehensive overview of recommender systems in digital platforms. Recommender system algorithms split into two paradigms, namely content-based and collaborative filtering.

Content-based filtering (CB) methods infer user-specific recommendations of additional items of interest by analyzing the properties of previous items and their respective user ratings (Basu et al., 1998; Park et al., 2012). A well-known example is the Music Genome Project, in which an expert scores individual songs alongside a set of characteristics (Koren et al., 2009). However, CB methods can be costly due to the explicit expert involvement and may be prone to overspecialized recommendations including only items that are very similar to those a particular user is already aware of (Lopez-Nores et al., 2008).

Initial research on recommender systems was spurred by advances in **collaborative filtering (CF)** in the mid-1990s (Resnick et al., 1994; Shardanand

& Maes, 1995). CF methods are based solely on past interactions between users and items, stored in a user-item interaction matrix. In this paradigm, the interaction matrix is deemed sufficient to detect similar users and products, and to derive predictions and new recommendations based on these estimated proximities (Breese et al., 2013). The properties of the items are not explicitly modeled. This way, data aspects can be adequately captured that would be elusive and difficult to profile with CB (Koren et al., 2009). In an e-commerce setting, the interaction matrix maps buyers to products via the probability of a particular type of buyer to purchase a particular type of product. However, CF methods have their own drawbacks, namely sparsity, scalability, and the cold start problem (Claypool et al., 1999; Koren et al., 2009; Sarwar et al., 2000a; Sarwar et al., 2000b). CF methods split into two primary strategies, neighborhood and latent factor models (Koren et al., 2009). Neighborhood models compute the relationships either between users or between items. Latent factor models, on the contrary, explain user preferences by characterizing both users and items on a third dimension, so-called factors, that are inferred from the rating patterns. This way, they provide a computerized alternative to expert scoring in CB methods.

2.1.5 Matrix factorization

Matrix factorization belongs to the class of latent factor models. Its initial idea is commonly ascribed to Funk (2006). Matrix factorization algorithms combine scalability, predictive accuracy, and flexibility, and are behind some of the most successful realizations of latent factor models, superior to neighbor techniques as demonstrated by the Netflix prize competition (Koren et al., 2009). The Netflix prize was a competition for the best algorithms to predict user ratings for movies, based solely on previous ratings. Whilst explicit feedback, such as persons' ratings for Netflix movies, is the most desirable, it usually comes in very sparse form, because only few people have seen a particular movie. One strength of matrix factorization is the incorporation of implicit feedback, i.e. inferring user preference from their observed behavior history. Since implicit feedback comes in the form of presence or absence of an event, i.e. whether a certain product is purchased by a certain customer or not, it typically results in a densely filled matrix (Koren et al., 2009). Matrix factorization decomposes the user-item interaction matrix into the product of two rectangular matrices of lower dimensionality by introducing latent factors stored in a vector F. This will be detailed for the research question of which airline customer segment purchases which airline product in the following.

Fundamental idea behind matrix factorization is that dependencies exist on two dimensions. First, customer searches are (dis)similar to each other. Although many customers do not have a history of multiple purchases, all possible *customer search types* exhibit some (dis)similarity. Second, products, i.e. the set of *possible bundles*, are (dis)similar to each other. Both (dis)similarities combined can be used to predict preferences of new customer searches and/or new products. Instead of directly mapping customer search types to bundles, an *intermediate latent factor vector* F is introduced. Typically, the number of latent factors f is much smaller than both the number of customer search types c, and the number of possible bundles available b as shown in Equations (1) and (2):

$$f \ll c \quad (1)$$

$$f \ll b \quad (2)$$

The relationship between the search type vector C and the bundle vector B is decomposed into the dot product of the relationships between C and the latent factor vector F, and between the transpose F^T and B, according to Equation (3).

$$C \cdot B = (C \cdot F) \cdot (F^T \cdot B) \quad (3)$$

Factorization requires estimation of much fewer model parameters due to the properties established in (1) and (2). The number of possible search types is exponential in the number of features and further depends on the number of values each can assume. Assuming ten features, each of which can take five possible values, results in approx. 9.8 million search feature combinations (5^{10}). Further assuming eight ancillary services results in 256 possible bundles (2^{8}). Factorization reduces the number of entries to be estimated and stored from 2.5 billion to less than 98 million (Figure 7).

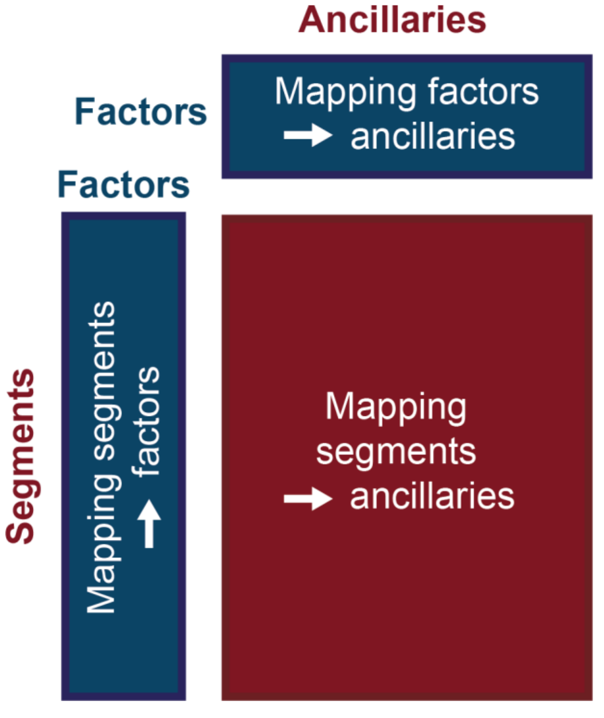

Figure 7: *Schematic depiction of the idea behind matrix factorization.*

The entries in F do not need to have an intuitive interpretation, although they could be thought of as some underlying customer preferences like convenience, flexibility, vacation, value of time, comfort, connectivity, entertainment, etc. During training, the algorithm learns the links from search types to the factors and from factors to the bundles. Learning happens via definition of an error function comparing model-predicted purchase decisions with actual ones, and subsequent minimization of the error function with the stochastic gradient descent method, which can be traced back to the Robbins-Monro algorithm of the 1950s (Bottou, 1998). The number of latent factors represents a hyperparameter, and the best choice needs to be determined during cross-validation.

The search type vector C contains one row for every possible combination of search feature. Such highly dimensional input space requires a large set of training data. This phenomenon is known as **“curse of dimensionality”,** an expression coined by Bellman (1957). Depending on how many features shall be included in the model and how many values each can assume, addressing the curse of dimensionality with dimensionality reduction strategies might be required before training the actual matrix factorization algorithm. The basic idea of dimensionality reduction is to compactly summarize data in fewer dimensions whilst preserving as much explanatory power as possible. Subsequently, the created compact description can be used as input to the actual matrix factorization model. A well-documented algorithm for dimensionality reduction is **principal component analysis** (PCA; Shlens, 2014). PCA projects each data point onto a new dimension of maximum variance. Afterwards, it serially projects each data point on additional, orthogonal dimensions. This transformation maintains dimensionality by simply sorting the new (and mutually orthogonal) parameters according to descending variance. With that information, one can then do filtering through feature selection, i.e. ignore all features with sufficiently low variance.

Given the power of matrix factorization to predict the attractiveness of all possible product combinations for new incoming searches, it is a candidate to enable customized offer management at much higher levels of granularity than existing approaches. All products could be ranked, and this ranking could be used as assortment optimization criterion. This way, the best out of both branded fares and *a la carte* ancillary offerings could be combined. Whilst branded fares are inflexible and unlikely to provide the best fit for most customer searches, a well-trained matrix factorization algorithm might be able to output customized bundle recommendations much closer to the actual preference of the incoming search. Also, as opposed to *a la carte* sales, customized bundle recommendations avoid overwhelming the customers. After all, customized airline offer management follows the logic of e-commerce platforms or movie recommendations. Matching consumers with their preferences is hence not only key to the success of Amazon, Netflix, YouTube, and many more, but possibly also to airlines or transportation companies in general.

After reviewing the history of AI and ML as well as possible applicability to the research question, the next section will review discrete choice models, how they could be applied to the research question, and how they compare to machine learning.

2.2 Discrete choice (DC) analysis

Discrete choice (DC) models are statistical models to explain choice behavior from a set of discrete and mutually exclusive alternatives. The clear definition of possible choices distinguishes the concept from standard regression models. Classic DC literature includes McFadden (1973), McFadden (1974), Domencich & McFadden (1975), Winston (1985) and Boyer (1998). DC models have a long history in transportation planning for mode choice modeling and demand analysis (Train, 1978), the choice of airports for cargo airlines (Kupfer et al., 2016) or customer preferences for mobility as a service (Polydoropoulou et al., 2020). Other applications include labor market economy and education choices (Fuller et al., 1982), conjoint analysis for market research (Train, 1986), recreation (Train, 1998) and energy systems (Goett et al., 2002; Revelt & Train, 1998). DC models aim to characterize the utility function for the population to allow statistical inference about the functional parameters.

This section starts with DC model setup and fitting (Section 2.2.1), continues with a discussion of the applicability to the customized airline offer management problem (Section 2.2.2) and closes with a comparison of DC and ML models (Section 2.2.3).

2.2.1 Discrete choice model setup and fitting

DC models estimate U_{ij}, i.e. the (latent) utility of choice j to person i, based on observed characteristics of the person X_i, observed characteristics of the choice Z_j, and an error term ε_{ij} representing unobservable influences. Equation (4) shows the generic DC model:

$$U_{ij} = F\left(X_i, Z_j, \varepsilon_{ij}\right) \tag{4}$$

F can assume several functional forms. Many empirical studies rely on logit models commonly estimated with maximum likelihood iterative processes (Blauwens et al., 2016). A simple structure of a logit model is given by Equation (5):

$$\ln\left(\frac{P_i}{1} - P_i\right) = \beta_0 + \sum_{k=1}^{n} \beta_k \cdot X_{ki} + \varepsilon_i \tag{5}$$

P_i represents the probability that person i chooses a certain alternative based on several explanatory variables X and estimated parameters β. Probabilities are introduced to account for the randomness of individual utilities.

Individuals are assumed to rationally choose the alternative with highest utility. To obtain the functional form, a specific distribution of the error term needs to be assumed. If errors are assumed to be distributed according to the extreme value distribution, the **multinomial logit (MNL) model** results, which is the most applied form in demand analysis (Fisher, 2000). MNL was originally introduced by Luce (1959) with McFadden (1973) popularizing its practical use. An overview of common MNL applications is provided by Winston (1985). If instead errors are assumed to follow the normal distribution, the **multinomial probit**

model results. Recent advances in DC models have been driven by growth in computer power and the use of simulation, allowing for unprecedented flexibility in model form (Walker & Ben-Akiva, 2011).

DC models implicitly assume independence from irrelevant alternatives (IIA). If IIA holds, the choice set presented to individuals does not influence their behavior. However, IIA is often violated in practice, with important implications on assortment optimization. Details on model fitting to handle situations in which IIA is deemed too strong can be found in Bruch & Mare (2012).

To obtain data of customer behavior, researchers can use stated (SP) or revealed preferences (RP) studies (Zamparini & Reggiani, 2007; Blauwens et al., 2016). In SP studies, researchers infer people's real-world behavior from their responses to specifically designed surveys in laboratory settings. As such, SP studies can help collect data on settings that do not exist in reality (yet). On the contrary, RP studies detect preferences directly from actual decisions people make, but RP data might not be available for completely new choice sets. In some settings, RP and SP can also be combined to improve accuracy (Zamparini & Reggiani, 2007).

2.2.2 Applicability to the customized airline offer management problem

With their history and previous applications, DC models are obvious candidates to be applied to customized airline offer management. Ben-Akiva & Gershenfeld (1998) provided examples for realistic DC models of customers choosing among different bundle options. Vulcano et al. (2012) applied MNL models explicitly in an RM context, and Ratliff & Gallego (2013) evaluated the sales and profitability impacts of airline branded fares and their pricing using a customer choice framework. In particular, the PODS research consortium has developed and applied DC to airline passenger choice behavior (PODS, 2021).

With this research objective, search features are assumed to represent the characteristics of the decision-maker. In the case of modeling choice for airline branded fares, the choice set typically includes three to four alternatives. In the case of modeling choice for all possible airline products, the choice set can include hundreds of potential bundles. In this case, data might become sparse, rendering statistical modeling more time-consuming or even impossible. To address, several techniques are available for the selection of choice subsets (Guo & Loo, 2013; Wang, 2020). However, selecting choice subsets implicitly assumes validity of the IIA assumption, which was shown to be violated due to various psychological effects that impact consumer choices (Jannach et al., 2010).

2.2.3 Comparing Discrete choice and Machine learning

Generally, literature reports that both DC and ML methods are suitable for both prediction and inference although there are some important differences (synthesized in Table 1). DC models aim to draw population inferences from samples focusing on explanation of the relationship between variables. They are widely used to explain how preferences impact a decision (Paredes et al., 2017). DC models create and fit a problem-specific probability model, which provides researchers with confidence that the discovered relationships adequately describe the “true” effect (Bzdok et al., 2018). If sufficient data are available, the underlying assumptions can be verified. ML models rather focus on prediction, trying to find generalizable patterns in rich and unstructured data (Bzdok, 2017; Bzdok et al., 2017; Feldman et al., 2018; Paredes et al., 2017). ML employs general-purpose learning algorithms not specifically designed for a particular problem.

To achieve representability, DC models deploy precisely designed data collection strategies. ML, on the contrary, makes minimal assumptions about the data generation process and does not require data gathering with carefully controlled experimental designs (Bzdok et al., 2018). ML mainly ignores potential sampling issues, whilst still implicitly assuming representability of the data. However, cross-validation and out-of-sample testing are integral part of ML, but often ignored in DC models (Paredes et al., 2017).

ML often achieves convincing prediction results (Bzdok et al., 2018). Feldman et al. (2018) compared an MNL model and recommender system algorithms for optimal product assortment on Alibaba's online marketplace. They found MNL to outperform recommender systems. Due to its generic nature, ML fails to capture critical problem-specific nuances, whereas MNL models are specifically built for the purpose of capturing customer purchasing behavior, and to model substitution patterns in particular. Paredes et al. (2017) compared MNL to various ML algorithms (random forest and support vector machines) for the prediction of car ownership. They reported that ML can be both inferior and superior to DC, depending on which features are used. Wang & Ross (2018) studied travel mode choice modeling. They concluded overall higher prediction accuracy for the ML model with the extreme gradient boosting (XGB) algorithm compared to MNL.

Strict statistical assumptions can limit MNL models, despite their closed-form mathematical structure with interpretable estimation results and solid foundation in random utility theory. ML, on the other hand, lacks explicit theoretical foundation, sound behavioral theory, and often interpretability (Bzdok et al., 2018; Paredes et al., 2017). Also, Kleinberg et al. (2015) criticize that the outperformance of ML over traditional econometric models on prediction often comes at the cost of explainability. The black box character of many AI applications makes it difficult to understand why a machine makes a certain decision. Since machines are unable to explain their thoughts and actions to human users, "explainable AI" has emerged as research strand over the last years. For further development of ever-more complicated AI applications, it becomes more and more essential that users understand, trust, and effectively manage AI (Gunning, 2017). Samek et al. (2017) proposed two approaches to explain predictions of deep learning models. The first method computes the sensitivity of the prediction with respect to changes in certain inputs, and the second method decomposes the decision in terms of various input variables.

Table 1: *Synthesis of academic literature comparing discrete choice analysis and machine learning.*

	Discrete choice	**Machine learning**
Original purpose	**Population inference** from samples by explaining relationship between variables: how do preferences impact a decision?	**Prediction** by identifying generalizable patterns in rich and unstructured data.
Pros	**Problem-specific** model increases confidence the "true" effect is captured. **Closed-form** mathematical structure with interpretable results. **Verifiable** if sufficient data available.	Can handle **large and unstructured data**. Minimal assumptions about the data generation process, hence, **does not require carefully controlled experiment design.** Often achieves **high prediction accuracy**.
Cons	Require **precisely defined data collection.** **Strict statistical assumptions** can limit applicability. Typically, **no cross-validation or out-of-sample testing.**	**General-purpose algorithms**, typically not designed for a particular problem. Often fails to capture problem-specific nuances. **Ignores potential sampling issues.** Can be perceived as black box **lacking interpretability.**

The reviews have highlighted mixed results in the literature whether ML or statistical models like DC should be preferred. The discussion seems to suggest that ML might represent cases of over-engineering for simple problems, and that for these sound statistical methods underpinned by strong theoretical support could be the better choice. On the other hand, ML methods offer superior predictive power and the ability to handle large and unstructured data. For these reasons, it is unclear which of the two is better suited to solve the research problem of this dissertation of customized offer management. This substantiates the question whether and how to combine both DC and ML; or specifically, how ML could enhance DC models by improving their prediction accuracy, and whether this also holds true in disruptive times such as the first global wave of the Covid-19 pandemic in the first quarter of 2020.

2.3 Research gap, strategy, and relevance

This section synthesizes the research problem, the context of airline offer management, the research objective, and the review of AI, ML and DC. The section starts with the research gap (Section 2.3.1). Subsequently, the research questions from Section 2.1 are refined based on the additional context (Section 2.3.2). The research strategy with a proposed solution and two validation steps on real airline data is presented next (Section 2.3.3). The section closes with the audience and relevance of the research in Section 2.3.4.

2.3.1 Research gap

Previous literature and practical airline implementation can be categorized into one of two categories. Either they segment customers into single digit distinct segments, typically deploying discrete choice or other statistical models. Or they use machine learning models that can facilitate segmentation with many more segments but present a black box character that airline users struggle to understand, trust, and ultimately adopt. For instance, Shukla et al. (2019) expand beyond single-digit customer segments. Instead, they model them with various machine learning models based on temporal, market-specific, journey-specific and price-related features. However, this might be viewed as black box by users. Visibility and interpretation of the customer segments is however important to achieve adoption and increase the likelihood of embedding into existing business process workflows and methods (Vinod, 2020).

To propose a novel solution to the customized airline offer management problem, this research develops a solution that meets the two criteria defined in Section 1.3. First, offers need to be relevant to the specific customer search and displayed in a convenient way. Due to the heterogeneity of airline customers, single digits of segments are not sufficient. Pure discrete choice models are not granular enough. Second, the solution needs to be usable to airlines. Because of the increased dimensionality, full-scale automation and change management within airline organizations are required (Daft et al., 2021). To gain trust from airline users, the solution must not be perceived as untransparent and unexplainable black box. Hence, many complex machine learning models do not seem to fit well with existing airline processes and users as they are likely not easily understandable and interpretable.

Putting both together, this research aims for discrete segmentation with a large number of segments. DC models are designed for discrete segmentation. To also achieve a large number of segments, the research tests whether ML can help identify these granular segments and solve potential data sparsity problems.

Table 2: *Selected contributions of the customized offer management literature since 2010.*

Paper	Problem addressed	Method used	Number of segments
Ratliff & Gallego (2013)	Bundling, Ancillary / Bundle pricing	Discrete choice	4-5
Fiig et al. (2016)	Flight pricing	Discrete choice	2
Madireddy et al. (2017)	Bundling, Ancillary / Bundle pricing	Various machine learning models	7
Bockelie & Belobaba (2017)	Flight pricing, Ancillary / Bundle pricing	Discrete choice	2
Wittman & Belobaba (2018)	Flight pricing	Discrete choice	2
Shukla et al. (2019)	Ancillary / Bundle pricing	Various machine learning models	Infinite
Wang et al. (2023)	Bundling, Ancillary / Bundle pricing	Markov chain choice model	2
This dissertation	Bundling, Assortment	Various machine learning models in Discrete choice	Thousands to millions

Table 2 visualizes the research gap this dissertation aims to fill: thousands to millions of segments, with a discrete segmentation logic that is still easily understood by airline users (e.g., Airline Pricing or Ancillary Managers), who can 1:1 map and trace every customer search into precisely one distinct customer segment. To facilitate segmentation in this order of magnitude, segmentation no longer builds on fare rules and booking classes, but on identifying structure in customer searches and bookings. This structure shall help estimating customer choice probabilities for all possible products, which ultimately allows airlines to display relevant customized products bespoke to the particular customer search. Methodologically, the gap is to combine the flexibility of machine learning for high-dimensional segmentation with the understandability of discrete choice model with clearly identifiable segments.

2.3.2 Refining the research questions

With the additional context, the five research questions from Section 2.1 can be refined:

1. Can airlines segment their customers into thousands to millions of distinct, clearly identifiable and MECE segments that exhibit significantly different choice behavior?
2. Can airlines use this segmentation to significantly improve the prediction accuracy of customer choice probabilities for searches in the future?
3. Can matrix factorization help solve the data sparsity problems when segmenting customers into thousands to millions of segments?
4. Can changes in customer behavior be captured, or how much of the prediction accuracy improvements can be achieved in a disruptive event like Covid-19 pandemic?
5. What is practical advice to balance cost and effectiveness: Which features should airlines train on? How complex should the prediction model be? How many segments should airlines use? How long should the training period be? How often should airlines retrain their model?

Answering these five research questions shall help airlines discuss practical implementation as well as indicate future research avenues for academic scholars.

2.3.3 Research strategy and data

With the research gap identified, this dissertation proposes a novel solution to the customized airline offer management research problem in Chapter 3. To validate the proposed solution, the dissertation uses a combination of inductive (Chapter 4) and deductive (Chapter 5) research as outlined in Table 3 and Figure 8. Real airline data is used for validation in both the inductive and deductive research. The author had access to hundreds of millions of bookings from a major network airline between 2018 and 2023. Compared to LCCs, network airline customers can be assumed to be more heterogeneous, hence segmentation offers larger opportunities and is of higher relevance.

These bookings cover global customers; however, due to the airline's network structure, its home markets constitute the biggest geographical share. The period between 2018 and 2023 allows validation in both stable and disruptive times considering the Covid-19 pandemic. This facilitates insights into how the solution performs when the market environment changes drastically and rapidly. The airline data include information on trip-specific features like route (origin, destination), travel weekday and season of the intended travel. Further, customer- or search-specific features include days before departure, search weekday, season of the search, number of travelers requested, and sales channel. Due to data confidentiality, no detailed descriptive statistics are shown.

Table 3: *Comparing inductive and deductive research as used in this dissertation.*

	Inductive research	**Deductive research**
Aim	Validate viability of proposed conceptual solution	Substantiate validation of the viability of the conceptual solution
Method	Find structure in the data. Generalize empirical observations to patterns	Test specific hypotheses on new data (out-of-sample test)
Outcome	Confirm fundamental hypothesis 1: *different segments exhibit significantly different choice behavior* Theorizing about patterns leads to tentative hypothesis to be tested in the follow-up deductive research	Confirm fundamental hypothesis 2: *segment-specific choice probabilities help airlines predict future customer choice with significantly higher accuracy*

Inductive research (Chapter 4) starts with empirical observations of the data. Then, it seeks to identify patterns in those observations. Theorizing about these patterns leads to generalizable hypotheses, which are then tested in the follow-up deductive research. In this dissertation, the goal of the inductive validation is to test the first fundamental hypothesis of the proposed conceptual methodology from Chapter 3, namely that different customer segments exhibit significantly different choice behavior.

The deductive research (Chapter 5) starts with the theory from the inductive part and tests the hypothesis on new data in out-of-sample tests. In this dissertation, the specific goal of the deductive validation is to test the second fundamental hypotheses of the proposed conceptual methodology from Chapter 3, namely that segment-specific choice probabilities help airlines predict customer choice behavior in the future significantly more accurately.

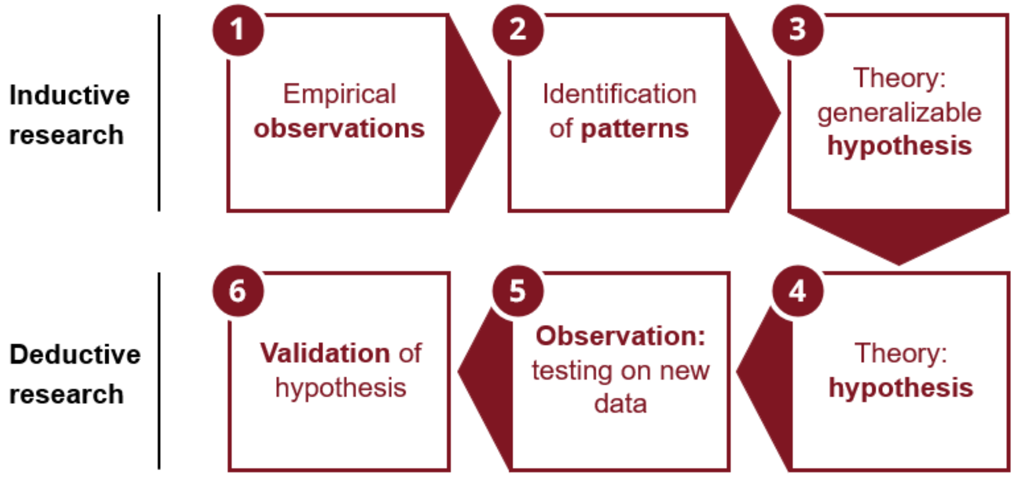

Figure 8: *Comparing inductive and deductive research and how they together validate the proposed solution.*

In summary, inductive and deductive research are designed as two-step validation of the two fundamental hypotheses of the conceptual methodology developed in Chapter 3.

2.3.4 Audience and relevance

Primary research audience are academic scholars as well as airline executives and managers in offer management or the subproblems segmentation, bundling, ancillaries, pricing, or revenue management. If airlines can improve prediction accuracy of discrete choice probabilities for future customer searches, then they can improve their assortment decision. Displaying more relevant offers can be expected to optimize customer or business outcomes. These outcomes can be higher search-to-book conversion, increased customer satisfaction, higher seat load factor, higher expected revenue, or higher expected profit per customer search.

Similar challenges are also present in other transportation companies. This was confirmed by an “Industrial Committee”, which the author consulted at the beginning of the PhD research project in January 2021. The Industrial Committee was comprised of representatives of passenger airlines, cargo airlines, ocean carriers, car rental companies, passenger train operators and mixed passenger/cargo ferry companies. Its representatives confirmed the offer management problem at their respective companies can also be described by the three characteristics identified in Section 1.1, though likely at different intensity. First, their customers are heterogeneous with respect to the product they are looking for as well as their willingness to pay. Second, the companies can customize their product to meet these diverse needs. And third, they are confronted with a large number of searches, which requires an automated and real-time response.

Hence, the research implications extend beyond airlines to scholars and practitioners of these transportation sectors as secondary audience. Potential research generalizations include studies on applications in these further transport modes like ferry operators, auto trains, high-speed rail, and possibly even pure freight carriers or logistics companies.

3 Conceptual methodology[4]

Revenue management has been a strategic priority for airlines for decades. More recently, ancillary revenues and targeted offers for diverse customers have become increasingly relevant. Hence, airlines need to evolve their revenue management into offer management, complementing flight pricing with dynamic bundling, ancillary pricing and assortment. This chapter designs a conceptual architecture for customized real-time offer management, expanding on existing literature in several directions. It proposes a machine learning framework for highly granular customer segmentation, whilst addressing three potential challenges of data sparsity, curse of dimensionality, and model understandability. The goal is to enhance customer satisfaction through customized offers and to improve pricing decisions by accurately predicting customers' willingness to pay. Embracing this approach shall position airlines to better meet customer needs and maximize revenue potentials.

In the context of the dissertation, Chapter 3 suggests a conceptual methodology to solve the customized offer management problem. It starts with an introduction and research objective in Section 3.1. Section 3.2 reviews academic and practical advancements that have enabled customized offer management. Section 3.3 describes the research contribution. Section 3.4 conceptualizes the high-level offer management system architecture proposed. Section 3.5 and 3.6 outline the two streams of the architecture, namely Product choice and Flight pricing. Section 3.7 summarizes innovations of the proposed architecture. Section 3.8 suggests strategies to validate and disseminate the conceptual architecture. Section 3.9 discusses practical implementation for airlines. Section 3.10 concludes the conceptual methodology.

[4] This chapter is an evolution of a paper published in the Journal of Revenue and Pricing Management (2021): Schubert, D., Sys, C. and Macário, R., 2021. Customized airline offer management: A conceptual architecture. *Journal of Revenue and Pricing Management*, pp.1-11.

