Modeling in Air Transportation:

Cargo Loading and Itinerary Choice

A Dissertation
Presented to
The Academic Faculty

by

Virginie LURKIN

In Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy in
Management Sciences

HEC-Ecole de gestion de l'Université de Liège

Jury:

Dr. Michael Schyns,	*Université de Liège*
Dr. Laurie A. Garrow,	*Georgia Institute of Technology*
Dr. Marie Lambert,	*Université de Liège*
Dr. Sabine Limbourg,	*Université de Liège*
Dr. Michel Bierlaire,	*École polytechnique fédérale de Lausanne*

Table of contents

List of Tables

List of Figures

Acknowledgments

When I was offered the opportunity to pursue a PhD on air transport project, I was both really excited and really afraid. After 3 years and half, I can say that it has been one of the most challenging things I've ever done and that this project would not have been possible without the advice and support of so many people.

First, I would like to acknowledge my doctoral advisor, Professor Michael Schyns, who offered me the opportunity to pursue a PhD in the field of air transport. I vividly remember our initial meeting and the passion in his voice when he was talking about airlines. Four years later, I feel incredibly privileged for having been able to pursue research on topics for which I am truly passionate and I want to thank him for that. Thank you for supporting me during these past four years. Thank you for putting your confidence in me and for believing in me (probably more than I believe in myself).

I am deeply grateful to Professor Laurie Garrow. My research has greatly benefited from her input, ideas and suggestions. Thank you for your constant availability for discussions and for sharing your international network of colleagues with me. Thank you also for being an ever present beacon of support, encouragement, and advice, despite the six hours and more than a thousand miles that separate us. I see the same drive and passion in your own research efforts, and I thank you for letting me do the same.

I would like to acknowledge all jury members, Professors Marie Lambert, Sabine Limbourg, and Michel Bierlaire, for agreeing to read the manuscript and to participate in the defense of this thesis. Thank you for your valuable questions, comments and suggestions that improved the presentation and contents of this dissertation.

I am also thankful to Jeffrey Newman and Matthew Higgins for the insights and knowledge they shared with me. The interest that you showed in my research is something for which I feel very grateful. Jeff, I am also so appreciative of your generosity with your time and resources.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the FNRS (Fonds National de la Recherche Scientifique). My work was also supported by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office (grant P7/36).

The research developed in the second chapter of this dissertation was initiated as part of a research project with TNT Airways. Special thanks are due to E. Meyer, M. Clety, H. Marchal, and JM. Urbani. I would also like to thank Angelo Guevara for his advice on instruments, Tulinda Larsen for providing us with schedule data, and Chris Howard and Asteway Merid of the Airlines Reporting Corporation for their patience and diligence in asnwering our many questions about their ticketing database.

Working within the research group "QuantOM" has been a great pleasure and I want to thank my excellent team of colleagues. In particular, thanks to Julien and Célia for all the time we spent together, discussing ideas about absolutely everything. Thank you for sharing both ups and downs, for always giving good advice and always being so supportive. This journey would not have been as colorful without you two. I also want to thank Stefano and Eli for all the good moments we shared together. Thank you for the support, the encouragement and shared laughter. You have brought the sun of your countries in our office. Thanks to Stéphanie for our *Freudian* conversations about research, teaching and life. Thanks to Professor Crama for all the things I've learned from him. Thank you for sharing your experience and expertise with all of the young padawans that we are. Thank you for your knowledge, your humility and your sense of humor. Thanks to Yasemin and Sabine for showing me that it is possible to be a good professor and a *Super-Mom*. Thanks to all the other members of QuantOM for the daily positive atmosphere and the numerous lively discussion sessions: Ashwin, Martine, Christine, Véronique, Guillaume, Thierry, Reza, Christian, Alessandro, Cédric, Anne-So and Hande.

I would like to thank all members involved in the organization of the AGIFORS conferences. I am very glad and honored for getting the opportunity to present my research at these conferences. The quality of the presentations there is amazing and all their social events are incredible. I've met many people whose knowledge and experience have helped me to progress in my research.

My research stay(s) in Atlanta would not have been as enjoyable without the many good friends I met there. I would like to thank them all. I also thank all my old friends for all the memories we share together and all the ones we will create.

Special thanks to my loving and caring family. Your support has been continuous fuel during my long journey in finishing this doctoral program. Thank you Dad for encouraging me to shoot for the moon and supporting me through every bump along the way. Thank you Mom for teaching me the real values in life and that my happiness is my own decision. Thank you, Mom and Dad, for showing me how to create balance in my lifestyle. I'm eternally grateful for the knowledge you shared with me and the irreplaceable life lessons that I live by. My gratitude for you is endless. Thanks to my twin, my best friend, my confidante and my greatest mentor, Céline. Your perpetual belief in my potential empowers me every single day. You're my *Jiminy Cricket*, the best conscience ever. I can't count the hours we spent together, the good advices you gave me and the laughs we shared. *You're simply the best, better than all the rest. Better than anyone, anyone I've ever met.* I also want to thank Jérome, for being like a brother to me. Thank you to my older sister, Stéphanie, and her daughter, my little sunshine Cassia, for being present in my life. Thanks to my grandparents for giving me constant love and encouragement. Thanks to my family-in-laws for welcoming me into your family.