3.1 Introduction and research objective

Airline customers behave differently with respect to both the services they demand and their willingness to pay (WTP) for these services. In response, airlines customize their product, complementing the mere seat with a multitude of ancillaries. These include seat reservation, baggage, flexibility, fast lane through airport security, lounge access, inflight entertainment, ground transportation, pollution compensation, and many more. Identifying which of these to be displayed to a given customer request at which price poses a challenge with various subproblems, namely bundling, pricing, and assortment. Bundling refers to creating a product bundle, i.e. an unbreakable entity of airline seat plus ancillaries with one single price tag. Pricing should match the individual customer WTP and can be decomposed into flight and ancillary pricing. Assortment refers to the selection of offers to be presented to the customer.

This chapter uses the term "customized offer management" to group the challenges of bundling, flight pricing, ancillary pricing, and assortment. Customized offer management processes the information customers explicitly and implicitly provide in their booking request to develop product and pricing strategies responsive to the particular request in real-time. Combining the convenience and simplicity of branded fares with the flexibility of unbundled ancillary sales, it expands on traditional rule-based segmentation, pricing and differentiation. More relevant offers might increase both customer satisfaction and WTP as well as help airlines escape the commodity trap. In addition, more precise WTP estimation could increase airline profits. Customized offer management requires new methods integrated into existing airline systems. Airlines need to evolve revenue management systems (RMS), focusing on flight pricing, into comprehensive offer management systems (OMS) including ancillary pricing, bundling and assortment.

This research develops a conceptual OMS architecture. It proposes machine learning (ML) for segmentation with much higher granularity than existing models, to support product and pricing strategies as customized as possible. To overcome resulting data sparsity problems and increase robustness, this research suggests application of the matrix factorization algorithm in a novel way. To address the curse of dimensionality [5] and improve both model understandability and applicability, the OMS is developed in modular form and breaks down segmentation into various subdimensions. Consequently, the proposed OMS aims to define a new balance between granularity/complexity and applicability/robustness of the segments identified.

The OMS is structured into two main *streams*, product choice and flight pricing. Product choice, i.e. which ancillaries which customers purchase at which prices, comprises bundling, ancillary pricing and assortment (Figure 9). Both streams could be independently embedded into existing airline RMS. To maximize practical relevance, the OMS processes data available to airlines and is consistent with the booking class logic of existing RMS. At the same time, it includes future-oriented developments such as continuous pricing that may eventually leave the booking class logic behind.

[5] Bellman (1957) coined the term “curse of dimensionality”. It refers to the observation that the volume space of numerous problems increases faster than the dimensionality of the inputs. Consequently, the amount of data required for sound analysis often grows exponentially with dimensionality.

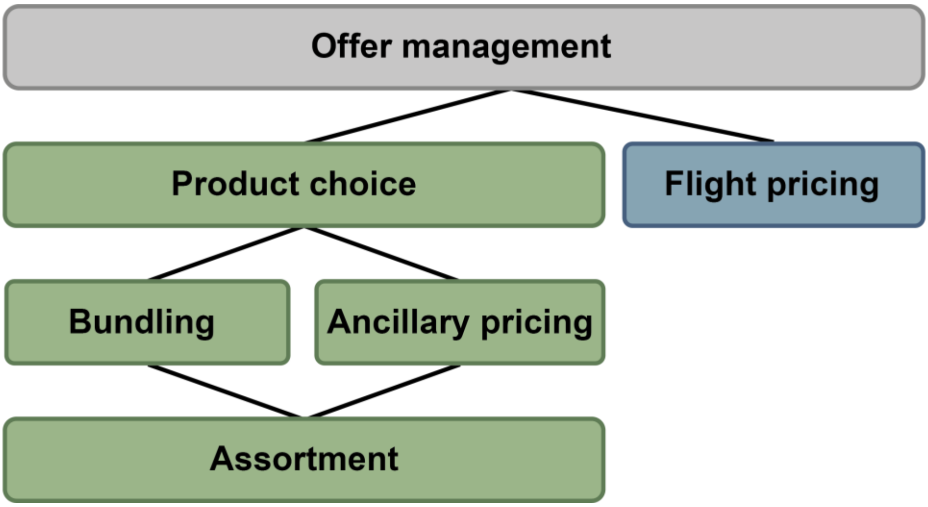

Figure 9: *Overview of the proposed offer management system (OMS).*

This chapter is structured as follows. Section 3.2 reviews existing research on different factors facilitating customized offer management to develop the research contribution in Section 3.3. Section 3.4 outlines the two-stream proposed OMS architecture, which is then detailed in the following Sections 3.5 for product choice and Section 3.6 for flight pricing, respectively. Section 3.7 highlights how the research expands on existing literature. Section 3.8 describes next steps for validation with actual airline data. Section 3.9 analyzes points of discussion for application in airline systems and potential extensions of the architecture. Section 3.10 summarizes the findings.

3.2 Enablers of customized offer management

Customized offer management builds on advancements in flight pricing, ancillary pricing, bundling, assortment, and distribution, as successively reviewed in the following.

3.2.1 Flight pricing

The origin of airline revenue management (RM) can be credited to Littlewood (1972) suggesting airlines to maximize revenue instead of load factors. Since then, both scholars and practitioners have constantly refined RM models (Vinod, 2021a). This chapter defines flight pricing to govern pricing of airline seats.

To match individual WTP, airlines segmented their customers with restrictive fare rules, leading to the assumption of independent demand in early RM models (Belobaba, 1987). The arrival of low-cost carriers (LCC) with unrestricted fares significantly reduced the effectiveness of restricted fares, requiring legacy airlines to adapt their RMS (Belobaba, 2011). Closely linked to advancements in "Distribution", leading airlines began to explore opportunities to overcome the rigidity of pre-defined price points. Fiig et al. (2016) proposed dynamic pricing responsive to airline strategy, real-time competitor information, and two different customer segments ("business" and "leisure"). Wittman and Belobaba (2018, 2019) noted that traditional airline pricing aims to optimize availability of given price points with rule-based segmentation instead of optimizing prices themselves. They developed a choice-based heuristic to dynamically amend pre-defined fares dependent on certain passenger and request characteristics. Continuous pricing, leaving price points entirely behind, is the ultimate vision of dynamic pricing (e.g., Lufthansa, 2018). It requires fundamental rethinking of airline RMS.

3.2.2 Ancillary pricing

Deregulation of air travel in the United States in 1978, before reaching nearly all parts of the world (Belobaba et al., 2016), fostered competition and stimulated the "pay-for-extras" business model of LCCs. Increased price competition led to uncoupled pricing of airline seats and ancillary services (termed "flight pricing" and "ancillary pricing", respectively) instead of earlier all-inclusive fares. This also concerns legacy airlines though many still position as “premium” by including some ancillaries in their base fares. Ancillary revenues, i.e. any airline revenue from services beyond the simple transportation of customers, have grown to 12% of total airline industry revenue pre-Covid (IdeaWorks, 2019). Ancillary services can be grouped into various types. Airline-own services include check-in baggage, seat reservation, increased legroom, priority boarding, and catering. Typically, baggage fees are the most relevant, composing 60% of total LCC ancillary revenues (IdeaWorks and CarTrawler, 2018). Other ancillaries concern facilities on the ground such as lounge access or fast lane through security. Finally, airlines sell third party services on commission. Examples are ground transportation, rental cars or hotel sales.

In contrast to quantitative models for flight pricing, most airlines employ static pricing for ancillary upsells. However, they have begun to experiment with dynamic ancillary pricing. An example is US LCC Spirit Airlines analyzing the impact of dynamic baggage fees based on search request, travel date, route and time of purchase (CAPA Centre for Aviation, 2019). Various providers developed dynamic pricing models with ML-driven pricing recommendations responsive to various request attributes. Examples are Shukla et al. (2019) and Kolbeinsson et al. (2021) reporting significant improvements in conversion and revenue over human rule-based approaches without processing any personalized information that might violate customer privacy. Since ancillary pricing might directly affect passengers’ choice for one airline over another, researchers designed joint models for flight and ancillary pricing (Bockelie, 2019; Hao, 2014; Ødegaard and Wilson, 2016).

3.2.3 Bundling

To differentiate from competitors and to manage their ancillary portfolio, airlines apply two strategies. First, unbundling with *a la carte* ancillaries offers maximal flexibility. Second, bundling with pre-defined *branded fares* is convenient and easier to comprehend for customers. The profitability of suitable bundling strategies goes back to early studies by Stigler (1963) and Schmalensee (1984).

Further, differential bundling, as opposed to differential pricing, might circumvent problems with anti-discrimination laws (Adams & Yellen, 1976). Ratliff & Gallego (2013) evaluated airline branded fares for four to five distinct customer segments in a discrete choice framework. They reported significant profitability differences across various branded fare design and pricing strategies. To improve comparability across airlines, Szymanski and Darrow (2021) developed a framework categorizing airline-specific branded fares into utility levels.

As a third strategy, dynamic bundling aims to combine convenience and flexibility. It requires methods to detect which specific bundle most closely resembles customer preferences for every individual request. Vinod (2020) proposed a combination of segmentation and personalization to create a "segment of ONE". He drafted a three-step framework to create personas, build recommendation engines, and adjust recommendations in case customer identity is declared. To integrate feedback effects and capture behavior changes, Vinod (2020, 2021a) advocated application of reinforcement learning techniques.

3.2.4 Assortment

Assortment aims to select a subset of all products to display to customers. It originated in the retail industry with the optimization of products on store shelves (Kök et al. 2008). Assortment optimization rests on the observation that consumers deviate from perfect rationality due to several psychological influences including framing, priming, positioning, and defaults (Jannach et al., 2010). Optimal pricing might also depend on which other products are made available for purchase, calling for integrated models (e.g., Ferreira & Wu, 2011).

3.2.5 Distribution

The relevance of traditional distribution through Global Distribution Systems (GDS) limits the ability of airlines to control the entire offer generation process and restricts pricing to 26 booking classes. However, the increased adoption of IATA's New Distribution Capability (NDC, IATA, 2020) and higher shares of direct distribution enable airlines to receive and process more information about every single search request, and finally to respond to each request with customized offers. Moreover, NDC paves the way for effective deployment of continuous pricing, eliminating the concept of fare classes in the long-term. Leading airlines consequently view NDC as important strategic pillar (e.g., Lufthansa, 2018).

3.3 Contribution: customized offer management

Due to high practical relevance, customized offer management has attracted scientific attention in the recent years. Table 4 summarizes relevant literature since 2013.

Table 4: *Summary of offer management literature since 2013.*

Paper	Bundling	Ancillary / Bundle pricing	Flight pricing
Ratliff and Gallego (2013)	Discrete choice	Discrete choice	-
Fiig et al. (2016)	-	-	Discrete choice
Madireddy et al. (2017)	Various machine learning models	Various machine learning models	-
Bockelie and Belobaba (2017)	-	Discrete choice	Discrete choice
Wittman and Belobaba (2018)	-	-	Discrete choice
Fiig et al. (2018)	-	A/B testing and machine learning	Discrete choice
Shukla et al. (2019)	-	Various machine learning models	-
Vinod (2020)	Various machine learning or statistical models	-	-
Srinivasan and Komirishetty (2021)	-	Deep neural networks	-
Kumar (2021)	-	Reinforcement learning	-
Ratliff (2021)	Machine learning	Machine learning	-

Most contributions either focused on bundling with given ancillary prices, ancillary pricing, or flight pricing. Notably, the PODS Research Project (PODS, 2021) develops discrete choice optimization models to integrate ancillary, itinerary and fare class choice. As part of this project, Bockelie & Belobaba (2017) distinguished two types of customers, namely simultaneous consumers selecting flight and ancillaries at the same time, and sequential consumers choosing flights first and evaluating ancillaries afterwards. This is how most flight search engines and the GDS work. Madireddy et al. (2017) studied offer construction for sequential customers, suggesting clustering algorithms for customer segmentation based on trip characteristics and, if available, past customer behavior. Comprehensive OMS were drafted by Vinod et al. (2018) with rule-based trip-purpose segmentation, and Fiig et al. (2018) propagating discrete choice models for flight pricing and ML for ancillary pricing.

In conclusion, some studies advocate discrete choice or generally statistical models, whereas others propose and/or apply different ML algorithms. Compared to problem-specific choice models, ML might be less theoretically sound and less comprehensible; however, ML could be more flexible to find generalizable patterns in unstructured high-dimensional data than assumption-driven choice models that require low-dimensional attributes for meaningful analyses (Bzdok et al., 2018; Fiig et al., 2018). Therefore, ML has various potential applications in travel, including customer segmentation with more segments than would be possible in choice models, upsell management, recommendation engines, and seat-based pricing (Dadoun et al., 2021; Vinod, 2021a, 2021b).

This research develops a conceptual OMS architecture comprising bundling, ancillary pricing, flight pricing, and assortment. It sequentially applies two ML algorithms, clustering and matrix factorization, to support segmentation with much higher granularity than existing research. At the same time, the modular design of the architecture ensures data sparsity, curse of dimensionality and model understandability are adequately addressed. The next section will describe the proposed OMS.

3.4 Offer management system architecture

Both streams introduced in Figure 9, product choice and flight pricing, are composed of various modules. Figure 10 shows the modules and how they interact with existing elements of airline RMS.

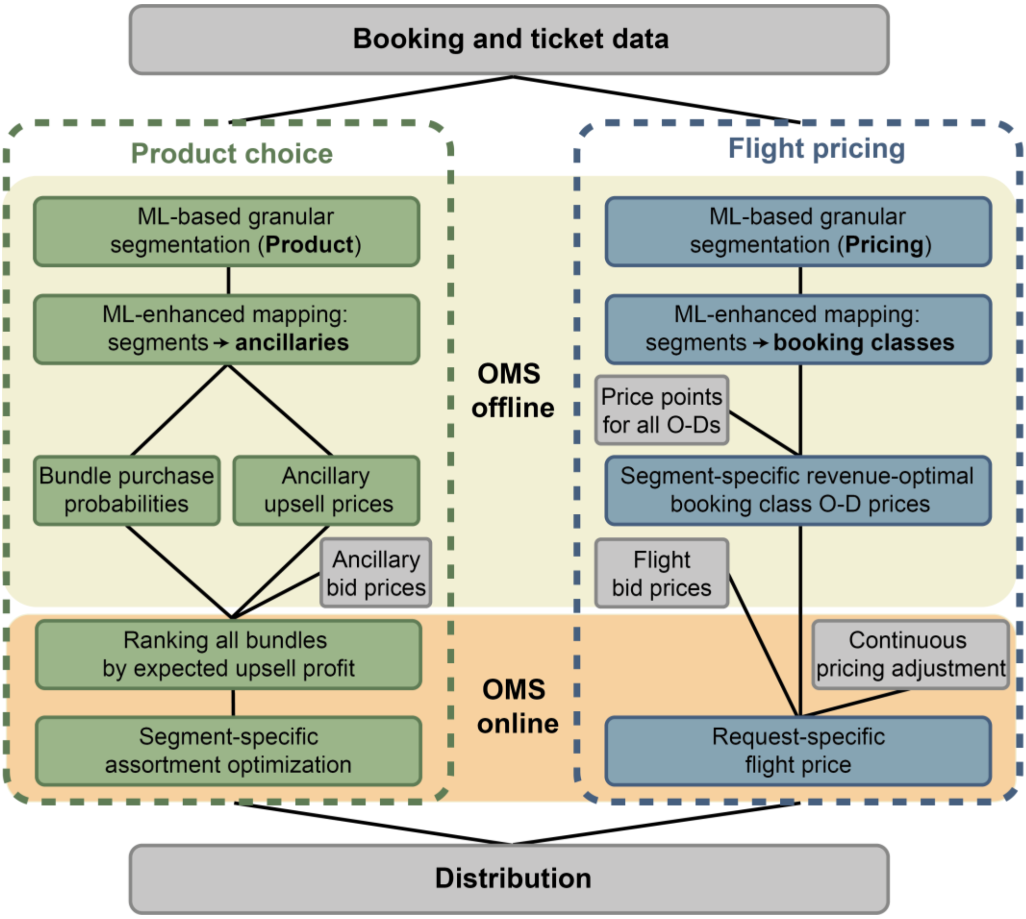

Figure 10: *Schematic overview of the modules contributing to the two streams (vertically), divided into the offline and online parts of the proposed OMS (horizontally).*

Both streams build on historical **booking and ticket data** including real-time request parameters assumed to have predictive power for product choice and prices paid. In all analyses, branded fares are treated as combination of flight plus all ancillaries included though this might raise some problems as addressed in “Practical airline implementation”.

For **product choice**, an ML-based segmentation model granularly clusters all booking requests with respect to ancillary purchases. To avoid curse of dimensionality and to keep the model understandable, segmentation is broken down into subdimensions. Subsequently, the segments are mapped to ancillary purchase probabilities and average ancillary spend, enhanced by ML to overcome data sparsity problems. Then, segment-specific bundle purchase probabilities and ancillary upsell prices are calculated. These modules run in the background with periodical updates (offline). In real-time, every incoming booking request runs through these steps (online). Together with ancillary bid prices, if existent, all possible bundles are ranked by expected upsell profitability. Finally, the assortment optimization module decides in real-time which of all possible bundles shall be displayed at which upsell prices to the particular segment.

Conceptually, the same two ML-driven optimization steps are performed for **flight pricing.** The segmentation model clusters with respect to booking classes purchased, and segments are mapped to nested booking class purchase probabilities. These are combined with price points for all origin-destination (O-D) pairs to determine the request-specific revenue-optimal booking class flight price for all O-Ds. Online, every new request runs through these steps. Together with flight bid prices and an additional module for continuous pricing, if applicable, the request-specific flight price is determined in real-time.

Together, the outputs of both streams govern the offers and corresponding prices to be returned to each incoming request in real-time through the **distribution** channels (direct sales, (online) travel agency, etc.).

The next two sections consecutively detail the modules of both streams, product choice and flight pricing.

3.5 Stream 1: Product choice

This stream refers to the left side of Figure 10 and sequentially details the modules shown from top to bottom. It aims to facilitate segment-specific bundle assortment at segment-specific upsell prices in real-time.

3.5.1 ML-based granular segmentation (Product)

Goal of this module is segmenting all booking requests with respect to product choice, i.e. ancillary purchase. This research proposes unsupervised ML to compactly represent unstructured data by identifying groups of similar behavior to enable much higher granularity than would be feasible for rule-based or choice models.

The different *parameters*, which are attached to each booking request and could predict customer choice, are categorized into C *features*, representing segmentation subdimensions. Figure 11 illustrates a representation with nine features. These fall into one of three categories. One, information that customers explicitly provide when searching for flights on airline websites (origin, destination, departure date, travelers, return dates). Two, information that implicitly comes with the search (weekday and time of day of search, sales channel, loyalty). Three, information related to the product (flight). Four, information airlines already collect and use for pricing or product decisions today (holidays, length of stay). Each feature can assume V_c possible values. This results in exponential scaling of the total number of possible *request combinations* r.

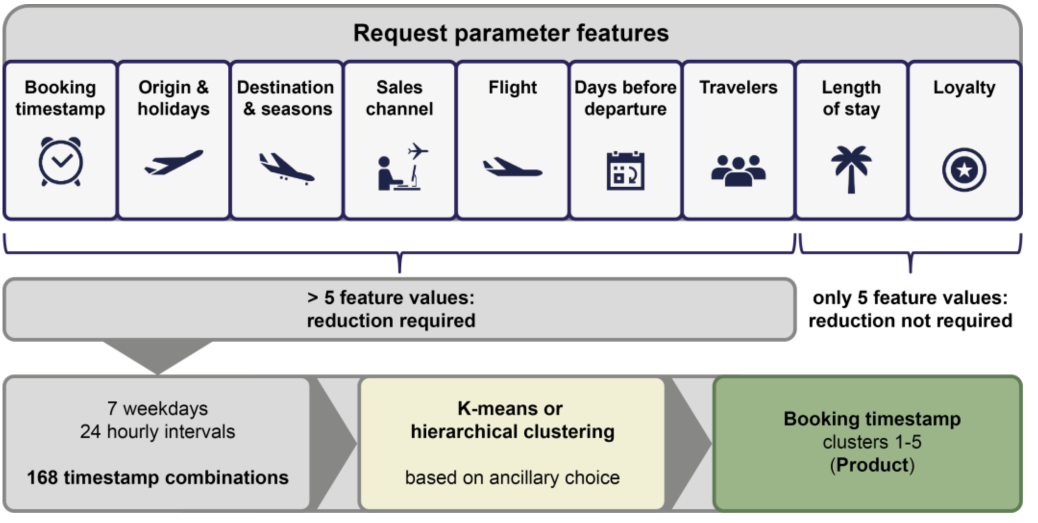

Figure 11: *Illustration of the granular request segmentation based on product choice.*

V_c might be small for some features, e.g. "length of stay", because a request is either made for one-way travel, same-day return, weekend, workweek, or full-week stay. However, V_c could be much larger for other features, such as "booking timestamp". For example, 7 weekdays and 24 hourly intervals result in $V_1 = 168$ timestamp combinations. To bring r to a tractable number of *segments* s, V_1 needs to be reduced to a much smaller number K_1. Two algorithms that seem suitable for this purpose are k-means and hierarchical clustering. K-means clusters requests into K distinct groups (Jain, 2010) based on similarity with respect to ancillary choice. Major benefit is that choosing K is left to the user. With $C = 9$ features and $K = 5$ possible values for each, there are still almost 2 million segments (K^C). Assuming $K = 4$ would already reduce this number to 262,144, which needs to be compared to the total number of airline bookings over the time period feeding the model, realistically 1-3 years. K_c can vary from feature to feature and needs to be optimized during validation on airline data. Advantage of hierarchical clustering is that the user can specify the maximum distance until which objects are clustered together instead of specifying the total number of clusters to form (Murtagh & Contreras, 2012). To capture changes in customer behavior, the segmentation model should be periodically recalibrated (e.g., monthly or quarterly). Also, higher weights can be put on recent observations with exponential smoothing techniques (Gardner 1985).

In real-time, any incoming request is clustered for each feature, i.e. a new request might be mapped to “booking timestamp cluster 1”, “origin & holidays cluster 2”, etc. A segment is the combination of clusters for all features, hence every incoming request gets mapped to precisely one segment.

3.5.2 ML-enhanced mapping: segments to ancillaries

For each segment, observed ancillary purchase probabilities and average ancillary spend form a segment-ancillary interaction matrix. Exponential smoothing can be applied to prioritize recent observations. Due to the highly granular segmentation in comparison to overall bookings, however, data sparsity problems may occur. To overcome, this research propagates a novel application of the well-studied matrix factorization algorithm. The initial idea behind this algorithm is ascribed to Funk (2006) in the context of the Netflix prize competition. So far, matrix factorization has mostly been applied to fill blanks in the user-item interaction matrix, i.e. to predict preferences of identified users for new movies. This research suggests application of matrix factorization on segments instead of individual users and ancillaries instead of movies. Goal is not to fill blanks, but to overcome data sparsity and increase prediction robustness for repeated interaction by updating the observed segment-ancillary matrix through factorization (Figure 12).

Figure 12: *Application of factorization to overcome data sparsity problems by updating the observed segment-ancillary interaction matrix.*

Mathematically, matrix factorization decomposes the relation between the segment vector S and the ancillary vector A into the dot product of the relations between S and a factor vector F, and between the transpose F^T and A (Figure 13).

$$S \cdot A = (S \cdot F) \cdot (F^T \cdot A) \tag{6}$$

The number of factors f to be introduced is found during validation. It needs to be smaller than the number of ancillaries a and significantly smaller than s. Next to overcoming data sparsity, factorization offers the additional advantage of factor interpretation as underlying preferences such as convenience, flexibility, vacation, value of time, comfort, connectivity, or entertainment.

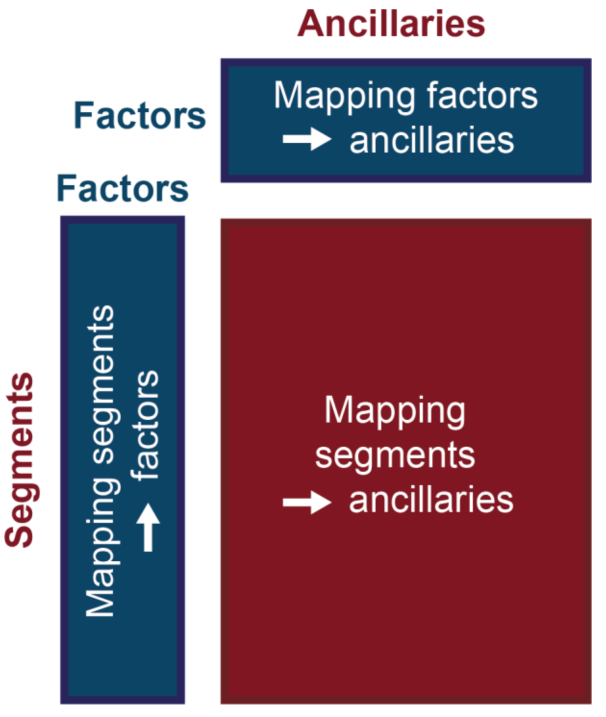

Figure 13: *Matrix factorization decomposes the mapping of segments to ancillaries.*

Starting with arbitrary initial values for the cells of both factor matrices $(S \cdot F)$ and $(F^T \cdot A)$, the matrix factorization algorithm iterates on these to minimize the differences between the originally observed matrix and the one resulting from factorization. Details on optimization techniques for iterative improvement can be found in Aggarwal (2016). The magnitude of the factorization update is inversely related to the number of observations for any particular segment. This is precisely what is intended. Segments with many observations are updated only slightly. Segments with few observations, i.e. those for which data sparsity can be problematic, are updated more heavily. Frequent recalibration of the factorization model and exponential smoothing can ensure capturing changes in customer behavior.

Any incoming request retrieves segment-specific ancillary purchase probabilities and average ancillary spend – updated with factorization – in real-time. These two outputs are independently processed in the next two modules.

3.5.3 Further modules

All possible bundles are all possible combinations of ancillaries. Segment-specific **bundle purchase probabilities** are obtained by multiplication of the segment-specific purchase probabilities of all ancillaries included and counter-probabilities of all ancillaries not included. These calculations are updated periodically.

Due to static ancillary pricing, many airlines lack data on how customers respond to dynamic ancillary prices. To generate such data, this chapter suggests making a very simple yet intuitive assumption to establish segment-specific **ancillary upsell prices.** The more a particular segment spends on ancillaries on average, the higher the ancillary WTP of that segment. Segments that do not spend much on ancillaries need cheap prices to be attracted to ancillary purchases. Segments that do spend much on ancillaries, are willing to pay more for additional services. In addition, airlines might want to specify floors and/or caps to restrict ancillary prices to reasonable ranges. This simplified logic assists first steps towards moving from static to segment-specific ancillary pricing. Once specific data is generated from actual testing, feedback effects can be collected. This would enable application of the same logic as suggested in "Stream 2: Flight pricing" to ancillary pricing as well. In the presented architecture, the optimal upsell price of each bundle equals the sum of the optimal ancillary prices of all ancillaries included. To ensure price consistency, airlines might choose to price any unbundled ancillary consistently as well.

The next module conflates bundle purchase probabilities and ancillary upsell prices to enable **ranking of all possible bundles by expected profitability,** responsive to each incoming request in real-time. In addition, airlines need to ensure that each ancillary at least covers its **ancillary bid prices,** i.e. ancillary opportunity costs on all flight legs traversed. These present an ongoing field of innovation and might not yet be available to most airlines.

After ranking, the **segment-specific assortment optimization** module decides in real-time which of all possible bundles are displayed to an incoming request at which upsell prices. Goal is to consider customers' purchase behavior driven by the psychological factors mentioned in "Assortment". Airlines can display some customized bundle options, e.g. the one with highest expected profitability, or the one with highest purchase probability, or a "premium" bundle including many ancillaries. In addition, airlines might want to display (some of) their established branded fares or the cheapest "no-frill" option to attract customers comparing airline offers for the cheapest price.

3.6 Stream 2: Flight pricing

This stream relates to the right side of Figure 10 and sequentially details the modules shown from top to bottom. Goal is real-time request-specific flight pricing.

Methodologically, this research proposes the same **ML-based request segmentation** as described in "Stream 1: Product choice" (Figure 11). However, requests are clustered with respect to booking class purchased. Notably, continuous pricing is often implemented through discounting and/or incrementing booking class prices to reach the continuous price. Such data-driven ML-based segmentation presents a major deviation from most airlines' practices with rule-based models. Whilst the clustering algorithm shall be updated periodically, every incoming request is mapped to precisely one segment in real-time.

In the next module, **segments are mapped to nested booking class purchase probabilities.** This research proposes updating observed booking class purchase probabilities with matrix factorization to increase prediction robustness. This resembles the methodology introduced in "Stream 1: Product choice" (Figure 12 and Figure 13). This update is performed periodically. For every incoming request, nested booking class purchase probabilities are retrieved from the matrix in real-time.

The subsequent modules of the stream process information from existing airline RMS elements. The **segment-specific revenue-optimal booking class flight price for all O-D pairs** is periodically calculated from segment-specific nested booking class purchase probabilities and pre-defined booking class **price points for all O-Ds**. For any new request, the corresponding booking class flight price is retrieved in real-time. If no continuous pricing adjustment exists with an airline (see below), nested probabilities could be interpolated between booking classes.

Flight bid prices represent seat opportunity costs on all flight legs traversed in a requested O-D. These are combined with the revenue-optimal booking class flight price and, if applicable, any **continuous pricing adjustment**, to obtain the optimal continuous **request-specific flight price** in real-time. The exact strategy to incorporate continuous pricing might vary from airline to airline, but generally follows the logic to adjust booking class prices to ensure consistency with existing RM practices.

As final step, the outputs of both streams determine the segment-specific set of offers and their corresponding request-specific prices, returned via the **distribution** channels to each incoming request in real-time. Bundle prices equal the sum of request-specific flight price and segment- and bundle-specific upsell price.

3.7 Innovations

The presented OMS architecture extends existing research and innovates airline RMS/OMS in various directions.

It combines real-time bundling, ancillary pricing, flight pricing, and assortment into a conceptual solution aimed at high practicality. First, the modularity allows incremental and independent embedding into existing RMS and processes data that already exist. Controllable modules conceivably support understandability and circumvent curse of dimensionality issues, both of which might occur with more complex deep learning models. Second, frequent recalibrations and exponential smoothing aim at timely adaptation to behavior changes. Feedback effects from actual live tests could further foster continuous model learning. Third, the separation of segmentation for product choice and flight pricing allows ancillary and flight WTP to behave differently. Still, correlations seem implicitly captured as both rely on the same data.

Most importantly, ML – as opposed to choice models – might enable data-driven highly granular and high-dimensional segmentation. Potentially resulting data sparsity problems could be addressed with the novel application of matrix factorization. Instead of relying on identified users, it focuses on repeated interaction between segments and ancillaries or booking classes to improve prediction robustness through learning over time.

Both streams are designed to independently add value to airlines. “Stream 1: Product choice” could combine the flexibility of unbundled ancillaries with the simplicity and convenience of branded fares. Whilst the bundling logic could already be implemented with static ancillary pricing, this research offers a simple intuitive logic as first steps towards customized ancillary pricing. “Stream 2: Flight pricing” aims to facilitate revenue optimization through customer segmentation and price-elasticity estimation with much higher granularity than existing approaches. The models can be trained with existing systems based on booking classes but enable integration of continuous pricing logic.

3.8 Validation and dissemination

The conceptual OMS architecture outlined in this chapter can be followed by several validation steps. First, a prediction model can be trained, validated and tested on actual airline data. Second, small-scope real-live testing could be attained through application on small route subsamples. This would enable incorporation of feedback effects and deployment of reinforcement learning techniques to enhance adaptability to possible behavior changes. Third, successful small-scope testing could result in gradual embedding into existing airline RMS/OMS, either of whole streams or selected modules only, allowing comparison to existing models. Fourth, application outside of passenger aviation, perhaps slightly adapted to the specific case, can yield valuable insights into the generalizability of the proposed architecture. Potential use cases include high-speed railways, ferry operators, or air cargo carriers.

3.9 Discussion for practical implementation

This section sketches points of discussion for practical airline implementation. It examines the impact of implicit OMS assumptions and proposes solutions. In addition, it outlines potential extensions of the proposed OMS and touches upon change management aspects.

The presented OMS makes several **implicit assumptions.** Practitioners should be aware of their impacts and how these could be addressed. First, “Stream 2: Flight pricing” assumes customers always choose the lowest offered booking class. This only holds true if product choice and booking class are independent, i.e. existing branded fares and ancillaries are offered in all booking classes. If not, an explicit correction might be required. Second, the OMS treats branded fares as if they represented independent customer choices for all ancillaries included. However, customers might not have purchased all branded fare components if offered unbundled, leading to inflated purchase probabilities of those ancillaries included in branded fares. The issue could be addressed with actual live tests. Third, different bundles might be displayed when otherwise identical requests are repeated with different time stamps. Airlines could think about an additional module able to store the requests and their respective returned bundles for some time. Alternatively, one might argue that airline customers are already used to dynamic flight pricing (and even some first attempts of dynamic ancillary pricing), so they might also familiarize themselves with being offered different products at different points in time. Fourth, in cases of new feature values (e.g., new destinations) or new ancillaries, no historical data exist for ML-based segmentation. Potential solution could be to manually cluster new destinations/ancillaries temporarily with existing ones until sufficient data is generated.

The modular OMS architecture allows **extension with various modules.** First, the OMS could process requests for which the no-purchase option was chosen in addition to those successfully converted into bookings. Second, third party ancillaries (hotels, rental cars, etc.) could be included, requiring interfaces to third party inventory and pricing systems. Third, some existing research (Bockelie & Belobaba, 2017; Fiig et al., 2016; PODS, 2021) explicitly model competition. The proposed OMS could be extended to include competitive pricing and offers. So far, it models competitive effects only indirectly as reflected in the booking data, which might work as long as competition intensity is constant for a given segment and/or O-D. Fourth, many leading airlines investigate into personalization although often the majority of requests remain anonymous. Clear focus of the proposed OMS is to segment requests independent of whether customer identity is declared. This also circumvents the problem of the same individual assuming different roles. However, one could think of enhancing the OMS with an additional personalization module similar to Vinod (2020) suggesting an adjustment of recommended offers for declared customers. Fifth, the ML-segmentation could be combined with discrete choice and/or econometric models to benefit from both the flexibility and adaptability of ML as well as the theoretical soundness of statistics.

Both streams could be independently integrated into airline RMS. Whereas “Stream 1: Product choice” would constitute novel elements to complement existing RMS for most airlines, “Stream 2: Flight pricing” would present a significant change from established practices. Both would require thorough **change management** initiatives within the RM organization including close coordination with Sales, Marketing, and Distribution departments. The same holds true to raise acceptance for ML adoption. The more understandable ML models are, the higher their acceptance. The proposed architecture appears more intuitive than deep learning models or artificial neural networks.

3.10 Conclusion

This research combines request-specific real-time bundling, ancillary pricing, flight pricing, and assortment into a conceptual architecture for customized airline offer management. To facilitate higher granularity than existing models, the proposed offer management system (OMS) leaves request segmentation entirely up to machine learning supported pattern recognition in booking data instead of applying customer choice models. The design of the OMS addresses three potential problems that might result from such segmentation with high granularity. First, a novel application of the well-studied matrix factorization algorithm could overcome data sparsity problems. Second, segmenting requests alongside subdimensions might circumvent curse of dimensionality issues. Third, the modularity of the architecture supports model understandability. Hence, the OMS might enable a new solution to the fundamental trade-off between granularity/complexity and model applicability/robustness.

The OMS is structured alongside two streams, both of which could be independently integrated into airline revenue management systems (RMS). "Stream 1: Product choice" would support the evolution of RMS into OMS to enhance customer satisfaction being presented more relevant offers. "Stream 2: Flight pricing" could facilitate willingness to pay estimation for highly granular segments to potentially improve pricing decisions and increase airline revenue.

Further academic research could either contribute to validation of the OMS or investigate into potential practical extensions as indicated.

4 Inductive research: first validation[6]

Airlines serve different customers. Developing customized offer and pricing strategies have become a strategic priority to airlines and have attracted operations research accordingly. This chapter tests the first fundamental hypothesis of the offer management system architecture of Chapter 3 that aims to combine the simplicity of branded fares with the flexibility of unbundled ancillaries. The hypothesis is that the wealth of data available to airlines enables segmentation orders of magnitude more granular than existing models based on anonymous features of a particular customer search.

The hypothesis is tested through inductive research based on 202 million flight coupons of a major network airline between 2018 and 2023. Chapter 4 observes customer choice probabilities differ depending on the weekday of the booking. Next, it identifies a pattern. Customer choice probabilities also differ based on sales channel, flight characteristics, and customer loyalty status. Generalizing these findings, this research suggests airlines can improve prediction accuracy of customer choices and hence display more relevant offers when segmenting on these search features. If true, it presents a low-cost process to increase prediction accuracy with data readily available to airlines. The findings are relevant to both airline practitioners and researchers for follow-up studies alike.

Section 4.1 briefly summarizes literature context, motivation, and contribution. Section 4.2 describes the research approach and provides details on the 496 million airline coupons used. Section 4.3 discusses a first observation, namely that the weekday of a booking affects customer choice. Building on that, Section 4.4 identifies a pattern. Not only booking weekday, but multiple features included in customer searches affect customer choice probabilities. Section 4.5

[6] This chapter is based on a paper that is written and approved by the airline partner for publication.

summarizes and discusses results. Section 4.6 concludes the validation of the first fundamental hypothesis of the dissertation.

4.1 Literature context, motivation, and contribution

In the last two decades, the evolution from revenue management (RM) to offer management (OM) and customization of offers have been two key themes for airlines. First, the rise of low-cost carriers has increased the importance of ancillary revenues. This has fueled the shift from RM aiming to optimize pricing to OM aiming to jointly optimize pricing and product decisions. In this notion, an offer is defined as the product plus the price. Second, advancements in pattern recognition through machine learning (ML) and online distribution enabled customization, or even personalization, of offers, with e-commerce spearheading innovation (Bakos, 2001; Kashyap et al., 2022). Airlines aim for similar personalized or customized pricing and product strategies as well.

Due to their practical relevance, both these themes have received ample attention in passenger air transport from operations research perspective. The same holds true for customer segmentation as key enabler.

4.1.1 Customers are different and want different things

Different customers demand different things. Firstly, because customers themselves are different from each other. These differences manifest themselves in their search behavior. Examples are customers searching on different weekdays or through different sales channels. Also, customers have different loyalty behavior and status. In addition, some customers shop for flights one year before departure, whereas others search flights that depart on the same day.

Secondly, because customers search for different products. They depart from different origins and want to go to different destinations. They look for different flight compartments, from Economy to First. Some customers travel alone, whereas others travel with a partner, with a family, with infants, or with a group. Customers might search for one-way options or return flights with length of stay varying between few hours and multiple months. Length of stay is among the oldest customer segmentation fences used in flight pricing (Belobaba, 2016). It aims to segment by travel purpose, assuming leisure travelers stay longer at their destination than business travelers.

4.1.2 Airlines differentiate their offers

Airlines aim to serve their differentiated customers with differentiated offers. To study differentiated OM, researchers and practitioners distinguish between personalization and customization. Personalization aims for offers catering to the individual person searching. This typically requires the customer to reveal themselves, e.g. by logging into the airline website or loyalty program. In this case, airlines can consider the past shopping history of the individual customer. On the contrary, customization does not require knowledge of the personal identify. Instead of past shopping behavior of this individual, it aims to identify patterns in customer searches. Goal is to respond with customized offers based on what customers with similar search behavior purchased in the past. Advantage of customization over personalization is that it is not in conflict with data privacy rules (Millet 2023). Also, it is more universally applicable precisely because it also works when customers do not declare their personal identify.

Airlines have multiple levers to differentiate, i.e. personalize or customize, their OM. Flight pricing is the oldest lever in traditional airline RM. Starting with Littlewood (1972) suggesting airlines to maximize revenue instead of load factors, airlines have gained decades of experience. Transitioning from RM to OM, product choice has become more relevant. It comprises which ancillaries customers purchase at which price on top of the mere seat. Within product choice, the differentiation levers can be expanded to bundling, ancillary pricing, and assortment as visualized in Figure 14.

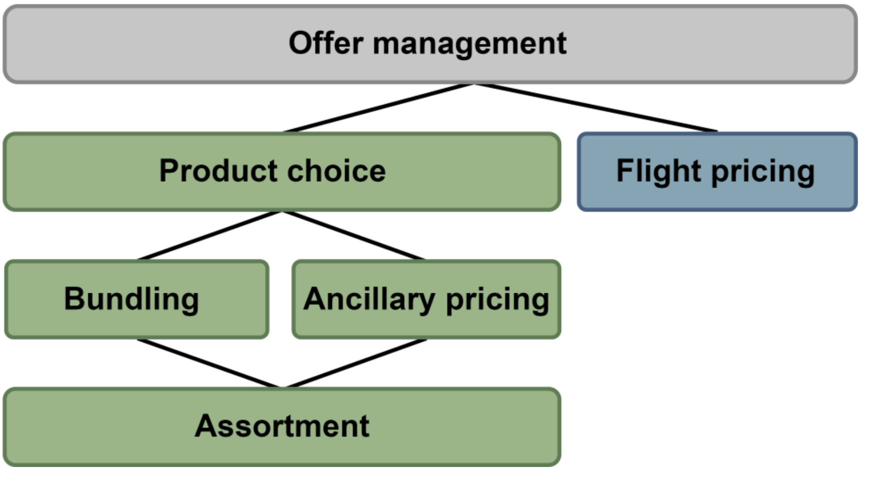

Figure 14: *Different levers to differentiate offer management.*[7]

Bundling refers to combining single ancillaries into unbreakable packages and was first academically researched by Stigler (1963). Bundling strategies can also result in some ancillaries or combinations of ancillaries being only available as part of bundles. Ancillary pricing comprises pricing of single ancillaries as well as bundles. In 2023, ancillary revenues were estimated at 15% of total airline revenues (IdeaWorks & CarTrawler, 2023). Finally, assortment refers to selecting which of the possible combinations of products to show to customers and in which order. It builds on several psychological behaviors that cause consumers to deviate from rational behavior (Jannach et al., 2010).

[7] Source: Schubert et al. (2021).

4.1.3 Models to optimize offers for a given search

Both academia and practitioners have applied different classes of models to use these levers in a way that maximizes customer or business outcomes for a given customer search. Customer outcomes can be offer relevance and satisfaction. Business outcomes can be conversion, revenue, seat load factor, and profitability. Broadly, the models can be categorized into statistical models like discrete choice analysis, and model-free ML algorithms. Established airline OM examples are Ratliff & Gallego (2013) and Wittman & Belobaba (2019) for discrete choice models, as well as Madireddy et al. (2017) and Shukla et al. (2019) for ML algorithms. Discrete choice models offer the advantage of theoretical soundness, whereas ML is more flexible to generalize patterns from unstructured and high-dimensional data (Bzdok et al., 2018; Fiig et al., 2018).

Chapter 3 presented a comprehensive OM architecture that aims to improve prediction accuracy of discrete choice models through application of different ML algorithms. The comprehensive architecture is visualized in Figure 15. To improve model understandability and applicability, they proposed a modular OM system, which is structured into two main streams for product choice and flight pricing. Both streams could be independently embedded into existing airline systems.

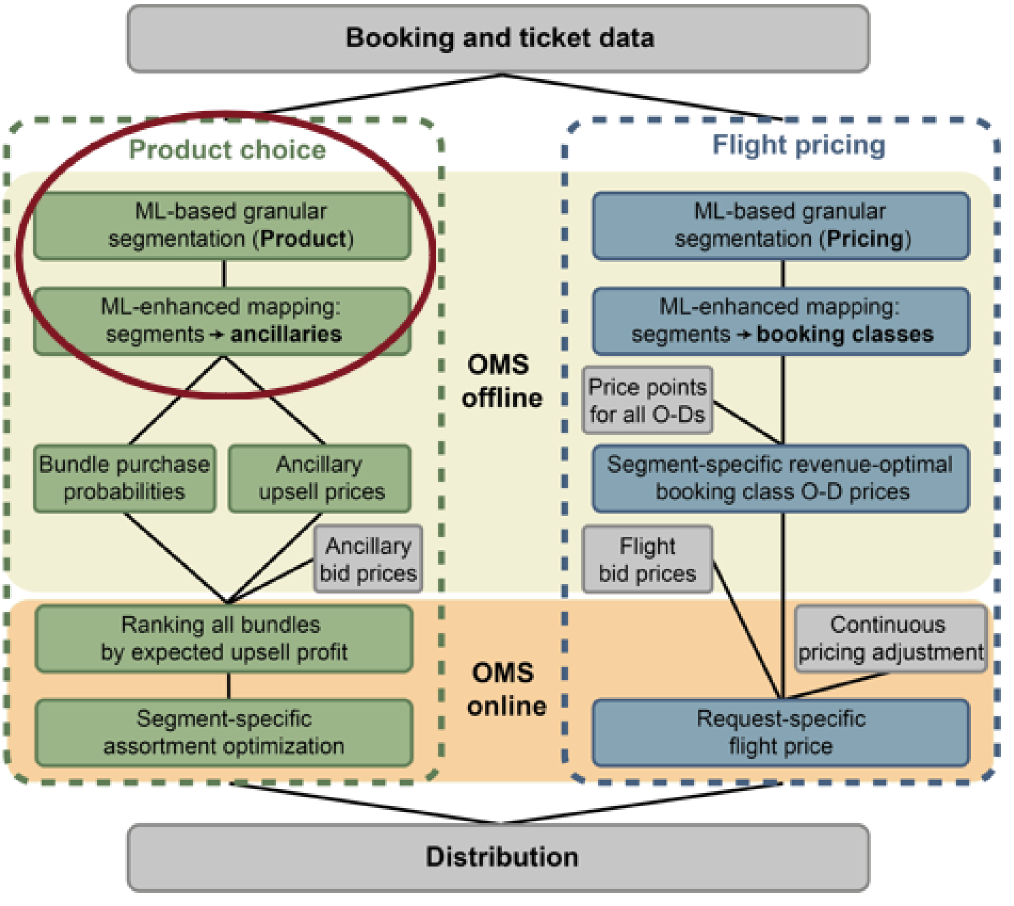

Figure 15: *Conceptual architecture put forward in this dissertation.*

The ability to improve prediction accuracy of customer discrete choice probabilities rests on the feasibility of highly granular segmentation to achieve a new balance between granularity/complexity and applicability/robustness of the segments. This would enable airlines to respond to customer searches with context-specific offers to increase outcomes for customers and business. It would combine the flexibility of unbundled ancillaries with the convenience and simplicity of branded fares. To resolve data sparsity, they conceptualized a novel application of the well-studied matrix factorization algorithm first introduced by Funk (2006) for the Netflix price competition. Matrix factorization belongs to the family of recommender system algorithms that are widely used to tailor offerings to individual customers or customer segments. Prominent use cases besides Netflix are Amazon, Spotify, and many more (Alamdari et al., 2020).

The viability of the proposed OM architecture rests on the following fundamental hypothesis:

Through a combination of ML algorithms, airlines can achieve segmentation orders of magnitude more granular and high-dimensional than previous models because of the wealth of data available to airlines included in customers searches.

Testing this hypothesis on the product choice levers requires zooming in on the red oval in Figure 15. Goal is to map search characteristics, representing customer segments, to products, thereby detecting segment-specific customer choice probabilities.

4.1.4 Research contribution

This chapter conducts inductive research on this hypothesis. It studies 202 million coupons of actual airline booking data from a major network airline, identifies patterns, and confirms the hypothesis. As such, it is the first proof-of-concept of the viability of the solution presented in Chapter 3. Customer choice behavior is investigated based on three customer and one product characteristics. The three customer characteristics are booking weekday, sales channel, and loyalty status. The product characteristics is whether a searched flight leg is an overnight flight or not. The scope of the research is limited to Economy searches and bookings.

The inductive research conducted is independent of personalized data, hence compliant with data privacy rules such as GDPR in Europe[8]. Instead of

[8] GDPR (General Data Protection Regulation) is an EU regulation governing information privacy in the European Union since 2018. Similar regulations exist in other parts of the world.

personalized OM, the research investigates the proposed solution for highly granular customized OM. Customer segments are clustered without using sensitive personal data, and the context of a search is used to connect a specific customer to one of the segments. The presented research is novel and innovative as the first to analyze the potential of highly granular segmentation with such comprehensive actual airline data.

4.1.5 Structure of the chapter

The remainder of the chapter is structured as follows: Section 4.2 describes the airline data and approach underlying this inductive research. Section 4.3 investigates customer choice probabilities based on the weekday their bookings were placed. It also details the methodology and metrics to quantify the differences between different choice probability distributions. Section 4.4 expands the analysis to customer choice probabilities based on sales channel, flight characteristics, and customer loyalty status. Section 4.5 generalizes the inductive findings. Section 4.6 summarizes and suggests how future research can validate these.

4.2 Research approach and data

This inductive research analyzes actual customer choice behavior of a major network airline in a period of five years between September 2018 and September 2023. Customer choice probabilities are investigated based on booking weekday, sales channel, flight characteristics, and customer loyalty status.

The customer choice process underlying the data follows the two-step process visualized in Figure 16. This chapter comprises the airline's customer journey with four branded fares (A, B, C, D) and five paid ancillaries (I, II, III, IV, V). The branded fare (BF) and ancillary names are anonymized to protect data confidentiality of the network airline.

This research focuses on airline-own ancillaries, i.e. excludes those offered by third parties such as hotels and rental cars. Some more expensive branded fares include some of the ancillaries that are available for *a la carte* purchase for cheaper branded fares (e.g., seat reservation and additional check-in baggage). Other value-added services (e.g., rebooking and cancellation options) are only available in more expensive branded fares, but cannot be purchased as *a la carte* ancillaries. Finally, not all combinatorial possibilities can be combined in practice. All possible combinations span a total of 100 products that customers can select from. A product is defined as a distinct combination of BF selection and zero, one or more paid ancillaries.

Figure 16: *Customer choice process, selecting branded fares in step 1 and amongst a list of a la carte ancillaries in step 2.*

The research is conducted based on 496 million coupons of a major network airline between September 2018 and September 2023. The data used is confidential information of the major network airline. The data is anonymized to comply with data privacy rules. After removing all compartments but Economy (i.e., removing First, Business, Premium Economy), 359 million coupons remained. To ensure data quality, additional cleansing removed coupons that did not have a branded fare associated, a create date, a coupon leg number, an operating airline, an origin airport, a destination airport, and/or any ticketed gross price information. After these cleansing steps, 202 million coupons remained. A coupon refers to one passenger and one flight leg. A booking can comprise multiple coupons. Geographically, the data spans origins and destinations worldwide. The hub structure of the airline is reflected in geographical concentration on its home markets.

Table 5 shows customer choice behavior with respect to BF and ancillary selection. The vast majority (89%) of all bookings did not purchase any paid ancillary. This is especially true for BF A, C and D. Only customers purchasing BF B exhibit a somewhat higher probability of paying for additional ancillaries. Also, there is strong concentration on ancillaries I and II. Any combination involving paid ancillaries III, IV or V was purchased by less than 0.5% of customers for all four branded fares.

Table 5: *Exploratory data analysis of 202 million coupons, showing the observed choice frequencies of customers (in %) based on their branded fare selection.*

Branded fare (BF)	**No paid ancillary**	**Paid ancillary I**	**Paid ancillary II**	**Paid ancillaries I and II**	**Other paid ancillaries**
BF A	97%	<0.5%	3%	<0.5%	<0.5%
BF B	81%	11%	6%	2%	<0.5%
BF C	94%	<0.5%	6%	<0.5%	<0.5%
BF D	94%	<0.5%	6%	<0.5%	<0.5%
Total	89%	6%	4%	1%	<0.5%

The analysis in this chapter covers a subset of the features presented in the conceptual architecture of Chapter 3 (Figure 17). Booking timestamp, sales channel, flight, and loyalty were selected as features for the inductive validation for two reasons. First, they are data readily available for airlines. Second, they are not used for segmentation, pricing, or product decisions by most airlines, unlike origin, destination, days before departure or length of stay. Additionally, these features comply with anti-discrimination laws.

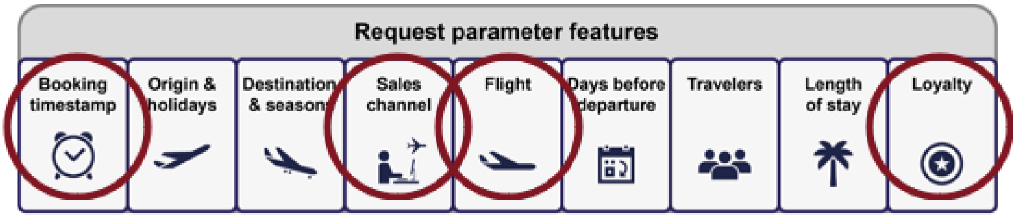

Figure 17: *Four features selected for validation of the conceptual solution.*

Figure 18 and Figure 19 visualize the distribution of booking weekdays and sales channels. For example, 8% of all bookings were made on a Saturday, which makes Saturday the day of the week with the lowest customer booking activity. These 8% on a Saturday compare to 17% of all bookings placed on Monday and another 17% on Tuesday, which are the two busiest days of the week with respect to customers placing their bookings. For sales channels, the real names of the channels are not included due to data confidentiality. However, the chart still shows interesting trends. For instance, 41% of all bookings are placed through one channel ("Channel 4"), which is by far the highest share. In contrast, five channels are used for less than 5% of all bookings.

Due to data confidentiality, further descriptive statistics, such as the distributions by flight type and loyalty status, are not included in the thesis. The same applies to descriptive statistics on other potential features, such as origin or destination countries or the number and type of travelers (adults, children, infants).

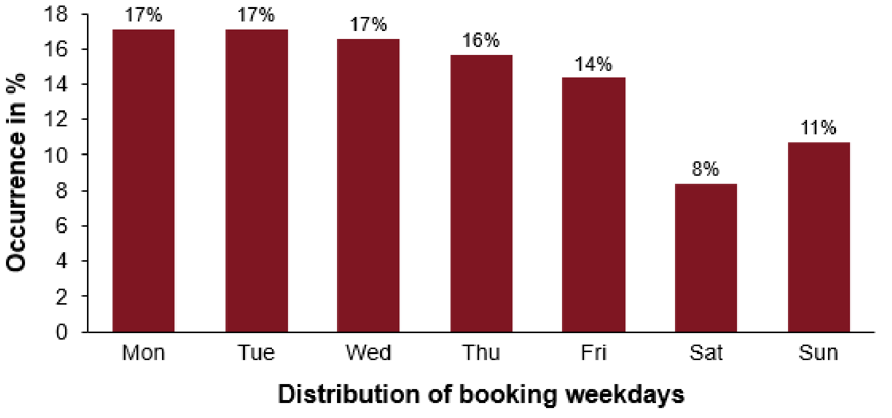

Figure 18: *Distribution of booking weekdays in the entire dataset used.*

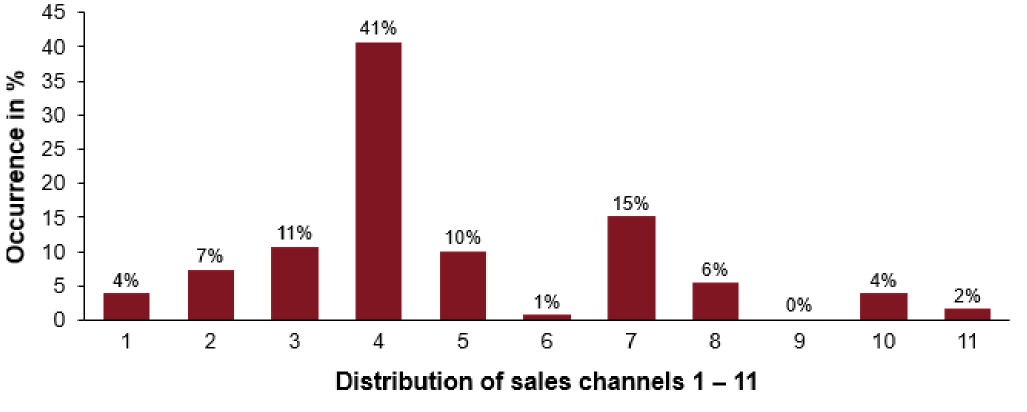

Figure 19: *Distribution of sales channels in the entire dataset used.*

In the next sections, first observations will be presented on different choice probabilities based on booking weekday. Thereafter, the three other characteristics sales channel, overnight flight yes/no, and customer loyalty status are sequentially analyzed, thereby generalizing to a pattern and building increasingly more confidence.

Figure 20 visualizes the sequence of the analyses.

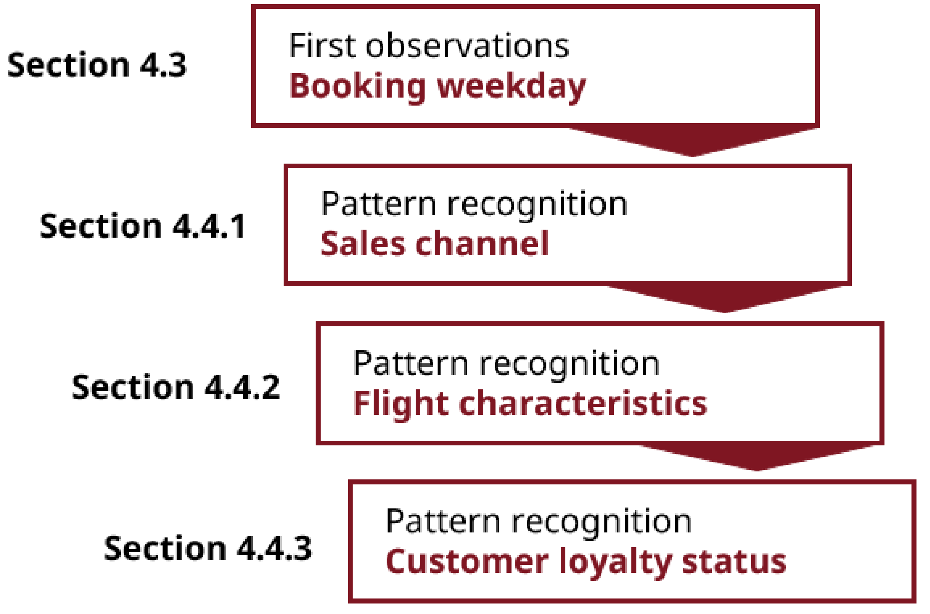

Figure 20: *Sequence of the analyses in this chapter.*

4.3 First observation

In this section, customer choice probabilities are analyzed based on booking weekday. The goal is to identify by how much customer choice probability distributions differ based on the weekday the booking was made.

Table 6 shows customer choice behavior depending on booking weekday. It shows observed customer choice probabilities for the four branded fares without paid ancillaries, and the four most often selected products that include paid ancillaries. Most notable is a change in customer behavior between weekdays (Monday-Friday) and weekend days (Saturday-Sunday). Although further details are not included in the thesis due to data confidentiality, this pattern may be explained by the different travel purposes of customers. Leisure travelers tend to book flights on all days of the week, whereas business travelers rarely do so on weekends. On weekdays, the most often selected product is BF A without paid ancillaries, whereas it is BF B without paid ancillaries on weekends. The weekdays vs. weekend days difference seems to come predominantly from BF A vs. BF B, both without ancillaries. The choice probabilities of other products seem less dependent of the weekday of the booking. An exception is the choice probability for BF D without paid ancillaries, which is more than double on weekdays compared to weekend days.

Table 6: *Observed product choice probabilities (in %) for different booking weekdays.*

Week-day	BF A + no anc.	BF B + no anc.	BF C + no anc.	BF D + no anc.	BF B + anc. I	BF A + anc. III	BF B + anc. III	BF B + anc. II	Other
Mon	46%	36%	0.2%	3.3%	5.0%	2.7%	1.1%	2.4%	2.9%
Tue	46%	37%	0.2%	3.4%	5.1%	2.6%	1.0%	2.4%	2.9%
Wed	45%	37%	0.2%	3.3%	5.1%	2.5%	1.0%	2.4%	3.1%
Thu	46%	37%	0.2%	3.4%	5.1%	2.5%	1.0%	2.4%	3.0%
Fri	46%	37%	0.2%	3.4%	5.1%	2.5%	1.0%	2.4%	3.0%
Sat	37%	46%	0.2%	1.4%	7.3%	1.2%	1.0%	3.2%	2.9%
Sun	37%	45%	0.2%	1.4%	7.4%	1.3%	1.1%	3.3%	2.9%

Note: Darker blue indicates higher product choice probabilities. Lighter gray indicates lower product choice probabilities. BF indicates branded fares. Anc. indicates ancillaries.

Next goal is to quantify differences in observed product choice probabilities between different weekdays. To establish the relationship between two probability distribution, a statistical distance metric is needed, termed **divergence** in information theory. Different divergence metrics exist (for a review, see Lin, 1991, and Taneja et al., 1989). Also, the metric definitions have evolved over time.

This chapter uses three divergence metrics. First, the Kullback-Leibler divergence (KL divergence) first introduced by Kullback & Leibler (1951).

Second, the Symmetrized divergence first used by Jeffreys (1948). Third, the Jensen-Shannon divergence (JS divergence) as symmetric and bound extension of the KL divergence with intuitive interpretability (Lin, 1991). Definitions of the three metrics will follow.

The KL divergence D_{KL} for discrete probability distributions P and Q on the sample space X is defined according to Equation (7) (MacKay, 2003):

$$D_{KL}\,(P\,||\,Q) = \sum_{x=1}^{X} \mathrm{P(x)}\; log\left(\frac{P(x)}{Q(x)}\right) \qquad (7)$$

D_{KL} is non-negative and can be interpreted as the information gain when using P instead of the currently used distribution Q. In the language of Bayesian inference, D_{KL} is the amount of information lost when P is approximated by Q.

With these properties, KL divergence is asymmetric as shown in Equation (8):

$$D_{KL}\,(P\,||\,Q) \neq D_{KL}\,(Q\,||\,P) \qquad (8)$$

This means the amount of information lost when approximating customer purchase behavior on Mondays with the customer choice probabilities observed on Tuesdays is different from the amount of information lost when approximating purchase behavior on Tuesdays with choice probabilities observed on Mondays. For that reason, Kullback & Leibler (1951) calculated the symmetrized divergence D_S, which was already introduced by Jeffreys (1948), as symmetrized and non-negative function according to Equation (9):

$$D_S = D_{KL}\,(P\,||\,Q) + D_{KL}\,(Q\,||\,P) \qquad (9)$$

Table 7 visualizes D_S between the different booking weekdays with the logarithm in Equation (1) calculated to the base two. In addition, it includes customer choice probabilities over all weekdays. Table 7 confirms the observations made in Table 6. The symmetrized divergences between weekdays (Mon-Fri) and weekend days (Sat-Sun) are 0.10-0.12. In contrast, all symmetrized divergences between one weekday and another weekday, as well as between Saturday and Sunday, are practically zero. In conclusion, customer purchase behavior seems different between weekdays and weekend days, but not within either weekdays or weekend days. As can be seen from symmetrized divergences with the probability distribution over all weekdays, weekdays dominate the entire population. This has the practical implication that hardly any information is lost when not using weekday-specific probability distributions on weekdays, whereas there is some information lost when doing the same on weekend days.

Table 7: *Symmetrized divergence between customer choice probability distributions on different booking weekdays, and between the different weekdays and the entire population (all weekdays).*

Weekday	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**	**Sun**	**All**
Mon	0.00							
Tue	0.00	0.00						
Wed	0.00	0.00	0.00					
Thu	0.00	0.00	0.00	0.00				
Fri	0.00	0.00	0.00	0.00	0.00			
Sat	0.12	0.11	0.11	0.11	0.11	0.00		
Sun	0.11	0.11	0.10	0.11	0.11	0.00	0.00	
All	0.01	0.00	0.00	0.00	0.00	0.08	0.07	0.00

Note: Darker blue indicates higher symmetrized divergences. Lighter gray indicates lower symmetrized divergences.

The symmetrized divergence is symmetric but lacks easy interpretability. It is not obvious whether a value of 0.12 is high or low. The JS divergence offers an intuitive interpretation as it is bound between zero and one if the logarithm for calculating the KL divergences is to the base two. Zero means the two probability distributions are identical, and one means the two probability distributions are maximally different, i.e. the two distributions have disjoint support (Lin, 1991). The JS divergence is calculated according to Equation (10):

$$D_{JS} = \frac{1}{2} D_{KL}\ (P \;||\; M) + \frac{1}{2} D_{KL}\ (Q \;||\; M) \qquad (10)$$

M is the average of the two distributions P and Q according to Equation (11):

$$M = \frac{1}{2}\,(P + Q) \tag{11}$$

Table 8 visualizes D_{JS} between the different booking weekdays with the logarithm in Equation (7) calculated to the base two.

Table 8: *Jensen-Shannon divergence (in %) between customer choice probability distributions on different booking weekdays, and between the different weekdays and the entire population (all weekdays).*

Week-day	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**	**Sun**	**All**
Mon	0.0%							
Tue	0.0%	0.0%						
Wed	0.0%	0.0%	0.0%					
Thu	0.0%	0.0%	0.0%	0.0%				
Fri	0.0%	0.0%	0.0%	0.0%	0.0%			
Sat	1.5%	1.4%	1.3%	1.4%	1.4%	0.0%		
Sun	1.4%	1.3%	1.2%	1.3%	1.3%	0.0%	0.0%	
All	0.1%	0.0%	0.0%	0.0%	0.0%	1.0%	0.9%	0.0%

Note: Darker blue indicates higher Jensen-Shannon divergences. Lighter gray indicates lower Jensen-Shannon divergences.

Using the JS divergence instead of the symmetrized divergence does not change the ordinality within the table. The JS divergence, however, shows that probability distributions between weekdays and weekend days are different, but they are much closer to being the same than to being disjoint.

This section has observed that customers booking on different weekdays have different choice probabilities, at least those booking Monday-Friday compared to those booking Saturday and Sunday. Next follows an investigation whether there is a pattern, i.e. whether customer choice probabilities depend on multiple features of their search, of which booking weekday is just one example.

4.4 More observations and pattern recognition

Having shown customer choice probabilities differ based on the weekday of their booking, this section expands the analysis to multiple features: sales channel (4.4.1), flight characteristics (4.4.2), and customer loyalty status (4.4.3). The analysis focuses on the JS divergence due to its intuitive interpretability established in the previous section.

4.4.1 Sales channels

Airlines distribute their offers through different sales channels. These can be grouped into airline-own (e.g., own website, own offices) vs. third parties (e.g., third party websites or comparison portals like skyscanner.com). Another differentiation can be made between online vs. offline distribution. The network airline analyzed distinguishes between eleven sales channels.

The analysis starts with observed customer choice probabilities for bookings placed through different sales channels.

Table 9 shows these for the four branded fares without paid ancillaries, and the four most often selected products that include paid ancillaries. This is the same choice set as analyzed for booking weekdays before, though now separating choice probabilities based on eleven different sales channels. The channel names are anonymized to protect data confidentiality of the network airline.

Table 9: *Observed product choice probabilities (in %) for different sales channels.*

Sales ch.	**BF A + no anc.**	**BF B + no anc.**	**BF C + no anc.**	**BF D + no anc.**	**BF B + anc. I**	**BF A + anc. III**	**BF B + anc. III**	**BF B + anc. II**	**Other**
1	58%	32%	0.2%	1.6%	4.6%	0.9%	0.3%	1.2%	1.7%
2	61%	14%	0.1%	11%	0.7%	8.7%	1.1%	0.3%	4.0%
3	42%	43%	0.1%	2.0%	4.9%	1.6%	1.2%	3.0%	2.5%
4	41%	41%	0.1%	2.0%	6.0%	2.1%	1.7%	3.2%	3.5%
5	64%	18%	0.1%	7.3%	2.1%	4.4%	0.5%	0.6%	3.1%
6	9%	79%	4.6%	1.1%	2.2%	0.2%	0.3%	3.0%	1.1%
7	19%	61%	0.3%	0.3%	12%	0.2%	0.4%	4.2%	2.6%
8	61%	27%	0.2%	4.1%	3.0%	1.9%	0.4%	1.1%	1.7%
9	78%	12%	0.0%	7.2%	1.1%	1.0%	0.2%	0.3%	0.9%
10	82%	11%	0.2%	2.0%	1.8%	1.1%	0.1%	0.4%	1.6%
11	5%	90%	0.3%	0.6%	1.0%	0.2%	0.2%	2.1%	0.3%

Note: Darker blue indicates higher product choice probabilities. Lighter gray indicates lower product choice probabilities. BF indicates branded fares. Anc. indicates ancillaries. Ch. indicates (sales) channels.

There seem to be substantial differences in customer choice behavior depending on sales channel. A notable example is BF D without paid ancillaries, which is selected 10.7% of bookings from Channel 2, whereas it is only selected 0.3% of bookings from Channel 7. BF B plus ancillary I is however selected 12.2% of bookings from Channel 7, but only 0.7% from Channel 2. Also, the choice probability with respect to the two most often selected products, i.e. BF

A and BF B without paid ancillaries, vary considerably between sales channels: between 5.3% and 81.8% for BF A without paid ancillaries, and between 11.0% and 90.0% for BF B without paid ancillaries.

As with booking weekdays, the JS divergence is used to quantify differences in observed product choice probabilities between different sales channels as well as across all channels.

Table 10 visualizes the JS divergence between the different sales channels. The biggest JS divergences are measured between channel 11 and channels 2, 9 and 10. In these cases, the two respective probability distributions are closer to having disjoint support than being identical, indicated by JS divergences greater than 0.5 (50%). Channels 3 and 4 seem to exhibit very similar customer choice behavior with their JS divergence close to zero. Further, the last row shows airlines incur the biggest error for Channel 11 when not differentiating choice probabilities based on sales channel.

Table 10: *JS divergence (in %) between customer choice probability distributions of bookings made through different sales channels, and between the different channels and the entire population (all channels).*

Sales ch.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	0%										
2	10%	0%									
3	2%	14%	0%								
4	3%	13%	0%	0%							
5	5%	1%	8%	8%	0%						
6	25%	45%	16%	17%	39%	0%					
7	13%	35%	7%	7%	27%	7%	0%				
8	1%	6%	4%	4%	2%	29%	19%	0%			
9	7%	5%	15%	15%	3%	49%	36%	4%	0%		
10	6%	7%	14%	14%	4%	48%	35%	4%	1%	0%	
11	31%	52%	21%	22%	46%	3%	10%	36%	56%	56%	0%
All	2%	11%	0%	0%	6%	18%	8%	3%	13%	12%	24%

Note: Darker blue indicates higher Jensen-Shannon divergences. Lighter gray indicates lower Jensen-Shannon divergences. Ch. indicates (Sales) channels.

Interestingly, JS divergences on sales channels are orders of magnitude higher than those on booking weekdays. Whereas most products, except for BF A and BF B without paid ancillaries, exhibit fairly similar choice probabilities independent of booking weekday, this is not the case for sales channels. On the one hand, this presents an argument to airlines to differentiate their offers based on sales channels. On the other hand, the JSD between all channels and the most frequently used sales channel, Channel 4, is 0%. This is not surprising given the fact that Channel 4 dominates the choice probability distribution for all channels more than any other channel, precisely because it is the most frequently occurring one. On the contrary, Channels 11, 6 and 9, which exhibit the largest JSD when compared to all channels, are used relatively infrequently. Whilst Channels 6 and 9 are used for only 1% and 0%, respectively, of all bookings, Channel 11 is used for 4% of all bookings. Still, it exhibits the highest JSD compared to all Channels. This suggests airlines gain much in terms of prediction accuracy when modeling customer choice behavior specifically for Channel 11 when they know a search comes through this channel, as opposed to ignoring the sales channel information.

4.4.2 Flight characteristics

Flights can be very different from one another. Flight lengths vary between few minutes and more than 15 hours. This analysis differentiates customer choice probabilities based on whether the flight leg was an overnight flight or not. Overnight flights are defined to leave before midnight (local time at origin) and arrive after midnight (local time at destination).

Table 11 shows observed customer choice probabilities based on whether the flight leg was an overnight flight or not. There seem to be notable differences in customer choice behavior between the two. For example, the choice probabilities of BF B plus ancillary I, and BF B plus ancillary II are 91% and 146% higher, respectively, when a flight leg is an overnight flight.

Table 11: *Observed product choice probabilities (in %) for different flight characteristics (overnight flight yes/no).*

Over-night flight?	BF A + no anc.	BF B + no anc.	BF C + no anc.	BF D + no anc.	BF B + anc. I	BF A + anc. III	BF B + anc. III	BF B + anc. II	Other
No	45.0%	37.5%	0.2%	3.1%	5.3%	2.3%	1.1%	2.4%	3.0%
Yes	16.7%	61.7%	0.0%	0.9%	10.1%	0.3%	0.4%	5.9%	4.0%

Note: Darker blue indicates higher product choice probabilities. Lighter gray indicates lower product choice probabilities. BF indicates branded fares. Anc. indicates ancillaries.

Table 12 shows the JS divergences between overnight flight and no overnight flight, and between both these categories and the entire population of all flights.

Table 12: *JS divergence (in %) between customer choice probability distributions based on flight characteristics (overnight flight yes/no), and between the flight characteristics and the entire population (all flights).*

Overnight flight?	No overnight flight	Overnight flight	All flights
No	0.0%		
Yes	9.5%	0.0%	
All flights	0.0%	8.9%	0.0%

Note: Darker blue indicates higher Jensen-Shannon divergences. Lighter gray indicates lower Jensen-Shannon divergences.

Separating customer choice probabilities based on overnight flight yes/no, the JS divergence is higher than between any booking weekdays, but lower than between some sales channels.

4.4.3 Customer loyalty status

Most airlines use loyalty programs to incentivize customer loyalty. Depending on how many flights a customer makes with the respective airline or airline group, distances flown, how much they pay, and in which compartment they travel, customers earn loyalty points. The more points they earn, the higher their loyalty category. At the same time, not all customers participate in loyalty programs. Table 13 shows observed customer choice probabilities based on customer loyalty status. There seem to be large differences in customer choice behavior depending on their loyalty status. For example, ancillary III is selected much more frequently in loyalty categories 2, 3 and especially 4. Also, these categories select uncommon choices with much higher likelihood, as indicated by the column “Other”. The names of the loyalty categories are anonymized to protect data confidentiality of the network airline.

Table 13: *Observed product choice probabilities (in %) for different loyalty categories.*

Loyalty cat.	BF A + no anc.	BF B + no anc.	BF C + no anc.	BF D + no anc.	BF B + anc. I	BF A + anc. III	BF B + anc. III	BF B + anc. II	Other
None	44.7%	41.5%	0.2%	2.4%	6.2%	0.0%	0.0%	2.8%	2.2%
1	44.8%	33.2%	0.1%	3.5%	4.6%	5.1%	2.6%	2.2%	3.9%
2	40.0%	23.3%	0.1%	6.0%	1.3%	14.9%	6.8%	1.1%	6.5%
3	32.2%	14.2%	0.1%	8.6%	0.7%	21.4%	6.9%	0.8%	15.1%
4	22.2%	16.5%	0.1%	7.5%	0.5%	24.2%	12.0%	0.8%	16.2%
5	41.9%	42.2%	0.1%	2.5%	7.6%	0.1%	0.1%	2.8%	2.7%

Note: Darker blue indicates higher product choice probabilities. Lighter gray indicates lower product choice probabilities. BF indicates branded fares. Anc. indicates ancillaries. Cat. indicates (loyalty) category.

Table 14 reports the JS divergences between different loyalty categories. It shows the biggest differences in observed customer choice behavior can be found between “no loyalty status” and loyalty categories 3 and 4, which also are most different from the entire population.

Table 14: *JS divergence (in %) between customer choice probability distributions from customers with different loyalty categories, and between the loyalty categories and the entire population (all loyalty categories).*

Loyalty category	**None**	**1**	**2**	**3**	**4**	**5**	**All**
None	0.0%						
1	4.5%	0.0%					
2	15.6%	4.6%	0.0%				
3	27.3%	12.9%	3.1%	0.0%			
4	32.6%	17.2%	5.4%	1.3%	0.0%		
5	0.2%	4.1%	15.3%	26.7%	31.7%	0.0%	
All	1.4%	1.2%	9.8%	20.2%	25.1%	1.2%	0.0%

Note: Darker blue indicates higher Jensen-Shannon divergences. Lighter gray indicates lower Jensen-Shannon divergences.

This section has identified a pattern. Customer choice behavior seems to differ based on multiple features of their search. The investigated features in this chapter include booking weekday, sales channel, flight characteristics, and customer loyalty status. Based on JS divergences, the relative importance between these can be ranked in descending order as sales channel, loyalty status, overnight flight yes/no, and booking weekday. The next section generalizes these findings.

4.5 Results and discussion

The fundamental hypothesis from Chapter 3 can be confirmed. The wealth of data available to airlines in customer searches enables highly granular and high-dimensional segmentation. The inductive research in this chapter shows the distribution of customer choice probabilities differs based on features of their search like sales channel, customer loyalty status, flight characteristics, and booking weekday during the tested period from September 2018 to September 2023.

Theorizing about the identified pattern, the inductive validation suggests a new hypothesis:

If airlines know and process search features like sales channel, customer loyalty status, flight characteristics, and booking weekday, then they can predict customer choice probabilities significantly more accurately. This way, and if this pattern continues to hold in the future, airlines could increase customer relevance by including these features in their assortment decision which customized offer to display.

Airlines need to solve several practical challenges to achieve differentiated offer management based on these features. Amongst others, airlines need to decide whether to differentiate by single features or multiple of them, and whether to move to real-time or repository-based offer management, or to develop a hybrid approach. As suggested by the findings of the inductive study, offer management differentiated by sales channel is an attractive starting point. Further differentiating features could then be iteratively added.

The next section concludes and outlines how future research can build on these findings.

4.6 Conclusion

Airlines serve a diverse range of customers but have not implemented effective and scalable strategies to differentiate their offer management with high degree of granularity. To solve this problem, a conceptual architecture has been proposed in Chapter 3 based on the fundamental hypothesis that airlines have a wealth of data in customer searches, which can enable highly granular and high-dimensional segmentation.

This chapter validates this hypothesis through an inductive study of 202 million coupons from a network airline spanning from 2018 to 2023. It calculates differentiated customer choice probabilities based on booking weekday, sales channel, whether the flight leg is overnight or not, and customer loyalty status. The key metric used to measure the statistical (dis)similarity between choice probability distributions is the Jenson-Shannon divergence, which is bound between zero (probability distributions are the same) and one (probability distributions are maximally different).

The inductive validation finds choice probabilities differ strongest based on sales channel. The Jensen-Shannon divergence is larger than 0.5 between some sales channels, indicating customer choice probabilities are more different than similar between these. Second strongest indicator of customer choice is loyalty status. Flight characteristics, concretely if a flight leg is an overnight flight or not, rank third. Booking weekday is the fourth strongest predictor with different customer behavior depending on whether a booking is made on a weekday (Monday-Friday) or weekend day (Saturday-Sunday).

Generalizing from these inductive findings, the chapter suggests airlines can improve prediction of customer choice probabilities, and hence display more relevant customized offers, when considering the features sales channel, loyalty status, flight characteristics, and booking weekday – in that order of importance – in their assortment decision.

Further academic research should seek a second validation of the results of this research. This could be done in different ways. First, the new hypothesis could be validated in deductive studies through offline or online tests. These studies could use different models ranging from simple forecasts to complex prediction models based on machine learning. Second, the inductive research could be repeated with other airlines or data from different time periods, or even in other industries that might exhibit similar behavior. These could include high-speed railways, ferry operators, air cargo or ocean carriers.

If confirmed from such follow-up studies, this research has important implications for airlines who want to display customized offers that are likely to match the need of the specific customer making a search. Notably, the segmentation method presented uses data readily available to airlines and included in customer searches, promising a low-cost process to increase prediction accuracy. Furthermore, the method developed is agnostic to individual customers. Instead, individuals are segmented based on anonymous characteristics of their search. As such, it can be applied to both customers that declare their identity and those that do not.

5 Deductive research: second validation[9]

Developing customized offer and pricing strategies have become a strategic priority to airlines and have attracted operations research accordingly. This chapter demonstrates feasibility of an offer management system architecture that aims to combine the simplicity of branded fares with the flexibility of unbundled ancillaries. Methodologically, it shows how discrete choice prediction accuracy can be improved with machine learning.

Chapter 5 confirms granular segmentation with data readily available to airlines and simple forecast models can significantly increase the prediction accuracy for future customer choice situations. These findings support a new balance between effectiveness and robustness of data-driven customer segmentation, increasing the number of segments from single digits in most traditional choice models to the magnitude of thousands.

Section 5.1 briefly summarizes literature context, motivation, and contribution. Section 5.2 describes a typical airline customer choice process and details on the real airline data used. It builds on the findings from Chapter 4 that search features like sales channel, customer loyalty status, flight characteristics, and booking weekday affect customer choice probabilities. Section 5.3 describes the methodology with three metrics to quantify prediction accuracy for future customer choices, three prediction models of increasing complexity, and five time periods analyzed. Section 5.4 contains results of the analysis. Section 5.5 discusses practical implementation for airlines. Section 5.6 suggests avenues for future research. Section 5.7 concludes the validation of the second fundamental hypothesis of the dissertation. Section 5.8 serves as an appendix, detailing the steps for the matrix factorization algorithm used.

[9] This chapter is based on a paper that is written and approved by the airline partner for publication.

5.1 Literature context, motivation, and contribution

The evolution from revenue management (RM) to offer management (OM) and customization of offers have been two key themes for airlines in the last two decades. RM aims to optimize airline revenue through flight pricing, whereas OM aims to optimize airline offers holistically. In this notion, an offer is defined as the price plus the product, which is the seat on an aircraft plus zero, one, or more ancillaries. The rise of low-cost carriers has increased the importance of ancillary revenues.

Moreover, advancements in pattern recognition through machine learning (ML) and online distribution enabled customization or personalization of offers, with e-commerce (Amazon) at the forefront of innovation (Bakos, 2001; Kashyap et al., 2022). Airlines aim for similar personalized or customized pricing and product strategies, responding to customer searches with context-specific offers that combine the flexibility of unbundled ancillaries with the convenience and simplicity of branded fares (BF). BF are pre-select bundles shown to customers including the flight ticket plus one or more ancillaries

5.1.1 Literature context and motivation

Both academia and practitioners have applied different classes of models to solve offer management to maximize customer or business outcomes for a given customer search. Customer outcomes can be offer relevance, i.e. to what degree the product offered matches customers' needs, and satisfaction, i.e. how satisfied customers are with the process of buying the product. Business outcomes can be search-to-book conversion, revenue, seat load factor, and profitability. The models can be categorized into statistical models like discrete choice analysis, and model-free ML algorithms. Established airline OM examples are Ratliff & Gallego (2013) and Wittman & Belobaba (2019) for discrete choice models, as well as Madireddy et al. (2017) and Shukla et al. (2019) for ML algorithms. Statistical models aim to draw inference from a sample, whereas ML aims to find generalizable patterns for prediction in high-dimensional data (Bzdok et al., 2018).

The novel OM architecture developed in Chapter 3 aims to define a new balance between effectiveness and robustness of high-dimensional segmentation. It goes beyond the traditional segmentation with travel purpose as one feature allowing airlines to distinguish two segments (Business vs. Leisure), by instead including up to nine features. The idea of the novel OM architecture is to improve prediction accuracy of discrete choice models through application of different ML algorithms, thereby allowing segmentation in an order of magnitude of thousands. The architecture rests on two key hypotheses:

- **Hypothesis 1:** The wealth of information in customer searches allows airlines to segment granularly and high-dimensionally, and the customers in the respective segments exhibit significantly different choice behavior.
- **Hypothesis 2:** Estimating segment-specific choice probabilities significantly increases the prediction accuracy of future customer choices.

With an inductive study of 202 million coupons from a major network airline between 2018 and 2023, Chapter 4 confirmed Hypothesis 1. Using the Jenson-Shannon divergence (Lin, 1991) to quantify differences between choice probability distributions, customer choice probabilities for 100 products seem to depend on all four features tested, in that order of priority: sales channel, customer loyalty status, whether the flight runs overnight or not, and booking weekday.

5.1.2 Contribution, aim, and research questions

The contribution of this chapter is to build on the inductive study through validation of Hypothesis 2 with deductive research with hundreds of millions of real airline coupons between 2018 and 2023. If airlines can improve prediction accuracy of discrete choice probabilities for future customer searches, then they can improve their assortment decision, display more relevant offers, and maximize expected business outcomes.

Specifically, five methodological research questions (RQ) will be tested.

- **RQ A:** Can segment-specific choice prediction, due to the informational value in customer searches, significantly improve prediction accuracy compared to a baseline of segment-agnostic choice prediction?
- **RQ B:** Can the inclusion of price, due to the informational value in prices paid, significantly improve segment-specific prediction accuracy even further?
- **RQ C:** Can a novel variant of the well-studied matrix factorization algorithm (Funk, 2006), due to overcoming data sparsity problems, significantly improve prediction accuracy of infrequent customer segments with less than ten bookings in the training data even further?
- **RQ D:** Can training the prediction model on the most recent month ensure, due to the high number of searches airlines generate, that changes in market and customer dynamics are captured timely, resulting in the highest prediction accuracy?
- **RQ E:** Can airlines even during the most disruptive times (training pre-Covid and testing in the beginning of Covid), achieve significant improvements in prediction accuracy?

These five RQ will be tested with three metrics of prediction accuracy, four prediction models with increasingly complex choice probability forecasts, and five experiments covering different time periods. With that, this chapter aims to suggest both the most accurate, and the most practical choices for airlines. This research is the first to substantiate an OM methodology with segmentation in the order of magnitude of thousands, validated on hundreds of millions of real airline coupons. The research is independent of personalized data, hence compliant with data privacy rules such as GDPR in Europe.

The remainder of the chapter is structured as follows: Section 5.2 describes the assumed customer choice process and provides explanatory statistics of the data used. Section 5.3 details the methodology with three metrics, four prediction models, and five experiments. Section 5.4 provides an in-depth view into the results of the three error metrics in the five experiments and for the four prediction models. Section 5.5 discusses the implications for practical airline implementation. Section 5.6 describes strands for future academic research building on top of this chapter. The chapter closes with a conclusion in Section 5.7.

5.2 Airline customer choice process and data

This section describes the assumed customer choice process, and the data used.

5.2.1 Customer choice process

The customer choice process is assumed to follow Figure 21, in line with the customer journey of the airline with four branded fares (A, B, C, D) and five paid ancillaries (I, II, III, IV, V) that can be combined. Except for the cheapest branded fare, they include some of the ancillaries that are available for a la carte purchase. Most airlines offer checked-in bags and seat reservation as part of some branded fares. Other value-added services, such as rebooking and cancellation options, are only available in more expensive branded fares but cannot be purchased as a la carte ancillaries. All possible combinations span a total of 100 products that customers can select from. A product is defined as a distinct combination of branded fare selection and zero, one or more paid ancillaries.

Figure 21: *Customer choice process, selecting branded fares in step 1 and amongst a list of a la carte ancillaries in step 2.*

5.2.2 Data preparation

The deductive research is based on 496 million coupons of a major network airline between September 2018 and September 2023 and actual customer purchase behavior. The data is anonymized to comply with data privacy rules.

72% of the airline coupons refer to Economy class. To validate whether airlines have sufficient data for high-dimensional segmentation, Economy is therefore the first class to be tested. After removing all compartments but Economy (i.e., removing First, Business, Premium Economy), 359 million coupons remained. A coupon refers to one passenger and one flight leg. To ensure data quality, additional cleansing removed coupons that did not have a branded fare associated, a create date, a coupon leg number, an operating airline, an origin airport, a destination airport, and/or any ticketed gross price information. After these cleansing steps, 202 million coupons remained. A booking can comprise multiple coupons. An example is a customer flying from Brussels to Montreal via London. This itinerary includes two coupons. Geographically, the data spans origins and destinations worldwide. The hub structure of the airline is reflected in geographical concentration on its home markets.

5.2.3 Exploratory data analysis

Table 15 shows customer choice behavior with respect to branded fare (BF) and ancillary selection. 89% of all bookings did not purchase any paid ancillary. This is especially true for BF A, C and D. Only customers purchasing BF B exhibit a somewhat higher probability of paying for additional ancillaries. Also, there is strong concentration on ancillaries I and II. Any combination involving paid ancillaries III, IV or V was purchased by less than 0.5% of customers for all four BF (see last column).

Table 15: *Exploratory data analysis of 202 million coupons, showing the observed choice frequencies of customers (in %) based on their branded fare selection.*

Branded fare (BF)	No paid ancillary	Paid ancillary I	Paid ancillary II	Paid ancillaries I and II	Other paid ancillaries
BF A	97%	<0.5%	3%	<0.5%	<0.5%
BF B	81%	11%	6%	2%	<0.5%
BF C	94%	<0.5%	6%	<0.5%	<0.5%
BF D	94%	<0.5%	6%	<0.5%	<0.5%
Total	89%	6%	4%	1%	<0.5%

In addition, the data allow observation of choice behavior depending on specific characteristics, or features, of the customer searches.

Table 16 shows substantial differences in customer choice behavior depending on sales channel. A notable example is BF D without paid ancillaries, which is selected 10.7% of bookings from Channel 2, whereas it is only selected 0.3% of bookings from Channel 7. BF B plus ancillary I is however selected 12.2% of bookings from Channel 7, but only 0.7% from Channel 2. Also, the choice probabilities with respect to the two most often selected products, i.e. BF A and BF B without paid ancillaries, vary considerably between sales channels.

Table 16: *Product choice probabilities (in %) for different sales channels.*

Sales ch.	BF A + no anc.	BF B + no anc.	BF C + no anc.	BF D + no anc.	BF B + anc. I	BF A + anc. III	BF B + anc. III	BF B + anc. II	Other
1	57.6%	31.9%	0.2%	1.6%	4.6%	0.9%	0.3%	1.2%	1.7%
2	60.7%	13.7%	0.1%	10.7%	0.7%	8.7%	1.1%	0.3%	4.0%
3	41.9%	42.8%	0.1%	2.0%	4.9%	1.6%	1.2%	3.0%	2.5%
4	40.9%	40.5%	0.1%	2.0%	6.0%	2.1%	1.7%	3.2%	3.5%
5	64.2%	17.7%	0.1%	7.3%	2.1%	4.4%	0.5%	0.6%	3.1%
6	8.7%	78.8%	4.6%	1.1%	2.2%	0.2%	0.3%	3.0%	1.1%
7	18.7%	61.1%	0.3%	0.3%	12.2%	0.2%	0.4%	4.2%	2.6%
8	61.1%	26.5%	0.2%	4.1%	3.0%	1.9%	0.4%	1.1%	1.7%
9	77.8%	11.5%	0.0%	7.2%	1.1%	1.0%	0.2%	0.3%	0.9%
10	81.8%	11.0%	0.2%	2.0%	1.8%	1.1%	0.1%	0.4%	1.6%
11	5.3%	90.0%	0.3%	0.6%	1.0%	0.2%	0.2%	2.1%	0.3%

Note: Darker blue indicates higher product choice probabilities. Lighter gray indicates lower product choice probabilities. BF indicates branded fares. Anc. indicates ancillaries. Ch. indicates (sales) channels.

To quantify differences between choice probabilities, the inductive study of Chapter 4 used the Jensen-Shannon divergence (JSD, Lin 1991). The JSD is bound between zero and one, with zero meaning two probability distributions are identical, and one meaning they are maximally different. Table 17 shows the JSD between some sales channels – e.g., between Channel 11 and both Channels 9 and 10 – is greater than 0.5. This means the respective probability distributions are closer to being maximally different than to being identical. On the contrary, Channels 3 and 4 have a JSD of only 0.2%, indicating customers booking through both these channels exhibit very similar choice behavior.

With a similar analysis, the inductive research concluded sales channel, customer loyalty status, overnight flight yes/no, and booking weekday all impact customer choice probabilities, in that order of priority.

Table 17: *JS divergence (in %) between customer choice probability distributions of bookings made through different sales channels, and between the different channels and the entire population (all channels).*

Sales ch.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	0%										
2	10%	0%									
3	2%	14%	0%								
4	3%	13%	0%	0%							
5	5%	1%	8%	8%	0%						
6	25%	45%	16%	17%	39%	0%					
7	13%	35%	7%	7%	27%	7%	0%				
8	1%	6%	4%	4%	2%	29%	19%	0%			
9	7%	5%	15%	15%	3%	49%	36%	4%	0%		
10	6%	7%	14%	14%	4%	48%	35%	4%	1%	0%	
11	31%	52%	21%	22%	46%	3%	10%	36%	56%	56%	0%
All	2%	11%	0%	0%	6%	18%	8%	3%	13%	12%	24%

Note: Darker blue indicates higher Jensen-Shannon divergences. Lighter gray indicates lower Jensen-Shannon divergences. Ch. indicates (sales) channels.

Building on these findings, this chapter tests whether airlines can use this knowledge to improve the prediction accuracy of customer choice probabilities for future customer searches. The next section presents the deductive methodology used.

5.3 Methodology

This section details the methodology used to predict choices of customized offers through highly granular segmentation and develops five methodological RQ to substantiate Hypothesis 2 holistically. Improvements in prediction accuracy are measured as a decrease in the error between predicted and observed choice probabilities, measured with three metrics. Four models consecutively refine the prediction from a segment-agnostic baseline to simple forecasts based on product and price to ML-enhanced prediction. The models are evaluated in an offline test with five experiments capturing five different time periods. In all analyses, choice probabilities are predicted and evaluated for the 100 products customers can select from.

5.3.1 Metrics

Prediction accuracy is evaluated with three metrics, all of which measure the prediction error between predicted and observed customer choice probabilities: Prediction Error (PE), Weighted Average Percentage Points Error (WAPPE), and Jensen-Shannon divergence (JSD).

First, the PE measures the error when predicting the most often selected product per segment for all customer searches of that segment. PE is calculated according to Equation (12):

$$PE_{test\ vs\ predict} = \sum_{s=1}^{S} w_s \left(1 - x_{s\_max,test}\right) \qquad (12)$$

s denotes the segments, *w* denotes the weights of the segments (relative occurrence of segment *s* in the test dataset), *s_max* refers to the product with highest prediction probability in the respective segment *s*, and *x* denotes the choice probabilities in the test data.

Second metric is the WAPPE. It extends PE by calculating the error when predicting choice probabilities for all products per segment. The WAPPE is calculated according to Equation (13):

$$WAPPE_{test\ vs\ predict} = \sum_{s=1}^{S} w_s \sum_{p=1}^{P} Abs(x_{s,p,test} - x_{s,p,predict}) \quad (13)$$

In Equation (13), *s* also denotes the segments, while *p* denotes the products, *w* denotes the weights of the segments (relative occurrence of segment *s* in the test dataset), and *x* denotes the choice probabilities in the test data and as predicted, respectively.

Third metric is the JSD, which is a statistical distance metric termed divergence in information theory. The JSD is a symmetric and bound extension of the Kullback-Leibler divergence and offers intuitive interpretability (Lin, 1991). It is calculated according to Equation (14):

$$JSD_{test\ vs\ predict} = \frac{1}{2} D_{KL}\ (P_{predict}\ ||\ M) + \frac{1}{2} D_{KL}\ (P_{test}\ ||\ M) \quad (14)$$

The Kullback-Leibler divergence D_{KL} for discrete probability distributions $P_{predict}$ and P_{test} on the sample space X is defined according to Equation (15) (MacKay 2003):

$$D_{KL}\ (P_{test}\ ||\ P_{predict}) = \sum_{x=1}^{X} P_{test}(\mathrm{x})\ log\left(\frac{P_{test}(x)}{P_{predict}(x)}\right) \quad (15)$$

Furthermore, M is the average of the two distributions $P_{predict}$ and P_{test} according to Equation (16):

$$M = \frac{1}{2}\ (P_{predict} + P_{test}) \quad (16)$$

With the logarithm in Equation (15) calculated to the base of two, the JSD shows is bound between zero and one, with zero meaning both probability distributions

$P_{predict}$ and P_{test} are identical, and one meaning $P_{predict}$ and P_{test} are maximally different, i.e. they have disjoint support.

5.3.2 Prediction models

This research applies three consecutive models to improve prediction accuracy from simple forecasts to more complex prediction models based on machine learning (ML): **segment-specific choice prediction** (Model 1), **segment-specific choice prediction with price information** (Model 2), and **ML-enhanced segment-specific choice prediction with price information for scarce segments** (Model 3). In addition, results are compared to a **baseline** (Model 0) as a benchmark for no segmentation.

With that analysis, the first three methodological RQ will be validated:

- **RQ A:** Can segment-specific choice prediction (Model 1), due to the informational value in customer searches, significantly improve prediction accuracy compared to the baseline (segment-agnostic choice prediction)?
- **RQ B:** Can the inclusion of price (Model 2), due to the informational value in prices paid, significantly improve segment-specific prediction accuracy compared to Model 1?
- **RQ C:** Can a novel variant of matrix factorization (Model 3), due to overcoming data sparsity problems, significantly improve prediction accuracy of infrequent customer segments with less than ten bookings in the training data compared to Model 2?

Prediction in the four models works as follows:

Model 0: In the **baseline,** predicted customer choice probabilities are equal to observed probabilities during training period. There is no segmentation, i.e. predicted choice probabilities are the same for all searches.

Model 1: In the **segment-specific choice prediction,** customer bookings are classified into segments based on four features: 7 booking weekdays, 11 sales channels, 2 flight characteristics (overnight flight yes/no) and 6 loyalty categories result in 924 segments (= 7 * 11* 2 * 6). Predicted choice probabilities are segment-specific and ignore price information. Model 1 predicts segment-specific observed probabilities during the training period when applied to the test period. Reductions in PE, WAPPE and JSD when going from the baseline to Model 1 allow inference on RQ A.

Model 2: In the **segment-specific choice prediction with price information,** segmentation is additionally done on four price buckets (very low, moderately low, moderately high, very high). Together with the previous four features from Model 1, this results in 3,696 segments (= 924 * 4). The bucketing is done by calculating the 25%, 50%, and 75% percentiles of total gross prices paid by the customers in the training dataset. Like Model 1, predicted segment-specific choice probabilities are equal to segment-specific observed probabilities during the training period. Reductions in PE, WAPPE and JSD when going from Model 1 to Model 2 allow inference on RQ B.

Model 3: In the **ML-enhanced segment-specific choice prediction with price information for scarce segments,** the number of segments is kept at 3,696 with the same features as with Model 2. As opposed to Model 2, for segments that occurred less than ten times in the training set, the predicted choice probabilities are updated through a novel application of the well-studied matrix factorization algorithm, originally proposed by Funk (2006). In this research, matrix factorization is applied for several reasons. First and foremost, Chapter 3 proposed applying it to improve choice probability prediction in the presence of data scarcity. The intuition is that one wants to use matrix factorization to update choice probabilities for infrequent segments, hypothesizing that these suffer from overfitting when predicting observed probabilities in the training set when applying on the testing set. Matrix factorization addresses this data sparsity problem by exploiting similarities between segments and products with four latent factors that are correlated to both segments and products. For segments that occurred at least ten times in the training set, no update is made to the predicted choice probabilities, i.e. these are exactly predicted like in Model 2.

Next to increasing prediction accuracy and model robustness, matrix factorization offers various additional benefits. One, factors can have interpretational value, such as value of time, convenience, business vs. leisure, etc. Two, matrix factorization reduces storage space required. Three, as opposed to Models 1 and 2, matrix factorization helps estimation of choice probabilities of new segments or for new products from day one. Whereas Models 1 and 2 can only give a data-based estimate for these once sufficient training data of the new segment or the new product have been generated. This is because matrix factorization can infer choice probability estimates for new segments or new products from the similarities identified between segments as well as between products. This ability to fill blanks in the segment-product matrix is precisely where matrix factorization has been shown extremely powerful in the case of movie, video, or product recommendation at Netflix, Youtube, and Amazon, respectively.

The run time of matrix factorization algorithms depends on multiple factors like the learning algorithm used, the hyperparameters such as the learning rate, the efficiency of the code and code libraries, as well as the hardware used. Also, it depends on the number of latent factors (set to 4 in this chapter), the number of columns (100 products), and the number of rows (3,696 segments). Section 5.8 describes an example configuration of the matrix factorization algorithm used. During the research, different configurations with different hyperparameters were tested in Python. Their results did not significantly differ. Resulting run times were between 5 and 30 seconds. Google Research (2022) provides details about computing times and how these scale when increasing the number of latent factors, columns, rows, or hyperparameters. Full optimization of the configuration is out of scope of this dissertation. It depends on many factors and will likely vary depending on how the using airline evaluates the complexity vs. effectiveness trade-off. For discussions how to optimize matrix factorization algorithms, the reader is referred to Aggarwal (2016) and Google Research (2022).

Reductions in PE, WAPPE and JSD when going from Model 2 to Model 3 allow inference on RQ C.

5.3.3 Time periods (experiments)

The entire five years span across both Covid and pre-Covid with potential major changes to customer behavior. To validate the methodology across different time periods and to test adaptability to changes in customer behavior, the whole period of five years is decomposed into five experiments: *Whole period*; *Pre-Covid*; *Covid*; *Disruption* (training on pre-Covid data and testing on Covid data); and *Recent month*. Experiments *Covid* and *Disruption* are specifically designed to analyze performance during times with rapidly changing customer behavior (see Yeoman, 2022, and Garrow et al., 2022, for an overview of how customer behavior changed with Covid).

Testing is conducted through an out-of-sample offline test, i.e. by training a prediction model on actual airline data and applying the predictions on different airline data not used for training. Table 18 summarizes the respective data split into training and test period for each experiment. Figure 22 in addition visualizes the timeline. With exception of the *Recent month* experiment, the training period includes at least 58 million airline coupons for each experiment.

Table 18: Details of the experiments.

Experiment	Training period; #observations	Test period; #observations
Whole period	Sep 1, 2018 – Mar 31, 2023; 171m	Apr 1, 2023 – Sep 30, 2023; 31m
Pre-Covid	Sep 1, 2018 – Aug 31, 2019; 58m	Sep 1, 2019 – Feb 29, 2020; 4m
Covid	Mar 1, 2020 – Mar 31, 2023; 109m	Apr 1, 2023 – Sep 30, 2023; 31m
Disruption	Sep 1, 2018 – Feb 29, 2020; 62m	Mar 1, 2020 – Aug 31, 2020; 6m
Recent month	Aug 1, 2023 – Aug 31, 2023; 5m	Sep 1, 2023 – Sep 30, 2023; 5m

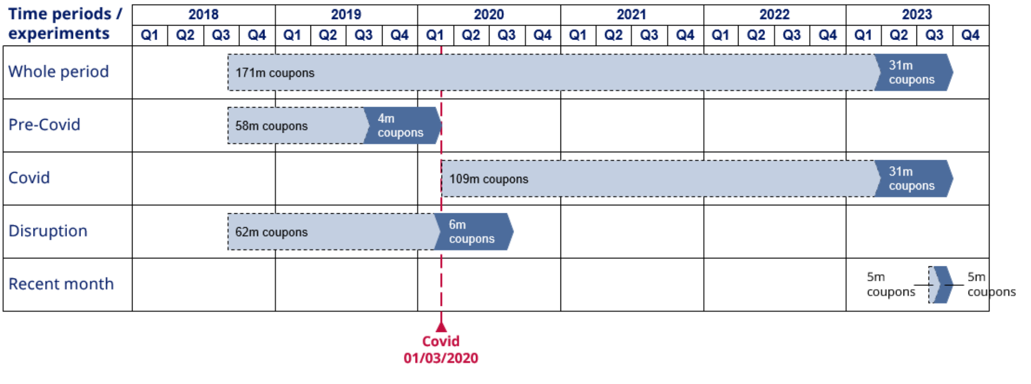

Figure 22: *Details of experiments, visualizing training and testing periods as well as the number of coupons in each.*[10]

With these five experiments, two further methodological RQ will be validated:

[10] Boxes in light blue and with dashed outlines indicate the training periods of the respective experiment. Boxes in dark blue with solid outlines indicate the test periods of the respective experiment.

- **RQ D:** Can training the prediction model on the most recent month ensure, due to the high number of searches airlines generate, that changes in market and customer dynamics are captured timely, resulting in the highest prediction accuracy?
- **RQ E:** Can airlines, even during the most disruptive times (training pre-Covid and testing on Covid periods), achieve significant improvements in prediction accuracy?

In the next section, results of the five experiments, four models, and three error metrics will be presented.

5.4 Results

This section presents results for PE, WAPPE and JSD for the four models and five experiments.

5.4.1 Model 0: segment-agnostic baseline

In Model 0 (baseline), predicted customer choice probabilities equal observed choice frequencies during the training period. As there is no segmentation, predicted choice probabilities are segment-agnostic and the same for all searches.

Table 19 compares the resulting PE, WAPPE and JSD in the five experiments.

Table 19: *Error metrics (PE, WAPPE and JSD, in %) without segmentation vary depending on experiment.*

Experiment	PE Model 0 (baseline)	WAPPE Model 0 (baseline)	JSD Model 0 (baseline)
Whole period	64.5%	51.4%	7.5%
Pre-Covid	55.8%	55.0%	8.3%
Covid	54.1%	49.6%	7.0%
Disruption	58.5%	51.6%	6.7%
Recent month	50.4%	48.5%	6.7%

All three metrics PE, WAPPE and JSD are lowest for the *Recent month* experiment. **This indicates validation of RQ D.** There seems to be sufficient data when training for one month only, and this seems to capture dynamics in customer and market trends best.

5.4.2 Model 1: segment-specific choice prediction

In Model 1, customer searches are classified into 924 segments based on booking weekday, sales channels, flight characteristics, and customer loyalty. Model 1 uses a simple forecast. For each segment, segment-specific predicted customer choice probabilities equal segment-specific observed frequencies during the training period. Table 20, Table 21 and Table 22 report on this product-based segmentation, comparing PE, WAPPE and JSD between Models 0 and 1 for the five experiments.

Table 20: *Error metric PE (in %) with segment-specific choice prediction vs. segment-agnostic baseline.*

Experiment	**PE Model 0 (baseline)**	**PE Model 1**	**Relative change PE**
Whole period	64.5%	51.1%	-20.7%***
Pre-Covid	55.8%	45.8%	-18.0%***
Covid	54.1%	50.1%	-7.4%***
Disruption	58.5%	51.0%	-12.8%***
Recent month	50.4%	41.3%	-18.0%***

Note: *** indicates significance at 99.9% confidence.

Table 21: *Error metric WAPPE (in %) with segment-specific choice prediction vs. segment-agnostic baseline.*

Experiment	**WAPPE Model 0 (baseline)**	**WAPPE Model 1**	**Relative change WAPPE**
Whole period	51.4%	25.4%	-50.7%***
Pre-Covid	55.0%	26.9%	-51.1%***
Covid	49.6%	25.2%	-49.2%***
Disruption	51.6%	32.2%	-37.6%***
Recent month	48.5%	21.7%	-55.3%***

Note: *** indicates significance at 99.9% confidence.

Table 22: *Error metric JSD (in %) with segment-specific choice prediction vs. segment-agnostic baseline.*

Experiment	JSD Model 0 (baseline)	JSD Model 1	Relative change JSD
Whole period	7.5%	1.2%	-84.3%***
Pre-Covid	8.3%	1.3%	-84.8%***
Covid	7.0%	1.1%	-84.9%***
Disruption	6.7%	1.1%	-83.5%***
Recent month	6.7%	0.3%	-96.3%***

Note: *** indicates significance at 99.9% confidence.

The results confirm RQ A. There is highly significant error reduction when going from Model 0 (segment-agnostic baseline) to Model 1 (segment-specific choice prediction). This effect is consistent across experiments and error metrics with 7.4%-20.7% relative PE, 37.6%-55.3% relative WAPPE, and 83.5-96.3% relative JSD decrease. For all three metrics, Model 1 performs strongest when trained on very recent data (*Recent month*), **substantiating the previous finding to RQ D.**

5.4.3 Model 2: enhancing predictions by inclusion of price

In Model 2, customer searches are classified into 3,696 segments based on booking weekday, sales channels, flight characteristics, customer loyalty, and price paid for the flight. For the latter, all bookings are categorized into four price buckets: very low, moderately low, moderately high, and very high. Each bucket is equally big, representing 25% of all bookings. In this analysis, the price paid is not adjusted for different origins, destinations, or distances.

The inclusion of price is motivated by observed differences in customer choice behavior depending on how expensive their flight ticket is. For example, over the entire 5 years, 70% of customers in the lowest 25% of price paid purchased the cheapest branded fare. This compares to 36% of customers in the highest 25% of price paid. On the contrary, the two most expensive branded fare options were chosen by customers with a likelihood 9 times higher when the flight was very expensive, compared to when the flight was very cheap. A detailed table of the branded fare choice probabilities depending on the price paid is not included in the thesis due to data confidentiality.

Like Model 1, Model 2 also uses a simple forecast. For each segment, segment-specific predicted customer choice probabilities equal segment-specific observed frequencies during the training period. Table 23, Table 24 and Table 25 report on this price-based in addition to product-based segmentation, comparing PE, WAPPE and JSD between Models 1 and 2 for the five experiments.

Table 23: *Error metric PE (in %) with segment-specific forecast that includes price information (Model 2) vs. Model 1.*

Experiment	PE Model 1	PE Model 2	Relative change PE
Whole period	51.1%	45.5%	-11.1%***
Pre-Covid	45.8%	40.7%	-11.2%***
Covid	50.1%	43.9%	-12.5%***
Disruption	51.0%	43.2%	-15.4%***
Recent month	41.3%	40.8%	-1.4%***

Note: *** indicates significance at 99.9% confidence.

Table 24: *Error metric WAPPE (in %) with segment-specific forecast that includes price information (Model 2) vs. Model 1.*

Experiment	WAPPE Model 1	WAPPE Model 2	Relative change WAPPE
Whole period	25.4%	19.0%	-25.2%***
Pre-Covid	26.9%	18.0%	-33.2%***
Covid	25.2%	17.4%	-30.8%***
Disruption	32.2%	17.4%	-45.9%***
Recent month	21.7%	7.3%	-66.2%***

Note: *** indicates significance at 99.9% confidence.

Table 25: *Error metric JSD (in %) with segment-specific forecast that includes price information (Model 2) vs. Model 1.*

Experiment	JSD Model 1	JSD Model 2	Relative change JSD
Whole period	1.2%	0.2%	-81.2%***
Pre-Covid	1.3%	0.3%	-77.9%***
Covid	1.1%	0.2%	-76.8%***
Disruption	1.1%	0.3%	-73.0%***
Recent month	0.3%	0.1%	-51.7%***

Note: *** indicates significance at 99.9% confidence.

The results confirm RQ B. There is highly significant error reduction when including price information, i.e. going from Model 1 to Model 2. This effect is consistent across experiments and for all three metrics with 1.4%-15.4% relative PE, 25.2%-66.2% relative WAPPE, and 51.7-81.2% relative JSD decrease. The strongest relative WAPPE decrease is observed for *Recent month,* whereas for both JSD and PE the relative decrease is weakest in *Recent month*. Just like Model 1, also Model 2 by far achieves the lowest WAPPE and JSD (and also almost lowest PE) when trained on very recent data (*Recent month*). In this case, it achieves WAPPE of 7.3% and JSD of 0.1%, which are less than half the WAPPE and JSD from the other experiments.

5.4.4 Model 3: resolving data sparsity with matrix factorization

In Model 3, customer searches are classified into the same 3,696 segments as in Model 2. For segments that occurred at least ten times in the training set, Model 3 uses the same forecast as Model 2, i.e. predicted customer choice probabilities equal segment-specific observed frequencies during the training period. The difference between Models 3 and 2 is the prediction for those segments that occurred less than 10 times in the training set. Whereas Model 2 still uses observed frequencies for those, Model 3 predicts choice probabilities based on the matrix factorization algorithm with four latent factors.

Table 26, Table 27 and Table 28 compare the resulting PE, WAPPE and JSD for 3,696 segments between Models 2 and 3 for the five experiments for those segments that occurred no more than nine times in the training set. Because Table 26, Table 27, and Table 28 report metrics for these infrequent segments only, Model 2 PE, WAPPE and JSD are different from Table 23, Table 24, and Table 25, in which the metrics were calculated for all segments.

Table 26: *Reducing error metric PE (in %) with matrix factorization (Model 3) for scarce segments vs. Model 2, calculated for the share of segments indicated (in %).*

Experiment	MF updates: % of segments; % of bookings	PE Model 2	PE Model 3	Relative change PE
Whole period	18.7%; 0.001%	71.2%	68.6%	-3.6%
Pre-Covid	25.8%; 0.01%	39.4%	28.7%	-27.0%***
Covid	22.2%; 0.002%	59.9%	62.9%	+5.0%
Disruption	24.9%; 0.01%	53.9%	45.5%	-15.7%**
Recent month	50.1%; 0.1%	53.8%	57.2%	+6.3%***

Note: ** and *** indicate significance at 99.0% and 99.9% confidence, respectively. No asterisk indicates no statistical significance of the relative change.

Table 27: *Reducing error metric WAPPE (in %) with matrix factorization (Model 3) for scarce segments vs. Model 2, calculated for the share of segments indicated (in %).*

Experiment	MF updates: % of segments; % of bookings	WAPPE Model 2	WAPPE Model 3	Relative change WAPPE
Whole period	18.7%; 0.001%	114.1%	71.9%	-37.0%*
Pre-Covid	25.8%; 0.01%	104.8%	64.8%	-38.2%***
Covid	22.2%; 0.002%	111.6%	82.9%	-25.8%
Disruption	24.9%; 0.01%	102.3%	55.9%	-45.3%***
Recent month	50.1%; 0.1%	96.8%	93.0%	-3.9%***

Note: * and *** indicate significance at 95.0% and 99.9% confidence, respectively. No asterisk indicates no statistical significance of the relative change.

Table 28: *Reducing error metric JSD (in %) with matrix factorization (Model 3) for scarce segments vs. Model 2, calculated for the share of segments indicated (in %).*

Experiment	**MF updates: % of segments; % of bookings**	**JSD Model 2**	**JSD Model 3**	**Relative change JSD**
Whole period	18.7%; 0.001%	44.0%	28.3%	-35.7%***
Pre-Covid	25.8%; 0.01%	38.5%	22.1%	-42.6%***
Covid	22.2%; 0.002%	43.2%	30.1%	-21.9%**
Disruption	24.9%; 0.01%	36.9%	18.8%	-51.3%***
Recent month	50.1%; 0.1%	36.6%	32.8%	-10.5%*

Note: *, ** and *** indicate significance at 95.0%, 99.0% and 99.9% confidence, respectively.

The results answer RQ C and indicate there can be improvements in prediction accuracy for infrequent segments when going from Model 2 to Model 3. Except for JSD, the effect is however not consistent across experiments, neither directionally nor with respect to significance. Especially interesting is the *Recent month* experiment reporting a highly significant increase in PE and significant decrease in WAPPE when going from Model 2 to Model 3. Looking at the five experiments, there are two (*Pre-Covid* and *Disruption*) for which all error metrics are significantly reduced in Model 3 compared to Model 2. There is one experiment (*Whole period)* with an indication of reduced error, one inconclusive experiment (*Covid)* and one experiment (*Recent month*) with the aforementioned conflicting result.

In summary, there is an indication that the matrix factorization can reduce the error for infrequent segments. However, since infrequent segments only account for 18.7%-50.1% of all segments representing 0.001%-0.1% of all bookings, airlines need to balance additional complexity from Model 3 compared to Model 2 vs. the gains in prediction accuracy for these infrequent segments.

5.4.5 Results summary

Table 29 summarizes the results for the five methodological RQ:

Table 29: *Summary of results for the five methodological research questions.*

	Methodological RQ	**Results**
A	Can segment-specific choice prediction (Model 1), due to the informational value in customer searches, significantly improve prediction accuracy compared to the baseline (segment-agnostic choice prediction)?	**Yes.** Product-based segmentation consistently and highly significantly decreases error metrics across all experiments by 7-21% (PE), 38-55% (WAPPE), and 84-96% (JSD).
B	Can the inclusion of price (Model 2), due to the informational value in prices paid, significantly improve segment-specific prediction accuracy compared to Model 1?	**Yes.** Price-based segmentation consistently and highly significantly decreases error metrics across all experiments by 1-15% (PE), 25-66% (WAPPE), and 52-81% (JSD).
C	Can a novel variant of matrix factorization (Model 3), due to overcoming data sparsity problems, significantly improve prediction accuracy of infrequent customer segments with less than ten bookings in the training data compared to Model 2?	**Inconsistent results with indication of error reduction,** although not consistent and not significant across all experiments. However, infrequent segments account for only 0.001-0.1% of all bookings.
D	Can training the prediction model on the most recent month ensure, due to the high number of searches airlines generate, that changes in market and customer dynamics are captured timely, resulting in the highest prediction accuracy?	**Yes.** The results suggest training on the recent month only. This gives consistently lowest errors for all three metrics for Models 0-2.
E	Can airlines, even during the most disruptive times (training Pre-Covid and testing during Covid), achieve significant improvements in prediction accuracy?	**Yes.** Across all five experiments, prediction accuracy can significantly be increased from Model 0 to 1 and Model 1 to 2.

In summary, the results suggest airline managers should adopt product- and price-based segmentation to capitalize on more accurate customer insights, they need to carefully evaluate whether the complexity of matrix factorization for sparse segments pays off in terms of additional accuracy, and they should continuously update their models with the latest data. Frequent model updates are particularly crucial during disruptive times.

In the next section, the practical implementation of the presented results will be discussed.

5.5 Discussion for practical implementation

This section discusses relevance and implications of the results for practical airline implementation and bridges to future research avenues.

The methodology presented is implementable as it only relies on data that are readily available to airlines as part of customer searches. Notably, the methodology does not require airlines to know the personal identity of a customer. Hence, it arguably presents a low-cost opportunity for airlines to significantly increase prediction accuracy of customer choice models. Specifically, airlines might consider the following questions for practical implementation:

Which prediction model should airlines use? The results show there are significant benefits for airlines going from Model 0 to Model 1, and from Model 1 to Model 2. The combined effect going from Model 0 to Model 2 is reported as relative PE decrease of 19.0%-29.5%, relative WAPPE decrease of 63.1%-84.9%, and relative JSD decrease of 95.5-98.2% depending on the experiment. This proves granular segmentation with simple forecast can be very effective in reducing prediction error. Going from Model 2 to Model 3 seems to bring further benefits, but only for max. 0.1% of customer bookings and not consistently across experiments. Due to the complexity of the matrix factorization performed, it seems questionable whether this pays off for airlines. The simple forecast could be the most practical trade-off between complexity and effectiveness. In addition, its simplicity can likely help navigate change management.

Is matrix factorization then to be neglected? Not necessarily. Matrix factorization could become more relevant if and when airlines further increase the number of segments to a magnitude of millions as suggested in Chapter 3. In this case, many more segments would become infrequent and there are indications that matrix factorization can indeed reduce the prediction error for these.

How many segments should airlines use? Airlines have a wealth of data available that enables highly granular and high-dimensional segmentation. This chapter confirms this for 3,696 segments, which is orders of magnitudes higher than previous models. It remains an open question whether the number of segments should be increased further. The results of the deductive validation suggest a risk of overfitting. Even with Model 3, JSD for infrequent segments are 18.8%-32.8%, whereas Model 2 achieves JSD of 0.1%-0.3% for all segments. For WAPPE and PE, it looks similar. When increasing the number of segments further, a higher share of them becomes infrequent, which might in turn lead to inferior prediction accuracy.

What time period should airlines train their prediction on? Model 2 results suggest major network airlines should train their prediction on the most recent month. This is the case in all three error metrics used. Both WAPPE and JSD are less than half of the other experiments with longer training periods for Model 2. Due to the wealth of searches major network airlines receive, the benefit from training on more data, i.e. going back in time six instead of one month, are outweighed by the advantage that recent month data allows airlines to capture changes in customer or market dynamics in a timely way.

How often should the prediction model be retrained? Whilst the results suggest training on the recent month, retraining the model once per week or once per day seems the appropriate trade-off. Various scholars (Shukla et al., 2019; Wittman & Belobaba, 2019) suggested transitioning from daily training to retraining on every event (e.g., customer transaction). For the use case at hand though, it seems practical that airlines keep predicted choice probabilities in a repository as part of their offline OM system and update the repository every week or day. Daily updates are also consistent with most network airlines' practice of recalculating bid price vectors every night. Only the application to actual customer searches happens in the online part of their OM system.

5.6 Future research

This research opens four main future research avenues, to (1) refine the methodology to achieve higher prediction accuracy, (2) expand to more comprehensive analysis of the conceptual architecture of Chapter 3, (3) derive policy and test effects on customer and business outcomes with live online tests, and (4) validate findings on other airlines or other sectors.

Improve methodology to achieve higher prediction accuracy. First, different features can be tested. This dissertation uses only select data available to airlines. One could either use different airline-own data (e.g., shopping context) or external contextual data (e.g., social media trends). Second, more features can be tested. This would increase the number of segments and reveal insights into improved prediction accuracy vs. overfitting, and whether matrix factorization could effectively resolve the resulting more severe data sparsity problems. Third, instead of deterministically assigning searches to segments, algorithms like Naïve Bayes could be used to model a probability distribution over all segments for each search. Fourth, a personalization module on top of the segmentation approach presented might increase prediction accuracy, as indicated by Vinod (2020). Fifth, techniques like exponential smoothing giving larger weight to more recent observations might improve prediction accuracy.

Sixth, for practical implementation airlines could refine matrix factorization, e.g. change the hyperparameter of latent factors to test whether consistent and significant error reduction can be achieved for infrequent segments.

Expand to more comprehensive analysis of conceptual architecture from Chapter 3. First, studies could target “Stream 1: Product choice” holistically through joint optimization of segmentation, bundling, ancillary/bundle pricing, and assortment strategies. Second, the “Flight pricing” stream could be independently analyzed with the segmentation approach presented in this dissertation to assess whether highly granular segmentation can lead to improved business outcomes from flight pricing decisions. Third, studies could aim to jointly optimize product choice and flight pricing, building on the simple first steps taken in Model 2.

Derive policy and test effects on outcomes with online tests. The deductive research proves prediction accuracy can be increased. However, only live tests with an assortment module will show whether airlines can derive a policy from this that improves customer (satisfaction) and business outcomes (conversion, revenue), and by how much customer discrete choice probabilities change based on how offers are presented to them. It would be especially interesting to measure whether a customized branded fare option will significantly decrease the probability of customers purchasing *a la carte* ancillaries after having selected their branded fares.

Validate findings on other airlines or other sectors. First, tests on smaller airlines could reveal whether less data is still sufficient to ensure frequent recalibration on one month of data whilst avoiding data sparsity problems. Second, research could investigate applicability on other transport modes, e.g. on high-speed railways, ferry operators, air cargo or ocean carriers.

5.7 Conclusion

Airline customers are different. Over the last decades, airlines have customized their offering to address different needs and to monetize ancillary revenues. At the same time, airlines aim to make the purchase experience convenient for their customers, which has led to the emergence of pre-selected bundles, so-called branded fares. To combine both customization and convenience, customer segmentation has become more relevant. Most existing airlines still today work with two segments, Business and Leisure, partly to avoid data sparsity problems when estimating customer choice probabilities.

This chapter proves segmentation orders of magnitude more granular than existing models can improve discrete choice probabilities of future customer choices. First, it proves product- and price-based segmentation with thousands of segments and simple forecasts without any machine learning can effectively and highly significantly lower prediction error for future customer choice probabilities compared to a baseline of no segmentation Second, for large airlines and segmentation in the magnitude of thousands, there seems positive yet limited benefit moving from simple forecast to machine learning with a novel application of matrix factorization. When further increasing the number of segments to the magnitude of millions, it will be an interesting research question whether matrix factorization can then reduce prediction error more significantly. Third, there is significant indication large airlines have sufficient data without incurring substantial data sparsity problems for at least 99.9% of all customer bookings. Fourth, because of the large number of data generated, the results suggest airlines to train their models on one month of past data as opposed to longer time periods to capture behavior changes in a timely way.

The findings support a new balance between effectiveness and robustness of the segments identified, and hence present an alternative approach to traditional customer choice models with typically two or single-digit numbers of segments and restrictive assumptions, such as the independence of irrelevant alternatives (IIA). It allows airlines to present their customers more relevant offers with simple segmentation approaches in a data-driven, automated, and cost-effective way. This allows airlines to combine the convenience and simplicity of branded fares with the flexibility of unbundled ancillaries. Notably, the models only process anonymous data readily available to airlines and in compliance with data privacy rules.

Further academic research could either refine the methodology to achieve higher prediction accuracy, expand to more comprehensive analysis of the conceptual architecture, derive a policy and test effects on customer and business outcomes with live online tests, or validate the findings on other airlines or other transport sectors.

5.8 Appendix: matrix factorization in Model 3

Step 1: Define the number of latent factors, rows, and columns.

Step 1.1: Define the number of latent factors as $f = 4$.

Step 1.2: Define the number of rows (customer segments) as $s = 3{,}696$.

Step 1.3: Define the number of columns (products) as $p = 100$.

Step 2: Create the matrices.

Step 2.1: Matrix $(S \cdot F)$ spanning rows and latent factors of size $s\ X\ f$.

Step 2.2: Matrix $(P \cdot F)$ spanning rows and latent factors of size $p\ X\ f$.

Step 3: Initialize the matrices.

Step 3.1: Factor matrices $(S \cdot F)$ and $(P \cdot F)$ are initialized with the absolute of random values drawn from a Gaussian distribution, multiplied by 0.1.

Step 3.2: Calculate the initial factorized matrix $(S' \cdot P') = (S \cdot F) \cdot (F^T \cdot P)$

Step 4: Set hyperparameters.

Step 4.1: Set the learning rate $\alpha = 0.05$.

Step 4.2: Set the regularization parameter $\beta = 0.01$.

Step 4.3: Set the number of epochs $n = 50$.

Step 5: Define the error function to be minimized.

Step 5.1: The error function ε measures the difference between the originally observed segment-product matrix $(S \cdot P)$ and the segment-product matrix from factorization $(S' \cdot P')$. It is calculated as $\varepsilon = Abs((S \cdot P) - (S' \cdot P'))$.

Step 6: Iterate through the training loop for each epoch.

Step 6.1: Alter the entries in the factor matrices $(S \cdot F)$ and $(F \cdot P)$.

Step 6.2: Compute the new error function ε.

Step 6.3: Update the matrices $(S \cdot F)$ and $(F \cdot P)$ with stochastic gradient descent to minimize the error function ε.

6 Discussion and synopsis of the research

This chapter synthesizes and reflects on the findings of the dissertation. The main insight is that machine learning can help improve prediction accuracy in discrete choice analysis. However, most of the benefit does not come from advanced machine learning models, but simply from high-dimensional segmentation using data readily available to airlines.

Section 6.1 answers the five PhD research questions introduced in Section 2.3.2, showing the path from problems to solutions. Section 6.2 discusses generalization of the findings to other airlines, other time periods, and other sectors. Section 6.3 summarizes implications for academic offer management literature and suggests future academic research avenues. Section 6.4 highlights practical implications for the airline industry and how airlines could capitalize on the research findings when implementing the proposed offer management system (OMS).

6.1 From research questions to answers

This section synthesizes the research results to answer the five distinct research questions from Section 2.3.2. The answers are explained in the subsections below. Each time, the research question is stated followed by the findings from this dissertation.

6.1.1 Can airlines identify distinct granular customer segments?

Research question 1: *Can airlines segment their customers into thousands to millions of distinct, clearly identifiable and MECE segments that exhibit significantly different choice behavior?*

This question was answered for both thousands and millions of segments.

Firstly, for thousands of segments, this question can be clearly answered with yes as confirmed by the inductive research of Chapter 4. The distinct segments, built exclusively on data available to airlines in customer searches, exhibit highly significantly different customer choice behavior in the period analyzed from September 2018 to September 2023. Specifically, the distribution of choice probabilities significantly differed based on the four features tested, in this order of significance. First, sales channels matter most. Measured by the Jensen-Shannon divergence of larger than 0.5, there are sales channels whose choice probability distributions are more different than identical. Second, customer loyalty status has a significant impact on customer choice behavior. Third, whether a flight is scheduled to run overnight or not impacts customer choice as well. Fourth, the weekday of the booking matters significantly between weekdays (Mondays-Fridays) and weekend days (Saturdays and Sundays).

Secondly, for millions of segments: this question cannot be answered with the research yet. It requires a follow-up study that increases the number of segments by increasing the number of features used. Due to the high significance for the magnitude of thousands of segments, it seems conceivable that customer choice probability distributions might also be significantly different for segments in the magnitude of millions.

6.1.2 Do granular segments improve future choice prediction?

Research question 2: *Can airlines use this segmentation to significantly improve the prediction accuracy of customer choice probabilities for searches in the future?*

Based on the deductive analysis presented in Chapter 5, the answer to this question is unequivocally affirmative. The prediction accuracy of customer choice probabilities, compared to a baseline of no segmentation, can be significantly improved when segmenting customer searches into thousands of segments. The results are highly significant across various experiments that cover five different time periods and the three error metrics Prediction error, Weighted average percentage point error, and Jensen-Shannon divergence.

With this, the research proves that segmentation on product choice alone, with thousands of segments, and using simple forecasts without any machine learning can effectively and highly significantly decrease the prediction error of future customer choice probabilities compared to a baseline of no segmentation. This holds true for all five experiments and all three error metrics.

Furthermore, expanding the segmentation model from product choice alone to include both product- and price-based segmentation can further significantly decrease the prediction error. This enhancement is consistent across all five experiments and all three error metrics.

6.1.3 Does matrix factorization solve the data sparsity problem?

Research question 3: *Can matrix factorization help solve the data sparsity problems when segmenting customers into thousands to millions of segments?*

This question can neither be answered with a clear yes nor a clear no. The results are inconsistent and not significant across experiments and error metrics, though there is an indication of error reduction. Still, the research allows insights into the potential of matrix factorization. These benefits can only be achieved for less than 0.1% of all bookings since the majority 99.9%+ do not suffer from data sparsity problems, at least not for the major network airline tested. This means there seems positive yet limited and not consistently significant benefit moving from simple forecast to machine learning with the novel application of the well-studied matrix factorization algorithm.

When further increasing the number of segments to the magnitude of millions, this research brings forward a new hypothesis that the number of segments suffering from data sparsity increases, the number of customers falling into a segment with data sparsity increases, and hence the error reduction with matrix factorization will be more significant.

6.1.4 Are changes in customer behavior captured timely?

Research question 4: *Can changes in customer behavior be captured, or how much of the prediction accuracy improvements can be achieved in a disruptive event like the Covid-19 pandemic?*

The Disruption experiment in Chapter 5 proves that, even in the most disruptive time training a model on pre-Covid and applying it on Covid data, significant improvements in customer choice prediction accuracy can be achieved. This is true for the thousands of segments tested.

Further, the different experiments show that training a model on only the recent month of data yields the highest prediction accuracy improvements. This did not come at the cost of data sparsity. Airlines as large as the major network airline tested seem to have more than enough data to allow segmentation in the magnitude of thousands even when only training on one month of data.

6.1.5 How should airlines implement the solution?

Research question 5: *What is practical advice to balance cost and effectiveness: Which features should airlines train on? How complex should the prediction model be? How many segments should airlines use? How long should the training period be? How often should airlines retrain their model?*

This research question includes various subquestions, which will be answered separately below.

In general, there is the obvious trade-off between complexity and effectiveness. The answers are intended to provide guidance for practical advice. If airlines evaluate the complexity-effectiveness trade-off differently, their optimal solution might vary.

Which features should airlines train on? This research has specifically investigated the explanatory power of four features, namely sales channel, customer loyalty status, whether a flight is scheduled to run overnight or not, and booking weekday. For the major network airline, these have been demonstrated to be significant in explaining customer choice behavior in this order of sequence mentioned. Additionally, the inclusion of price has been shown to highly significantly improve prediction accuracy further. Whilst the methodology itself is reproducible, other airlines who consider testing the model themselves may likely get different results in terms of which features can improve prediction accuracy and by how much. In general, it is recommended to first test potential features separately, and then build more complex models with multiple features.

How complex should the prediction model be? The results in this research prove that granular segmentation with simple forecasts can be very effective in increasing the prediction accuracy of customer choice probabilities. These simple forecasts predict that the past behavior will also hold true in the future. Machine learning, more precisely the novel application of the matrix factorization algorithm, seems to further improve on these simple forecasts. However, this was only shown for the last 0.1% of customer bookings, and even there the effect cannot be consistently and significantly measured across experiments and metrics. Based on this research, it seems questionable whether this justifies the much higher complexity of the matrix factorization model, at least as long as the goal is to enable segmentation in the magnitude of thousands as tested in this research. In this case, simple forecasts look like the most practical trade-off between complexity and effectiveness. Its simplicity will likely also help with change management within airlines. For segmentation in the magnitude of millions, however, many more of the segments will suffer from data sparsity, and matrix factorization will likely become more effective.

How many segments should airlines use? In the case of the major network airline tested, the wealth of data available in customer searches suggests thousands of segments are not necessarily the limit. The analysis in Chapter 5 proves that increasing the number of segments from 924 to 3,696 by including the price as additional feature, further improves the prediction accuracy. However, the research also emphasizes the risk of overfitting when increasing the number of segments. This in turn might lead to inferior prediction accuracy. In summary, answering this question critically depends on how many customer searches airlines have.

How long should the training period be? The results in this research are clear and suggest airlines to train their model on one month only of data. This has been shown to lead to higher prediction accuracy as changes in customer behavior are captured. In the case of the major network airline, this advantage clearly outweighs the disadvantage of fewer data to train on. Large network airlines simply have more than enough customer searches within just one month. Training the model on even shorter periods than one month has not been studied in this research.

How often should airlines retrain their model? Various academic scholars, e.g. Shukla et al. (2019) and Wittman & Belobaba (2019) suggest retraining the model after every event, i.e. every customer search or booking. Whilst this intuitively makes sense, as every new event carries new informational value, it would also come at high cost. An appropriate trade-off could be to retrain the model every night. This would be consistent with most network airlines' practice to recalculate bid price vectors every night.

Table 30 summarizes the key learnings for airline managers considering practical implementation of the solution based on the findings of this dissertation.

Table 30: *Summary of key learnings for practical airline implementation.*

Question	Key learnings
Which features should airlines train on?	Sales channel, customer loyalty status, overnight flight yes/no, booking weekday, and price.
How complex should the prediction model be?	Simple forecasts are effective. Matrix factorization slightly improves accuracy but adds complexity.
How many segments should airlines use?	Depends on the number of customer searches. Trade-off between accuracy and overfitting.
How long should the training period be?	Train on one month of data to capture changes in customer behavior.
How often should airlines retrain their model?	Retrain the model every night to balance the arrival of new informational value with retraining cost.

After answering these five research questions, the next section will discuss generalization of the research and its findings outside of major passenger network airlines.

6.2 Research generalization

This dissertation has shown the value of the suggested solution to the major network airline tested on. This section discusses whether the findings will likely generalize to other time periods, other airlines, or other transport sectors.

With respect to other time periods, the prediction accuracy improvement has been demonstrated for five different time periods. This makes it likely the research will generalize to a sixth, seventh, or n-th different time period.

With respect to other airlines or other transport sectors, it makes sense to first evaluate why it works for the major network airline, i.e. why it can solve the research problem for them. One, the tremendous number of searches and bookings generate tremendous amounts of data to train the model on. Two, customer preferences are heterogenous and can be split into distinct segments with significantly different purchase behavior. Three, this segmentation can be performed based on explicit or implicit information in the search. Four, network airlines control their own offer, but only do so when distributing via direct channels or using the NDC and One Order standards.

Another angle was given in Section 1.1 introducing three characteristics that define the customized offer management problem. First, customers differ with respect to both the product they demand and their WTP. This seems to be a main criterion to focus on. For example, it seems intuitive that freight companies likely have less heterogenous customers, and price likely plays a bigger role in their decision-making for or against certain products. The latter is likely also true for LCCs. Second, the product is customizable as base product plus ancillaries. This is likely true for other transport sectors as well. Third, there is a large amount of customer searches and bookings. This obviously depends on the size of the respective company analyzed.

Table 31 takes these criteria and hypothesizes about research generalization. These hypotheses are derived from relative comparisons to the presented results for the major network airline analyzed in the dissertation. These hypotheses will need to be tested before one can conclude they are actually true. The line of argumentation for how the hypotheses are derived is further detailed after the table.

Table 31: *Hypothesizing about generalization of the research findings to other transport sectors.*

Transport sector	Hetero-genous customers (product)	Hetero-genous customers (WTP)	Custom-izable product	Many searches and bookings	Hypothesis: will findings generalize?
Network airline analyzed	Yes	Yes	Yes	Yes	n/a
Other big network airlines	Yes	Yes	Yes	Yes	Very well
Smaller network airlines	Likely	Likely	Yes	Maybe	Likely well
Large LCCs	Likely	Much less	Yes	Yes	Likely well
Smaller LCCs	Somewhat	Much less	Yes	Maybe	Unclear
High-speed rail	Likely	Yes	Rather yes	Yes (if big enough)	Likely well
Long-distance buses	Likely	Much less	Yes	Yes (if big enough)	Unclear-likely
Car rentals	Yes	Yes	Yes (even more)	Yes (if big enough)	Likely very well
Passenger ferries	Somewhat	Likely	Somewhat	Yes (if big enough)	Maybe
Air cargo carriers	Limited	Rather not	Somewhat	Yes (if big enough)	Very unclear
Ocean carriers	Somewhat	Rather not	Somewhat	Yes (if big enough)	Unclear

Other major network airlines: There is no indication the proposed solution will not work with other major network airlines. It can be expected that all criteria are met: their customers are different with respect to the services they demand as well as their WTP for these services, they can customize their product, and they receive a large number of searches and bookings.

Smaller network airlines: The main difference to major network airlines is the number of searches and bookings. The practical implication might be that these airlines would need longer training periods than the suggested recent month. This will likely come at the cost of less fast adaptation to changing customer preferences. Still, the prediction accuracy improvements were also highly significant when training on longer periods. Also, customer heterogeneity might be less with smaller network airlines due to their concentration on certain geographies, markets, or customer segments. In conclusion, most of the research findings can likely be expected to generalize well.

Large LCCs: LCCs typically have customers that are much less heterogeneous with respect to their WTP. Also, they might be less heterogeneous with respect to the services they demand. For example, the LCC customer travel purpose is mostly leisure, whereas most network airlines cater to both leisure and business travelers. In summary, the presented findings can be expected to generalize when it comes to suggesting customized bundle options to LCC airline customers. This is conceivably true even though the benefit of a more convenient customer experience is arguably less important when selling to LCC customers. One aspect not included in Table 31 is distribution. LCCs often have a higher share of direct distribution. One of the biggest LCCs, Ryanair, for example exclusively distributed via their own website before signing a GDS agreement with Amadeus in late 2022 (Aviation Direct, 2022). Lastly, while the proposed solution might be more directly applicable to network airlines, LCCs are more dependent on ancillary revenues, and hence, would profit more from their optimization (CarTrawler, 2023; IdeaWorks, 2024).

Smaller LCCs: In addition to the discussion on large LCCs, smaller LCCs likely have less heterogenous customers and fewer searches and bookings to train their models on. This renders a hypothesis about research generalization unclear.

High-speed rail: Compared to major network airlines, the product might be less customizable. For example, rail operators typically do not charge for luggage. Otherwise, it can be hypothesized the research will likely generalize well to larger operators with sufficiently many searches and bookings. Hypothesizing further, search features might help operators to predict which customer browsing for tickets in second class might be open to an upgrade to first class at which price.

Long-distance buses: Similar to LCCs, customers can be assumed to be much more homogenous in their quest for low fares. This renders the value of the research to long-distance bus operators anything between unclear and likely yes.

Car rentals: Car rentals exhibit characteristics that make generalization of the findings highly likely. The product is arguably more customizable with all the different types of cars on offer. Car rental companies could define the smallest, cheapest car as equivalent to a seat on a flight, with anything bigger, faster, or more expensive an ancillary. Customers are heterogeneous from business to leisure and from looking for small all the way to luxury cars.

Passenger ferries: It seems conceivable that customers are less heterogeneous in their preferences for product, and passenger ferries have less ancillaries to differentiate their offers. This makes it difficult to hypothesize in favor of or against research generalization.

Air cargo carriers: Cargo customers are typically B2B, as opposed to all the sectors discussed before. B2B customers are typically very price-sensitive and also know the market price. Also, most air cargo customers likely choose this option because of its unmatched speed as single most important criterion. These characteristics make it very unclear whether a similar solution as presented in this dissertation generalizes to air cargo carriers.

Ocean carriers: Like air cargo carriers, but customers could be more heterogeneous. Generalization to ocean carriers seems more likely than for air cargo carriers, but still unclear. It could still add value when considering different routing options to customers.

6.3 Implication on academia

This section concludes the implication of the research on the academic offer management literature. It starts with the contribution made (Section 6.3.1) before discussing limitations of the proposed solution and suggesting future research avenues (Section 6.3.2).

6.3.1 Contribution to the academic literature

This research fills the gap between single digits of segments in existing (discrete) choice models, and infinite segmentation in some machine learning models. The research is novel as the proposed solution combines viability, usability and feasibility, validated with hundreds of millions of real airline data. First, viability has been established as data-driven cost-effective segmentation leads to significantly higher prediction accuracy for future customer choices that is able to adapt to changes in customer behavior. Second, it is usable due to the transparency, understandability, and explainability from mapping each search to precisely one segment. Third, the proposed solution is feasible as it can be implemented on data that already exists for airlines and is compliant with data privacy regulation. The results of this research are based on hundreds of millions of coupons from a major network airline between September 2018 and September 2023. With this, the proposed solution delivers on the criteria postulated in Section 1.3 when backwards engineering from the goal of solving the customized offer management problem.

Methodologically, it combines the strengths of the existing approaches by addressing the existing models' respective limitations: Existing DC models are not granular enough to represent the full heterogeneity of airline customers. And infinite segmentation with ML is not easily understandable, interpretable, and explainable, hence might struggle to get adoption from airline users. This research has developed a proposed solution for discrete, explainable segmentation with many segments that adequately represent customer heterogeneity. Specifically, the research has demonstrated the proposed solution can highly significantly increase prediction accuracy for a magnitude of thousands of segments, and it can do so for all customer searches, independent whether the personal identity of the customer is declared or not. Further, the research has shown that ML, more precisely a novel application of matrix factorization, can help solve data sparsity problems, and hence improve DC in high-dimensional spaces. However, for the major network airline, data sparsity problems only affected at most 0.1% of all customer bookings.

Zooming out from the customized airline offer management problem, there could be an implication for the wider academic community. The research has developed and tested a new use case for matrix factorization. Previously, this algorithm has been proven to successfully and highly effectively solve the problem of filling a sparsely populated user-item matrix for Netflix and other companies who want to recommend relevant content to their customers. Now, the algorithm has been shown to improve prediction robustness through repeated interactions. Here, the matrix is not sparse with many missing entries, but some entries are based on a handful of actual observations only. With the latent factors, this research has given an indication that matrix factorization can improve the robustness of these infrequent segments.

In summary, the academic contribution of this dissertation is that high-dimensional and data-driven segmentation, potentially aided by machine learning to solve data sparsity, can be combined with the understandability of discrete choice models with clearly identifiable segments. These findings have implications on the different subproblems in the offer management literature, as will be discussed next.

Segmentation. The goal of segmentation is to identify groups of customers with similar behavior. The proposed solution has shown this in a data-driven way using revealed instead of stated preferences. Instead of asking customers what they want, their actual behavior is measured. Customers are not assigned to a segment based on few manual rules, but purely based on features of their search. This is particularly beneficial when customer preferences and behavior change quickly. For example, in the beginning of the Covid pandemic, many airlines gave full flexibility with refund options to all their branded fares. In result, the main value proposition of the more expensive branded fares lost its value. The proposed solution has been shown to learn quickly and adapt. It can be expected that it would have suggested customized bundles based on other attributes.

Bundling. The goal of bundling is to create a product (bundle) that most closely resembles customer preferences. This goal has not directly been addressed in this research due to the assumption that all possible product combinations offered by the airline are always available with the *a la carte* ancillaries. Methodologically, this is not real-time bundling, although it might appear so for airline customers. It is also not pure bundling (Kobayashi, 2005), which would offer certain ancillaries as part of bundles only.

Pricing. The goal of pricing is to match customer WTP. The conceptual solution of Chapter 3 is aimed to help airlines achieve that. However, this hypothesis has not been validated in this dissertation. It remains to be tested in follow-up research.

Assortment. The goal of assortment is to optimize the selection of products to be displayed to a customer search. That is at the core of this dissertation. 11% of customers of the major network airline purchased a product that includes at least one ancillary in addition to their branded fare selection. Out of these 11%, more than two thirds purchased the product with the highest purchase probability in their respective segments when excluding branded fares. Put differently, these 7.5% of the total customers chose a product that was not displayed in step 1 of the customer journey but would have been when additionally displaying the most likely product as customized bundle in step 1. These 7.5% are the primary target group for which this research can make their customer journey more pleasant. They would see their preferred product already in step 1 as customized bundle, instead of having to navigate through the list of *a la carte* ancillaries in step 2. This can be hypothesized to increase customer satisfaction, which can be hypothesized to increase search-to-book conversion. The actual effect on customer and business outcomes remains to be tested with online tests in the airline's customer journey.

Customized offer management. The goal of customized offer management is to solve the subproblems in a way that allows airlines to use the information customers implicitly and explicitly provide in their search to develop customized product and pricing strategies. This research contributes to solving this problem by showing that assortment can be improved through data-driven, high-dimensional segmentation.

In conclusion, the OMS proposed and validated in this research has been demonstrated to increase prediction accuracy for future customer choices. What remains to be tested is whether this increased prediction accuracy leads to improved outcomes for customers and airlines.

6.3.2 Limitations and future research avenues

The research described in this dissertation has limitations. The findings indicate seven avenues for future research. These include testing the prediction accuracy against other models; refining the methodology; relaxing some of the assumptions; designing and conducting online tests; validating the findings on other airlines or other sectors in line with the generalization hypotheses developed in Section 6.2; testing the proposed solution to improve flight pricing decisions; and validating and testing the conceptual architecture from Chapter 3 in a holistic way. These seven pathways will be discussed in the following.

Test prediction accuracy against other models. Whilst the research has shown the proposed solution can significantly increase prediction accuracy compared to the base case of no segmentation on the features sales channel, customer loyalty status, overnight flight yes/no, and booking weekday, it has not tested prediction accuracy against models other than the baseline of no segmentation based on these features. This should happen on the same data. Three other models seem particularly interesting. First, deep learning ML models like neural networks can achieve the same magnitude of segments, or maybe even perform infinite segmentation like in Shukla et al. (2019). However, these often suffer from a lack of explainability as well as transparency, and hence lack of usability.. Despite this black box character, if these deep learning models outperform the proposed solution with simple forecasts in this dissertation, then it becomes a trade-off for airline executives to resolve. Second, DC models with single digits of segments are explainable, transparent, and hence usable by airline users. However, the findings from this dissertation suggest the DC models' prediction accuracies to be inferior to the proposed solution as the DC models neglect a large amount of the informational value explicitly and implicitly provided in customer searches. In other words, they do not make use of characteristics or features that this research has shown to have explanatory power for customer choice. Third, simpler ML models like Naïve Bayes would model probability distributions over all segments for each search, instead of deterministically assigning each search to precisely one segment. This arguably again comes at the cost of explainability, hence would be interesting to conclude whether their prediction accuracy would be higher than the solution proposed in this dissertation.

Refine the methodology. Validation in this dissertation has been performed on four features and based on customers who have converted their search into a booking. There are various conceivable strategies to refine this methodology. First, the no-purchase option could be included in the choice set. This would increase the number of searches that can be tested on and improve the understanding of customers who do not proceed from a search to making a booking. Second, different or more features could be tested than sales channel, customer loyalty status, overnight flight yes/no, and booking weekday. These could be airline-own features or external contextual data such as social media trends or weather forecasts. Adding more features will likely reveal important insights into the trade-off between prediction accuracy and the risk of overfitting. Third, the weighting of recent observations could be improved. The validation in this research simply defines a distinct training period. Either a booking is part of the training set, or it is not. Bookings that are part of the training set have all received the same weight. There are techniques, such as exponential smoothing, who give observations different weights. For example, more recent observations could be weighted higher, whilst those from the more distant past are still not completely discarded. Finally, the matrix factorization approach could be improved. For instance, the hyperparameter of the number of latent factors could be tweaked. Tuning hyperparameters is a typical step in refining ML models.

Relax some of the assumptions. The proposed solution makes various assumptions. Solving how to relax them and testing a new solution could add value to the scientific community. Two of these assumptions are that customers always choose the lowest-price booking class, and that branded fares represent an independent customer choice for all ancillaries included. It seems likely that the latter assumption is not true in reality. In the proposed solution, this would lead to an overestimation of customer choice probability of branded fares. Since the core goal is to offer a customized bundle that is not a branded fare, this however has no implication as long as the ordinality of the products that are not branded fares is unchanged.

Design and conduct online tests. Online tests would display the customized bundle in step 1 of the customer journey. Only then it can be proven that customer and business outcomes can in fact be improved from showing a customized bundle in addition to the static branded fares in step 1 of the customer journey. The basic hypothesis would be that showing customized bundles significantly decreases the selection of *a la carte* ancillaries in step 2 of the customer journey, precisely because the relevant product for the specific customer is already displayed as customized bundle in step 1. There are more research questions: Will this lead to higher revenues, higher search-to-book conversion, and higher customer satisfaction? Can the airline increase its market share because its product is perceived superior by its customers? Will it lead to higher customer loyalty? Also, will it lead to higher customer WTP? This last question might be more difficult to test as it is not directly observable though. In addition, it remains to be tested whether automation benefits from increased productivity hence cost saving can be achieved.

Validate on other airlines or other sectors. To expand the relevance of the research beyond major passenger network airlines, the analyses presented in this dissertation could be repeated with other airlines or other transport sectors. Potential candidates were discussed in Section 6.2 alongside with hypotheses how well the research might generalize. Three questions are of particular scientific interest. First, do smaller airlines have sufficient data to effectively train the model? Second, do LCCs have sufficiently heterogeneous customers? Due to their higher dependency on ancillary revenues than legacy carriers, this would be highly relevant to study. And third, does the proposed solution generalize beyond passenger airlines?

Test the proposed solution to improve flight pricing. Validation in this dissertation has been performed for the product choice stream, i.e. the left side of Figure 10. Future academic research could test the solution on revenue optimization potential from better flight pricing decisions. Research question would be whether segment-specific WTP estimates can outperform existing WTP estimation methods. This is also an interesting question with respect to potential automation benefits. Could it suffice if airlines, in the long term, estimate only one WTP band for each flight type instead of defining and manually maintaining a large number of price points? Flight types could for instance be a combination of origin-destination, flight weekday, and season.

Validate and test the conceptual architecture holistically. This is likely the biggest untapped research field. In addition to testing the proposed solution on flight pricing separately, the conceptual architecture can also be tested holistically, i.e. for joint optimization of product and price. Utilizing detected patterns for customized offer management rests on the assumption that customers have an overall WTP for the entire product instead of single WTPs for specific ancillary services. As such, it is congruent to the "simultaneous consumer" according to Bockelie & Belobaba (2017). If predictive power of such models turns out high, this could greatly reduce complexity of airline RM model landscapes, since cross-elasticities between different ancillary WTPs and substitution probabilities would not need to be explicitly modeled. Whilst the proposed models could already allow some derivations into customized WTP estimation, an explicit prediction of customized WTP could extend the proposed research. Whilst this dissertation does not detail out the final path to be taken, one could think of final model outputs as probability scores for each product and WTP point. This can be represented by a two-dimensional product space that contains all possible combinations of product at all possible WTP points. For each possible combination, the model would estimate the probability that this is the optimal one for the particular customer search. This can be visualized as three-dimensional product surface (Figure 23).

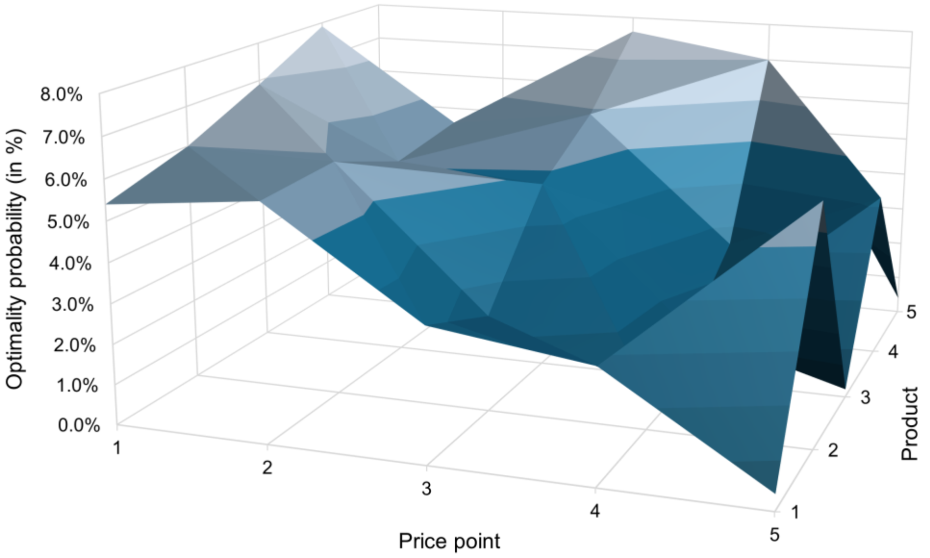

Figure 23: *Stylized example of the three-dimensional product space, which is defined by a combination of product, price point, and optimality probability*.

The product surface could allow for multi-modal distributions with various peaks. This could be plausible in case of “ambiguous” searches consistent with more than one segment. Going further, this could offer airlines the opportunity to think of resolving the ambiguity by asking customers bespoke questions or offering them a specific ancillary very cheaply. The customer’s reaction to such bespoke question or cheap ancillary offering could then help detect which of the various consistent segments is prevalent.

Finally, applying the holistic architecture can be hypothesized to contribute to automation benefits from simplifying airline processes and workflows. The proposed conceptual architecture largely eliminates the need to price specific ancillaries, branded fares, or upsells since WTP methodologically would be treated in the same fashion as product choice.

6.4 Implication on the airline industry

Airlines clearly had been at the forefront of revenue management decades ago. However, with respect to offer management, the airline industry has rather moved slowly for a while. This is partly due to airlines' dependency on legacy systems like Global Distribution Systems (GDS). In addition, it would not be uncommon that previous frontrunners exhibit some hesitation to experiment with novel approaches since the problem is perceived to be well-understood. Hence, the opportunity cost is assessed as greater than the potential value gained from experimentation.

The OMS proposed in this dissertation is different from industry practice and innovations propagated by academic and industrial scholars. Doing something differently from their competitors obviously poses a risk to airlines who consider experimenting with and implementing the proposed OMS. The demonstrated higher prediction accuracy has not been proven yet to lead to improved customer and business outcomes. On the other hand, it also presents an opportunity to gain a competitive advantage over other airlines in a cost-effective way.

From a methodological point of view, this dissertation aims to solve an operations research (OR) problem. For that reason, the implications on academia from Section 6.3 are also relevant to the airline industry, and their OR experts in particular. In addition, this section discusses practical implementation with benefits (Section 6.4.1), limitations (Section 6.4.2), and concluding remarks highlighting trade-off decisions to be made when implementing (Section 6.4.3).

6.4.1 Benefits for practical implementation

This subsection discusses four main benefits for airlines who decide to implement the proposed solution into their Offer Management System (OMS).

First, the proposed OMS **combines the convenience of branded fares with the flexibility of *a la carte* ancillaries.** If implemented, customers could benefit from both an intuitive and seamless customer journey that is still fully customizable to their needs. Airlines can gain this benefit at low cost since the proposed OMS only processes data that are readily available to airlines. Airlines do not need to purchase external data or set up completely new data infrastructure. The segmentation logic rests on customers' actual purchase decisions instead of customer surveys. This saves both cost to the airlines and, when available, revealed preferences are to be preferred over surveys asking customers how they would react in a hypothetical setting (Zamparini & Reggiani, 2007).

A second major advantage is that the conceptual architecture is designed in a way to **fit into existing RMS.** For distribution via direct channels, it is implementable right away. The holistic OMS is designed in modules to allow for simple and gradual embedding into RMS. The solution is also robust to work with both existing limitations (for example, twenty-six booking classes) and innovations (for example, continuous pricing). Due to its modularity, the architecture can support loosely coupled Application programming interfaces (APIs). Adding to its cost-effectiveness, the proposed solution uses a repository-based offer management instead of requiring real-time computation with fast response at the time of a customer search. When a customer searches, the corresponding customized bundle is simply read from the stored matrix. This is a relevant difference to true real-time bundling or complex deep learning models, in which the actual bundle is only calculated when a customer searches.

Usability to airline users like Revenue Analysts, Ancillary Analysts, or Offer Management Specialists is a third benefit. The segmentation of searches into segments is traceable and visible to airline users. This is of paramount importance to achieve adoption and embedding into existing business processes, workflows, and methods (Vinod, 2020). Further, users can be given steering opportunities within the proposed solution. The customer choice matrix is stored offline, which offers explainability as a key advantage over deep learning models. At the same time, airlines are advised to be deliberate about how much steering they want to give to users and balance the steering benefits with their cost in terms of lower automation and potential over-steering due to human biases.

As a fourth advantage, the proposed solution is **agnostic to individual customers.** It does not require customers to reveal their identity but works equally well with customers that do declare their identity and those that do not. Also, the proposed solution is adaptable to changing customer preferences. This dissertation has shown that significant prediction accuracy improvements can be achieved even in the most disruptive times, such as training the model on pre-Covid and applying the model on Covid data. Lastly, the model is also flexible to incorporate new ancillaries. These can initially be added as *a la carte* ancillaries. Over time, the model will learn which ones to display to which customer segment as part of the customized bundle.

6.4.2 Limitations for practical implementation

There are various limitations airlines need to be aware of when considering implementation of the proposed solution.

First, in the current distribution landscape for most airlines, and legacy carriers in particular, the full value of the proposed solution **can only be realized for direct channels, or once NDC and One Order are implemented.** Similarly, the full value of the conceptual architecture for flight pricing assumes full-fledged continuous pricing capabilities.

Second, the proposed solution maps each customer search into precisely one segment. In general, this greatly increases explainability and understandability, both substantially removing the barriers to adopt. At the same time, there might be **customer searches who could fit into several segments.** This problem might be solvable by designing the customer journey in a way to extract targeted information from the search. This could mean search-specific questions or ancillary offerings, hypothesizing that the customer's reaction to those will reveal the true segment.

Third, the proposed OMS models flight pricing, ancillary pricing, and product choice probabilities separately. This is a deliberate choice to keep the overall problem and solution tractable. Also, correlations between customer WTP for the mere flight seat and ancillaries, or between WTP and product choice decisions might be implicitly captured because the modules are trained on the same data. Still, it seems likely this leads to **suboptimization when comparing to the hypothetical scenario of an integrated optimization.**

Fourth, as mentioned in Section 6.3.2, airlines should be aware of the **assumption that branded fares represent independent customer choices** for the ancillaries included in the branded fares. Which is likely not true in reality. It is also likely not a major problem if the ordinality of products that include ancillaries does not change.

Fifth, airlines need to pay attention to the **risk of customers playing the airline.** For displaying customized bundles, there is no obvious incentive for customers to do so. They profit from being presented more relevant offers. However, when airlines use the proposed solution also to differentiate flight and/or ancillary prices, then customers have an incentive to pretend to belong to a segment with low WTP.

Sixth, airlines are obliged to **comply with anti-discrimination laws.** When the proposed OMS is implemented for bundling and assortment, there does not seem to be a potential conflict with these laws as long as all customer segments can still purchase the exact same products. The difference is only how the exact same products are presented to them, i.e. whether in step 1 or 2 of the customer journey. When the proposed OMS is implemented for flight and/or ancillary pricing, then airlines need to ensure adherence with price discrimination laws. They need to monitor which features airlines are allowed to price-differentiate on and which they are not. And they need to ensure this for each jurisdiction they operate in. The proposed OMS has taken various measures of precaution. It does not rely on personalization. The used features sales channel, loyalty status, flight characteristics, and booking weekday are less sensitive than personal data like nationality, age or gender. There might, however, be correlations between these sensitive criteria and the features mentioned. Airlines need to manage this risk.

6.4.3 Concluding remarks

In summary, the findings of this research support a new balance between effectiveness and robustness of customer segmentation, and how this can be used to improve assortment decisions. The proposed OMS presents an alternative approach to traditional customer choice model with typically two or single-digit numbers of segments and restrictive assumptions, for example the independence of irrelevant alternatives (IIA). It allows airlines to present their customers more relevant offers, using simple, understandable segmentation in a data-driven, automated, and cost-effective way. In particular, the usability and explainability differentiate this research from previously proposed ML models. The proposed OMS only processes anonymous data readily available to airlines and in full compliance with data privacy rules. Hence, the proposed OMS allows airlines to combine the convenience and simplicity of branded fares with the flexibility of unbundled, *a la carte* ancillaries.

Airlines could gain their own first experiences with the proposed OMS by starting with single features. Subsequently, more features can be tested and, if proven effective, added to the model. Based on the findings from the inductive research, this dissertation suggests starting with sales channel as the first feature.

For concrete practical implementation, airlines need to solve various trade-offs. Is the simple forecast good enough and maybe the better deal than matrix factorization to balance prediction accuracy with cost, complexity, and effort? In the dissertation it does 99.9% of the job, whereas complex matrix factorization might only improve on the last 0.1% of customer bookings. If the matrix factorization is implemented, then how many latent factors should be modeled? How often should airlines retrain their model, balancing frequent retraining with longer training periods, hence more data? Should airlines include more or fewer features, leading to a higher or lower number of segments? Should airlines maybe evaluate these two decisions together, and trade-off thousands of segments with monthly retraining versus millions of segments with half-yearly retraining? When and how should users have the option to steer parameters of the model or to overwrite the model's output? Examples could be when adding new ancillaries, or when existing ones should be promoted to a wider target audience. Another example could be when the competitive situation changes, for instance when a competitor publishes a new schedule, or introduces a new product or pricing model.

Zooming out from these methodological questions, the findings of this research can encourage airline senior leaders to rethink their organizational structure. The subproblems of assortment, segmentation, bundling and pricing are linked. How customers are segmented, how flights and ancillaries are priced, and how offers are presented to customers all together determine whether customers like the offer and ultimately convert their search into a booking, and how much revenue the airline can generate from this customer search. This research has not studied the organizational setup in detail. However, the link between these subproblems hints towards a close organizational alignment with a single-threaded leader accountable for converting customer searches into maximum possible outcomes, thereby using the different levers of how to segment customers, how to price flights, ancillaries, and bundles, and which product to offer to which customer at which step of the customer journey at which price.

Lastly, airline leaders are recommended to closely follow the ongoing (Gen)AI (r)evolution. (Gen)AI has the potential to stand at the forefront of transforming how humans and technology interact and jointly solve increasingly complex problems. For instance, it does not seem unthinkable that airline customers in the future enter their search request in a free text instead of pre-defined fields like origin, destination, departure date, length of stay, etc. And then a GenAI engine translates this unstructured text into the structured information required for the airline offer management to process. To capitalize on these developments, airlines will likely need to constantly rethink and upgrade their OMS to stay at the leading front of innovation.

References

General references

The following references were consulted throughout the research.

Doganis, R., 2019. *Flying off course: Airline Economics and Marketing*. 5th ed. Routledge.

Belobaba, P.P., Odoni, A., Barnhart, C. and Bamber, G., 2016. *The global airline industry*. 2nd ed. West Sussex: Wiley.

Blauwens, G., De Baere, P. and Van de Voorde, E., 2016. *Transport Economics*. 6th ed. Berchem: De Boeck.

Macário, R. and Van de Voorde, E. eds., 2010. *Critical issues in air transport economics and business* (Vol. 84). Routledge.

Nagle, T.T. and Müller, G., 2017. *The strategy and tactics of pricing: A guide to growing more profitably*. Routledge.

References cited in the thesis

Adams, W.J. and Yellen, J.L., 1976. Commodity bundling and the burden of monopoly. *The Quarterly Journal of Economics*, 90(3), pp. 475-498.

Aggarwal, C.C., 2016. *Recommender systems*. Vol. 1. Cham: Springer International Publishing.

Alamdari, P.M., Jafari Navimipour, N., Hosseinzadeh, M., Safaei, A.A. and Darwesh, A., 2020. A systematic study on the recommender systems in the E-commerce. *IEEE Access*, 8, pp. 115694-115716.

Albawi, S., Mohammed, T.A. and Al-Zawi, S., 2017. Understanding of a convolutional neural network. In: *2017 International Conference on Engineering and Technology (ICET)*. IEEE, pp. 1-6.

Albers, S., Gibb, J., Stabenow, S. and Daft, J., 2024. Breaking Bottlenecks: Power distribution dynamics in industry evolution. *Organization Studies*, 45(8).

Alpaydin, E., 2020. *Introduction to machine learning*. 4th ed. Boston: MIT.

Aviation Direct, 2022. About Amadeus: Ryanair enters GDS distribution. [Online] Available at: Via Amadeus: Ryanair enters GDS distribution – Aviation.Direct [Accessed 29 Oct. 2024].

Avionics International, 2020. The First 100 Days: How Airlines Responded to the COVID-19 Crisis. [Online] Available at: https://interactive.aviationtoday.com/avionicsmagazine/august-september-2020/the-first-100-days-how-airlines-responded-to-the-covid-19-crisis/ [Accessed 5 Oct. 2024].

Avram, B., 2019. Airlines customer segmentation in the hyper-competition era. *Expert Journal of Marketing*, 7(2), pp. 137-143.

Bahdanau, D., Cho, K. and Bengio, Y., 2014. Neural machine translation by jointly learning to align and translate. *arXiv*, arXiv:1409.0473.

Bakos, Y., 2001. The emerging landscape for retail e-commerce. *Journal of Economic Perspectives*, 15(1), pp. 69-80.

Bakos, Y. and Brynjolfsson, E., 1999. Bundling information goods: pricing, profits, and efficiency. *Management Science*, 45(12), pp. 1613-1630.

Basu, C., Hirsh, H. and Cohen, W., 1998. Recommendation as classification: Using social and content-based information in recommendation. In: *Proceedings of the 15th National Conference on Artificial Intelligence*, pp. 714–720.

Bellman, R.E., 1957. *Dynamic programming*. Princeton University Press.

Belobaba, P.P., 1987. Air travel demand and airline seat inventory management. PhD thesis, Massachusetts Institute of Technology.

Belobaba, P.P., 2011. Did LCCs save airline revenue management? *Journal of Revenue and Pricing Management*, 10(1), pp. 19–22.

Belobaba, P.P., Odoni, A., Barnhart, C. and Bamber, G., 2016. *The global airline industry*. 2nd ed. West Sussex: Wiley.

Bengio, Y., Courville, A.C., Goodfellow, I.J., Mirza, M., Ozair, S., Pouget-Abadie, J., Warde-Farley, D. and Xu, B., 2014. Generative adversarial nets. *Neural Information Processing Systems*.

Ben-Akiva, M. and Gershenfeld, S., 1998. Multi-featured products and services: analysing pricing and bundling strategies. *Journal of Forecasting*, *17*(3-4), pp.175-196.

Blauwens, G., De Baere, P. and Van de Voorde, E., 2016. *Transport Economics*. 6th ed. Berchem: De Boeck.

Bockelie, A., 2019. Ancillary services in the airline industry: Passenger choice and revenue management optimization. PhD thesis, Massachusetts Institute of Technology.

Bockelie, A. and Belobaba, P.P., 2017. Incorporating ancillary services in airline passenger choice models. *Journal of Revenue and Pricing Management*, 16(6), pp. 553–568.

Boser, B.E., Guyon, I.M. and Vapnik, V.N., 1992. A training algorithm for optimal margin classifiers. In: *Proceedings of the 5th Annual Workshop on Computational Learning Theory (COLT '92)*.

Bottou, L., 1998. Online algorithms and stochastic approximations. In: *Online Learning and Neural Networks*.

Breese, J.S., Heckerman, D. and Kadie, C., 2013. Empirical analysis of predictive algorithms for collaborative filtering. *arXiv*, preprint arXiv:1301.7363.

Bruch, E.E. and Mare, R.D., 2012. Methodological issues in the analysis of residential preferences, residential mobility, and neighborhood change. *Sociological Methodology*, 42(1), pp. 103-154.

Brumelle, S.L. and McGill, J.I., 1993. Airline seat allocation with multiple nested fare classes. *Operations Research*, 41(1), pp. 127–137.

Bulut, Z., Gürler, U. and Şen, A., 2009. Bundle pricing of inventories with stochastic demand. *European Journal of Operational Research*, 197(3), pp. 897–911.

Bzdok, D., 2017. Classical statistics and statistical learning in imaging neuroscience. *Frontiers in Neuroscience*, 11, p. 543.

Bzdok, D., Krzywinski, M. and Altman, N., 2017. Machine learning: a primer. *Nature Methods*, 14(12), pp. 1119-1120.

Bzdok, D., Altman, N. and Krzywinski, M., 2018. Points of significance: statistics versus machine learning. *Nature Methods*, 15(4), pp. 233-234.

Camilleri, M.A., 2018. Market segmentation, targeting and positioning. In: *Travel Marketing, Tourism Economics and the Airline Product*. Chapter 4, pp. 69-83. Springer, Cham, Switzerland.

CAPA Centre for Aviation, 2019. Spirit Airlines pushes up ancillaries, offsetting other challenges. [Online] Available at: https://centreforaviation.com/analysis/reports/spirit-airlines-pushes-up-ancillaries-offsetting-other-challenges-501102 [Accessed 25 Jan. 2021].

CarTrawler, 2023. Airline ancillary revenue reaches record $117.9 billion worldwide in 2023. [Online] Available at: https://corporate.cartrawler.com/wp-content/uploads/2023/10/The-CarTrawler-Worldwide-Estimate-of-Ancillary-Revenue_Final.pdf [Accessed 21 Sep. 2024].

Chiambaretto, P., 2021. Air passengers' willingness to pay for ancillary services on long-haul flights. *Transportation Research Part E*, 147, https://doi.org/10.1016/j.tre.2021.102234.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. and Sartin, M., 1999. Combining content-based and collaborative filters in an online newspaper. In: *Proceedings of the ACM SIGIR'99 Workshop on Recommender Systems*.

Cohen, J., Cohen, P., West, S.G. and Aiken, L.S., 2013. *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Cooper, W.L., Homem-de Mello, T. and Kleywegt, A.J., 2006. Models of the spiral-down effect in revenue management. *Operations Research*, 54(5), pp. 968–987.

Cramer, J.S., 2002. The origins of logistic regression (Technical report). 119. Tinbergen Institute, pp. 167-178. doi:10.2139/ssrn.360300.

Curry, R.E., 1990. Optimal airline seat allocation with fare classes nested by origins and destinations. *Transportation Science*, 24(3), pp. 193–204.

Dadoun, A., Defoin-Platel, M., Fiig, T., Landra, C. and Troncy, R., 2021. How recommender systems can transform airline offer construction and retailing. *Journal of Revenue and Pricing Management*, (2021), pp. 1-15.

Daft, J., Albers, S. and Stabenow, S., 2021. From product-oriented flight providers to customer-centric retailers: a dynamic offering framework and implementation guidelines for airlines. *Journal of Revenue and Pricing Management*, *20*(6), pp.615-625.

De, S.S., Dehuri, S. and Wang, G.-N., 2012. Machine learning for social network analysis: a systematic literature review. *The IUP Journal of Information Technology*, 8, pp. 30-51.

Dean, J. and Ghemawat, S., 2004. MapReduce: Simplified data processing on large clusters.

Deng, L. and Yu, D., 2014. Deep learning: methods and applications. *Foundations and Trends in Signal Processing*, 7(3-4), pp. 1-199.

Eliot, L.B., 2017. New advances in AI autonomous driverless self-driving cars: artificial intelligence and machine learning (1st ed.). LBE Press Publishing.

Fawagreh, K., Gaber, M.M. and Elyan, E., 2014. Random forests: from early developments to recent advancements. *Systems Science & Control Engineering*, 2(1), pp. 602-609.

Feldman, J., Zhang, D., Liu, X. and Zhang, N., 2018. Customer choice models versus machine learning: Finding optimal product displays on Alibaba. Available at SSRN 3232059.

Ferreira, K.D. and Wu, D.D., 2011. An integrated product planning model for pricing and bundle selection using Markov decision processes and data envelope analysis. *International Journal of Production Economics*, 134(1), pp. 95–107.

Fiig, T., Goyons, O., Adelving, R. and Smith, B., 2016. Dynamic pricing - the next revolution in RM? *Journal of Revenue and Pricing Management*, 15(5), pp. 360–379.

Fiig, T., Le Guen, R. and Gauchet, M., 2018. Dynamic pricing of airline offers. *Journal of Revenue and Pricing Management*, 17(6), pp. 381–393.

Fisher, M.M., 2000. Travel demand. In: Polak, J. and Heertje, A., eds.*Analytical transport economics*. Cheltenham: Edward Elgar, pp. 51-78.

Freiberger, O., 2024. Airline Revenue Management. [Online] Available at: https://www.linkedin.com/pulse/airline-revenue-management-olivier-freiberger-exwoe [Accessed 21 Sep. 2024].

Frias-Martinez, E., Magoulas, G., Chen, S.Y. and Macredie, R., 2006. Automated user modeling for personalized digital libraries.*International Journal of Information Management*,26 ,pp234–248.

Frias-Martinez,E ., Chen,S.Y.and Liu,X .,2009.Evaluationof a personalized digital librarybased on cognitive styles :Adaptivity vs.adaptability.*International Journalof Information Management*,29 ,pp48–56.

Fuller, W.C., Manski, C.F. and Wise, D.A., 1982. New evidence on the economic determinants of postsecondary schooling choices. *Journal of Human Resources*, pp. 477-498.

Funk, S., 2006. Netflix Update: Try This at Home. [Online] Available at: https://sifter.org/~simon/journal/20061211.html [Accessed 28 Jan. 2021].

Gallego, G., Li, A., Truong, V.-A. and Wang, X., 2016. Online resource allocation with customer choice. *arXiv preprint*, arXiv:1511.01837.

Gardner Jr, E.S., 1985. Exponential smoothing: The state of the art. *Journal of Forecasting*, 4(1), pp. 1-28.

Garrow, L.A., Lurkin, V. and Marla, L., 2022. Airline OR innovations soar during COVID-19 recovery. *Operations Research Forum*, 3(1). Cham: Springer International Publishing.

Golrezaei, N., Nazerzadeh, H. and Rusmevichientong, P., 2014. Real-time optimization of personalized assortments. *Management Science*, 60(6), pp. 1532–1551.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. and Bengio, Y., 2014. Generative adversarial networks. *arXiv preprint*, arXiv:1406.2661.

Google Research, 2022. Large-Scale Matrix Factorization on TPUs [Online] Available at: https://research.google/blog/large-scale-matrix-factorization-on-tpus/ [Accessed 21 Nov. 2024].

Gürler, U., Öztop, S. and Şen, A., 2009. Optimal bundle formation and pricing of two products with limited stock. *International Journal of Production Economics*, 118(2), pp. 442–462.

Gunning, D., 2017. Explainable Artificial Intelligence (XAI). [Online] Available at: https://www.cc.gatech.edu/~alanwags/DLAI2016/(Gunning)%20IJCAI-16%20DLAI%20WS.pdf [Accessed 28 Jan. 2021].

Guo, Z. and Loo, B.P., 2013. Pedestrian environment and route choice: evidence from New York City and Hong Kong. *Journal of Transport Geography*, 28, pp. 124-136.

Hao, E., 2014. Ancillary revenues in the airline industry: Impacts on revenue management and distribution systems. Master's thesis, Massachusetts Institute of Technology.

Hinton, G.E., Osindero, S. and Teh, Y., 2006. A fast learning algorithm for deep belief nets. *Neural Computation*, 18, pp. 1527-1554.

Hinton, G.E. and Zemel, R.S., 1994. Autoencoders, minimum description length and Helmholtz free energy. *Advances in Neural Information Processing Systems*, 6, pp. 3-10.

Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural Computation*, 9(8), pp. 1735-1780.

IATA, 2020. Together, let's build airline retailing – NDC program update. [Online] Available at: https://www.iata.org/contentassets/6de4dce5f38b45ce82b0db42acd23d1c/ndc-standard-presentation.pdf [Accessed 23 Jul. 2021].

IATA, 2023. Modern Airline Retailing. A Business Case. [Online] Available at: https://storage.pardot.com/123902/1680153969rCfTs5nF/Modern_Airline_Retailing___A_Business_Case____March_2023_.pdf [Accessed 1 Feb. 2025].

IATA, 2025. Airline Retailing. Together, Let's Build Modern Airline Retailing! [Online] Available at: https://www.iata.org/en/programs/airline-distribution/retailing/#tab-3 [Accessed 1 Feb. 2025].

IATA & McKinsey, 2022. Understanding the pandemic's impact on the aviation value chain. [Online] Available at: https://www.iata.org/en/iata-repository/publications/economic-reports/understanding-the-pandemics-impact-on-the-aviation-value-chain [Accessed 5 Oct. 2024].

IdeaWorks, 2019. Airline ancillary revenue projected to leap to $109.5 billion worldwide in 2019. [Online] Available at: https://ideaworkscompany.com/wp-content/uploads/2020/04/Press-Release-142-Global-Estimate-2019.pdf [Accessed 25 Jan. 2021].

IdeaWorks, 2024. The CarTrawler Yearbook of Ancillary Revenue. [Online] Available at: https://ideaworkscompany.com/reports/ [Accessed 6 Oct. 2024].

IdeaWorks and CarTrawler, 2018. Airline ancillary revenue projected to be $92.9 billion worldwide in 2018. [Online] Available at: https://www.cartrawler.com/ct/ancillary-revenue/ancillary-revenue-projections-2018/ [Accessed 25 Jan. 2021].

IdeaWorks and CarTrawler, 2023. Airline ancillary revenue reaches record $117.9 billion worldwide for 2023. [Online] Available at: https://corporate.cartrawler.com/wp-content/uploads/2023/10/The-CarTrawler-Worldwide-Estimate-of-Ancillary-Revenue_Final.pdf [Accessed 24 Feb. 2024].

Jain, A.K., 2010. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), pp.651-666.

Jannach, D., Zanker, M., Felfernig, A. and Friedrich, G., 2010. *Recommender Systems: An Introduction*. Cambridge: Cambridge University Press.

Jeffreys, H., 1948. *Theory of Probability*. 2nd ed. Oxford University Press.

Joyce, J., 2003. “Bayes' Theorem”, in Zalta, E.N., ed. *The Stanford Encyclopedia of Philosophy*. Spring 2019 ed. Metaphysics Research Lab, Stanford University.

Kashyap, A.K., Sahu, I. and Kumar, A., 2022. Artificial intelligence and its applications in E-commerce – a review analysis and research agenda. *Journal of Theoretical and Applied Information Technology*, 100(24), pp. 7347-7365.

Kaur, P., Krishan, K., Sharma, S.K. and Kanchan, T., 2020. Facial-recognition algorithms: A literature review. *Medicine, Science, and the Law*, 60(2), pp. 131-139.

Kim, H.N., Ji, A.T., Ha, I. and Jo, J.S., 2010. Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. *Electronic Commerce Research and Applications*, 9, pp. 73–83.

Kleinberg, J., Ludwig, J., Mullainathan, S. and Obermeyer, Z., 2015. Prediction policy problems. *American Economic Review*, 105(5), pp. 491-95.

Kobayashi, B.H., 2005. Does economics provide a reliable guide to regulating commodity bundling by firms? A survey of the economic literature. *Journal of Competition Law & Economics*, 1(4), pp. 707–746.

Kök, A.G., Fisher, M.L. and Vaidyanathan, R., 2008. Assortment planning: Review of literature and industry practice. In: *Retail Supply Chain Management*, pp. 99–153. Springer.

Kolbeinsson, A., Shukla, N., Gupta, A., Marla, L. and Yellepeddi, K., 2021. Galactic Air improves airline ancillary revenues with dynamic personalized pricing. Available at SSRN 383694.

Koren, Y., Bell, R. and Volinsky, C., 2009. Matrix factorization techniques for recommender systems. *Computer*, 42(8), pp. 30-37.

Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. *Neural Information Processing Systems*.

Kullback, S. and Leibler, R.A., 1951. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), pp. 79-86.

Kumar, R., 2021. Dynamic pricing of ancillaries using reinforcement learning. Presentation at the AGIFORS 2021 Revenue Management Study Group online conference, 4 May 2021.

Kupfer, F., Kessels, R., Goos, P., Van de Voorde, E. and Verhetsel, A., 2016. The origin–destination airport choice for all-cargo aircraft operations in Europe. *Transportation Research Part E: Logistics and Transportation Review*, *87*, pp. 53-74.

Lake, B., Ullman, T., Tenenbaum, J. and Gershman, S., 2017. Building machines that learn and think like people. *Behavioral and Brain Sciences*, 40.

Lautenbacher, C.J. and Stidham, S., 1999. The underlying Markov decision process in the single-leg airline yield-management problem. *Transportation Science*, 33(2), pp. 136–146.

LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), pp. 2278-2324.

Lin, J., 1991. Divergence measures based on the Shannon entropy. *IEEE Transactions on Information Theory*, 37(1), pp. 145-151.

Littlewood, K., 1972. Forecasting and control of passenger bookings. *Airline Group International Federation of Operational Research Societies Proceedings*, 1972(12), pp. 95-117.

Lopez-Nores, M., Garca-Duque, J., Frenandez-Vilas, R.P. and Bermejo-Munoz, J., 2008. A flexible semantic inference methodology to reason about user preference in knowledge-based recommender systems. *Knowledge-Based Systems*, 21, pp. 305–320.

Lu, Y., 2019. Impacts of ancillary services on airline revenue management. Master’s thesis, Massachusetts Institute of Technology.

Luce, R.D., 1959. Individual choice behavior: A theoretical analysis. *Frontiers in Econometrics*, 2, pp. 105-142.

Lufthansa, 2018. Conference & roadshow presentation. [Online] Available at: https://investor-relations.lufthansagroup.com/fileadmin/downloads/en/charts-speeches/presentations/RS_Presentation_Q3_CS_Conference_29Nov.pdf [Accessed 25 Jan. 2021].

MacKay, D.J.C., 2003. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press.

Madireddy, M., Sundararajan, R., Doreswamy, G., Nia, M.H. and Mital, A., 2017. Constructing bundled offers for airline customers. *Journal of Revenue and Pricing Management*, 16(6), pp. 532–552.

Mahendru, M., Singh, A. and Ranjan, J., 2024. Enhancing customer-centric retailing through AI-driven total offer management strategies for airline users. *International Journal of System Assurance Engineering and Management*, pp.1-18.

Mari, A., 2019. The rise of machine learning in marketing: goal, process, and benefit of AI-driven marketing. doi:10.13140/RG.2.2.16328.16649.

McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior.

McFadden, D., 1974. The measurement of urban travel demand. *Journal of Public Economics*, 3(4), pp. 303-328.

McKinsey, 2019. Airline retailing: The value at stake. [Online] Available at: https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/airline-retailing-the-value-at-stake [Accessed 1 Feb. 2025].

McKinsey, 2020a. Getting the price right in logistics. [Online] Available at: https://www.mckinsey.com/industries/travel-logistics-and-transport-infrastructure/our-insights/getting-the-price-right-in-logistics?cid=other-eml-alt-mip-mck&hdpid=bf8195b1-ae43-4655-bc7d-c381c28fa2eb&hctky=11253864&hlkid=5626f8c4504c45888b1eb7477703ecbf [Accessed 25 Jan. 2021].

McKinsey, 2020b. Executive's guide to developing AI at scale. [Online] Available at: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/executives-guide-to-developing-ai-at-scale?cid=other-eml-alt-mip-mck&hdpid=d36c6b61-313b-431c-b3d4-141bc805e7e2&hctky=11253864&hlkid=29289d1f7d714aed988781da33b394d4#intro [Accessed 28 Jan. 2021].

McKinsey, 2020c. Potential total annual value of AI and analytics across industries. [Online] Available at: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-executives-ai-playbook?page=industries/ [Accessed 28 Jan. 2021].

McKinsey, 2022. Taking stock of the pandemic's impact on global aviation. [Online] Available at: https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/taking-stock-of-the-pandemics-impact-on-global-aviation [Accessed 5 Oct. 2024].

McKinsey, 2024. Ready for takeoff: The airline retailing opportunity. [Online] Available at: https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/ready-for-takeoff-the-airline-retailing-opportunity?utm_medium=DSMN8&utm_source=LinkedIn&utm_user=14419233730516768 [Accessed 5 Oct. 2024].

Mikolov, T., Karafiát, M., Burget, L., Černocký, J. and Khudanpur, S., 2010. Recurrent neural network based language model. In: *11th Annual Conference of the International Speech Communication Association*.

Mikolov, T., Chen, K., Corrado, G. and Dean, J., 2013. Efficient estimation of word representations in vector space. [Online] Available at: https://arxiv.org/pdf/1301.3781.pdf [Accessed 14 Jul. 2020].

Millet, H., 2023. Leveraging booking information for offer personalization and GDPR consequences. *Journal of Revenue and Pricing Management*, 22(2), pp. 152-156.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, A.G., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, A., King, H., Kumaran, D., Wierstra, D., Legg, S. and Hassabis, D., 2015. Human-level control through deep reinforcement learning. *Nature*, 518, pp. 529-533.

Moore, G., 1965. Moore’s law. *Electronics Magazine*, 38(8), p. 114.

Murtagh, F. and Contreras, P., 2012. Algorithms for hierarchical clustering: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(1), pp. 86–97.

Nagpal, A., Jatain, A. and Gaur, D., 2013. Review based on data clustering algorithms. In: *Proceedings of the 2013 IEEE International Conference on Information and Communication Technologies (ICT 2013)*, pp. 298-303.

Nvidia, 2016. A visual representation of AI, machine learning, and deep learning. [Online] Available at: https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/ [Accessed 13 Jul. 2020].

Ødegaard, F. and Wilson, J.G., 2016. Dynamic pricing of primary products and ancillary services. *European Journal of Operational Research*, 251(2), pp. 586–599.

Page, L., Brin, S., Motwani, R. and Winograd, T., 1999. The PageRank citation ranking: Bringing order to the web. *Stanford InfoLab*.

Paredes, M., Hemberg, E., O'Reilly, U.M. and Zegras, C., 2017. Machine learning or discrete choice models for car ownership demand estimation and prediction? In: *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pp. 780-785. IEEE.

Park, D.H., Kim, H.K., Choi, I.Y. and Kim, J.K., 2012. A literature review and classification of recommender systems research. *Expert Systems with Applications*, 39(11), pp. 10059-10072.

PODS, 2021. MIT/PODS Revenue Management Consortium. [Online] Available at: https://www.podsresearch.com/consortium.html [Accessed 2 Feb. 2021].

Polydoropoulou, A., Tsouros, I., Pagoni, I. and Tsirimpa, A., 2020. Exploring individual preferences and willingness to pay for mobility as a service. *Transportation research record*, *2674*(11), pp.152-164.

Raghubir, P. and Valenzuela, A., 2006. Center-of-inattention: Position biases in decision-making. *Organizational Behavior and Human Decision Processes*, 99(1), pp. 66–80.

Raschka, S., 2018. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv*, arXiv:1811.12808.

Ratliff, R., 2021. An approach to airline offer management: Dynamic bundling and pricing. Presentation at the AGIFORS 2021 Revenue Management Study Group online conference, 5 May 2021.

Ratliff, R. and Gallego, G., 2013. Estimating sales and profitability impacts of airline branded-fares product design and pricing decisions using customer choice models. *Journal of Revenue and Pricing Management*, 12(6), pp. 509-523.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788.

Relay 42, 2024. How to increase airline ancillary revenue more effectively. [Online] Available at: https://relay42.com/resources/blog/how-to-increase-airline-ancillary-revenue-more-effectively [Accessed 21 Sep. 2024].

Resnick, P., Iakovou, N., Sushak, M., Bergstrom, P. and Riedl, J., 1994. GroupLens: An open architecture for collaborative filtering of netnews. *Computer Supported Cooperative Work Conf.*

Revelt, D. and Train, K.E., 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*, 80(4), pp. 647-657.

Ricci, F., Rokach, L. and Shapira, B., 2011. Introduction to recommender systems handbook. In: *Recommender Systems Handbook*, pp. 1-35.. Springer, Boston, MA.

Rodway, P., Schepman, A. and Lambert, J., 2012. Preferring the one in the middle: Further evidence for the centre-stage effect. *Applied Cognitive Psychology*, 26(2), pp. 215–222.

Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), p. 386.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature*, 323(6088), pp. 533-536.

Samek, W., Wiegand, T. and Müller, K.R., 2017. Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint*, arXiv:1708.08296.

Samuel, A.L., 1959. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3), pp. 210-229.

Sarwar, B., Karypis, G., Konstan, J.A. and Riedl, J., 2000a. Application of dimensionality reduction in recommender system - a case study. In: *Proceedings of the ACM WebKDD-2000 Workshop*.

Sarwar, B., Karypis, G., Konstan, J.A. and Riedl, J., 2000b. Analysis of recommendation algorithms for e-commerce. In: *Proceedings of the ACM E-Commerce*, pp. 158–167.Schmalensee, R.L., 1984. Gaussian demand and commodity bundling. *The Journal of Business*, 57(1), pp. 211-230.

Schmalensee, R.L., 1984. Gaussian demand and commodity bundling. *The Journal of Business*, 57(1), pp. 211-230.

Schubert, D., Sys, C. and Macário, R., 2021. Customized airline offer management: A conceptual architecture. *Journal of Revenue and Pricing Management*, pp.1-11.

Shardanand, U. and Maes, P., 1995. Social information filtering: Algorithms for automating 'Word of Mouth'. In: *Human Factors in Computing Systems Conference*.

Shlens, J., 2014. A tutorial on principal component analysis. *arXiv*, arXiv:1404.1100.

Shukla, N., Kolbeinsson, A., Otwell, K., Marla, L. and Yellepeddi, K., 2019. Dynamic pricing for airline ancillaries with customer context. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2174–2182.

Skift.com, 2017. US Millennials Travel the Most but Gen Z is on the rise. [Online] Available at: https://skift.com/2017/10/02/u-s-millennials-travel-the-most-but-gen-z-is-on-the-rise/ [Accessed 24 August 2019].

Srinivasan, B. and Komirishetty, G., 2021. Dynamic pricing & personalized ancillaries – Machine learning engine. Presentation at the AGIFORS 2021 Revenue Management Study Group online conference, 4 May 2021.

Stigler, G.J., 1963. United States v. Loew's Inc.: A note on block-booking. *The Supreme Court Review*, 1963, pp. 152-157.

Strauss, A.K., Klein, R. and Steinhardt, C., 2018. A review of choice-based revenue management: Theory and methods. *European Journal of Operational Research*, 271(2), pp. 375–387.

Szeliski, R., 2010. *Computer vision: Algorithms and applications*. London: Springer.

Szymanski, T. and Darrow, R., 2021. Shelf placement optimization for air products. *Journal of Revenue and Pricing Management*, (2021), pp. 1–8.

Talluri, K. and van Ryzin, G., 2004. Revenue management under a general discrete choice model of consumer behavior. *Management Science*, 50(1), pp. 15–33.

Taneja, I.J., Pardo, L., Morales, D. and Menéndez, M.L., 1989. On generalized information and divergence measures and their applications: A brief review. *Qüestiió: Quaderns d'estadística i investigació operativa*.

Taubmann, H., 2016. One year since Lufthansa’s distribution cost charge – what we’ve learned. [Online] Available at: https://amadeus.com/en/insights/blog/one-year-since-lufthansas-distribution-cost-charge-weve-learned [Accessed 25 Jan. 2021].

Teichert, T., Shehu, E., and Wartburg, I., 2008. Customer segmentation revisited: the case of the airline industry. *Journal of Transportation Research*, 42(1), pp, 227-242.

Train, K.E., 1978. A validation test of a disaggregate mode choice model. *Transportation Research*, 12(3), pp. 167-174.

Train, K.E., 1986. *Qualitative choice analysis: Theory, econometrics, and an application to automobile demand*. Vol. 10. MIT Press.

Train, K.E., 1998. Recreation demand models with taste variation over people. *Land Economics*, 74(2), pp. 230-239.

Turing, 2024. Top Generative AI Industry Applications: An In-Depth Look. [Online] Available at: https://www.turing.com/resources/generative-ai-applications [Accessed 22 Sep. 2024].

Van Ryzin, G. and McGill, J., 2000. Revenue management without forecasting or optimization: An adaptive algorithm for determining airline seat protection levels. *Management Science*, 46(6), pp. 760–775.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N. and Polosukhin, I., 2017. Attention is all you need. *arXiv preprint*, arXiv:1706.03762.

Vinod, B., 2020. The age of intelligent retailing: personalized offers in travel for a segment of ONE. *Journal of Revenue and Pricing Management*, 2020, pp. 1-7.

Vinod, B., 2021a. Advances in revenue management: the last frontier. *Journal of Revenue and Pricing Management*,20(1), pp.15–20.

Vinod, B., 2021b. Artificial intelligence in travel. *Journal of Revenue and Pricing Management*, (2021), pp. 1–8.

Vinod, B., Ratliff, R. and Jayaram, V., 2018. An approach to offer management: Maximizing sales with fare products and ancillaries. *Journal of Revenue and Pricing Management*, 17(2), pp. 91–101.

Walker, J.L. and Ben-Akiva, M., 2011. Advances in discrete choice: mixture models. In: De Palma, A., Lindsey, R., Quinet, E. and Vickerman, R., eds. *A Handbook of Transport Economics*. Cheltenham: Edward Elgar, pp. 160-187.

Wang, F. and Ross, C.L., 2018. Machine learning travel mode choices: Comparing the performance of an extreme gradient boosting model with a multinomial logit model. *Transportation Research Record*, 2672(47), pp. 35-45.

Wang, K.K., 2020. Airline Revenue Management with Dynamic Offers: Bundling Flights and Ancillary Services. Master's thesis, Massachusetts Institute of Technology.

Wang, K.K., Wittman, M.D. and Fiig, T., 2023. Dynamic offer creation for airline ancillaries using a Markov chain choice model. *Journal of Revenue and Pricing Management*, 22(2), pp. 103-121.

Wang, Z., Schaul, T., Hessel, M., Van Hasselt, H., Lanctot, M. and De Freitas, N., 2016. Dueling network architectures for deep reinforcement learning. In: *Proceedings of the 33rd International Conference on Machine Learning*, vol. 48.

Weber, F. and Schütte, R., 2019. A domain-oriented analysis of the impact of machine learning – the case of retailing. *Big Data and Cognitive Computing*, 3(11).

Werbos, P.J., 1990. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10), pp. 1550-1560.

Winston, C., 1985. Conceptual developments in the economics of transportation: an interpretive survey. *Journal of Economic Literature*, 23(1), pp. 57-94.

Wittman, M., 2018. Dynamic pricing mechanisms for airline revenue management: Theory, heuristics, and implications. PhD thesis, Massachusetts Institute of Technology.

Wittman, M.D. and Belobaba, P.P., 2018. Customized dynamic pricing of airline fare products. *Journal of Revenue and Pricing Management*, 17(2),pp. 78–90 .

Wittman, M.D. and Belobaba, P.P., 2019. Dynamic pricing mechanisms for the airline industry: a definitional framework. *Journal of Revenue and Pricing Management*, 18(2), pp. 100–106.

Wittmer, A. and Hinnen, G., 2016. Airline Passengers. In: *Air Transport Management – An International Perspective*. New York: Routledge, pp. 139-150.

Yadav, M.S., 1994. How buyers evaluate product bundles: A model of anchoring and adjustment. *Journal of Consumer Research*, 21(2), pp. 342–353.

Yeoman, I., 2022. The value of revenue management and pricing in a challenging environment. *Journal of Revenue and Pricing Management*, 21(5), pp. 485-486.

Zaharia, M., Chowdhury, M., Franklin, M.J., Shenker, S. and Stoica, I., 2010. Spark: Cluster computing with working sets. *HotCloud*, 10(10-10), p. 95.

Zamparini, L. and Reggiani, A., 2007. The value of travel time in passenger and freight transport: an overview. In: *Policy Analysis of Transport Networks*, pp. 145-161. London: Routledge.

Zhifei, S. and Joo, E., 2012. A review of inverse reinforcement learning theory and recent advances. In: *2012 IEEE Congress on Evolutionary Computation*, CEC 2012, pp. 1-8.