Finally, I want to thank the man who has been with me through it all, Cédric. You never complained about the time I spent in the office or in front of my computer at home. You have always encouraged me and have always been present for me. I have no words to express how grateful I am for the sacrifices you have made to allow me to achieve this project. Thank you for making me laugh. Thank you for wiping my tears away. Thank you for being you. You rock my world.

"The crucial lesson was that the scope of things I didn't know wasn't merely vast; it was, for all practical purposes, infinite. That realization, instead of being discouraging, was liberating. If our ignorance is infinite, the only possible course of action is to muddle through as best we can."

Martin A. Schwartz

Summary

We examine two problems as part of this dissertation. The first is a cargo loading problem. The aim is to load a set of containers and pallets into a cargo aircraft that serves multiple airports. Our work is the first to model cargo transport as a series of trips consisting of several legs at the end of which pickup and delivery operations might occur. This problem is crucial for airlines because in an attempt to reduce their costs, most airlines prefer to load as many containers as possible, even if all the loaded containers do not have the same final destination. Our results demonstrate that it is possible to quickly find near optimal or excellent feasible loading plans, and that our approach leads to substantial savings with respect to typical manual approaches currently used in practice.

The second problem we examine involves the estimation of itinerary choice models that include price variables and correct for price endogeneity using a control function that uses several types of instrumental variables. The motivation for developing these models is to demonstrate the importance of accounting for price endogeneity and to estimate different price sensitivities as a function of advance purchase periods. This is important as the airline industry can use our results to incorporate different customer segments as revealed through high-yield and low-yield booking curves when evaluating the profitability of airline schedules.

Results based on Continental U.S. markets for May 2013 departures showed that models that fail to account for price endogeneity overestimate customers' value of time and result in biased price estimates and incorrect pricing recommendations. The advanced models estimated (nested logit and ordered generalized extreme value (OGEV) models) are shown to outperform the baseline multinomial logit model with regard to statistical tests and behavioral interpretations. Additionally, results show that price sensitivities vary as a function of advance purchase periods, with those purchasing high-yield products being less price sensitive than those purchasing low-yield products (across any advance purchase periods) and those purchasing closer to departure being less price sensitive. Results also indicate that inter-alternative competition is strong for itineraries that share similar departure times.

Finally, as part of the itinerary choice model developed in this dissertation, we estimate highly refined departure time of day preferences. Results are intuitive and show that departure time of day preferences vary across

many dimensions including the length of haul, direction of travel, number of time zones crossed, departure day of week, and itinerary type (i.e., outbound, inbound, and one-way itineraries). To the best of our knowledge, these curves represent the most refined publicly-available estimates of airline passengers' time of day preferences.

Chapter 1

Introduction

"When you realize that aviation, if it were a country, would be the 19th largest economy in the world, supporting 56.6 million jobs[1] and over two trillion dollars in economic impact, you really see the scale of air transport."

-Paul Steele, Executive Director of the Air Transport Action Group (ATAG).

1.1 Background

On the morning of 1 January 1914, the world's first scheduled commercial airline flight took to the air. The flight took place between St. Petersburg and Tampa, flying a total of 23 minutes with only one passenger on board. According to the Air Transport Action Group (ATAG), the world's airlines welcome today over three billion passengers a year and carry 50 million tons of freight. In addition, aviation provides 8.7 million direct jobs worldwide and contributes over $2.4 trillion to global gross domestic product (GDP) (ATAG (2014)). The progress that has been made over the past 100 years is quite remarkable. Airlines have evolved into a sophisticated business in an industry that has become a major driver of the global economy activity.

As shown in Table 1.1, airlines are expected to generate $29.3 billion in net profits in 2015, representing an increase of more than 75% compared to $16.4 billion in 2014. Despite 2015 record profits, the net profit margin is only 4.0% of revenues, or $8.27 per passenger (IATA (2016b)). Small profit margins, even in the best of times, are due to the complex nature of the business. Airlines face high costs, intense competition and are vulnerable

[1] Around 8.7 million people work directly in the industry itself and other jobs are provided by the economic activity that air travel supports (ATAG (2014)).

to exogenous events such as oil-price-shocks, accidents, infectious diseases or terrorism.

Worldwide airline industry	**2013**	**2014**	**2015**
Net post-tax profits, $billion	10.6	16.4	29.3
Net profit margin, % revenues	1.5	2.2	4.0
Net profit margin, $ per passenger	3.37	4.94	8.27

Table 1.1: Profits increase but margin remains small

To enhance their net profit margin or to simply remain competitive and viable in this challenging environment, airlines have to focus on both cost reduction and revenue growth through a better understanding of demand. Indeed, with few exceptions, history has shown that the most successful airlines are those with the strictest cost controls and the deepest knowledge of individual customers' preferences.

This dissertation has been motivated by current research problems faced by the airline industry and focuses on the development of models and algorithms that address these industry-related problems. Specifically, it contains three papers, each written in journal format. The first paper presents a mathematical model to assist aircraft cargo load planning. The objective is to minimize fuel and handling operation costs. The second and third papers focus on estimating demand for itinerary choice models. A particular emphasis is placed on including price in these models, correcting for price endogeneity, and assessing the robustness of price estimates to different utility function specifications and discrete choice model structures.

Therefore, the rest of this chapter will present more precisely each of the two main themes of this dissertation:

- A reduction of fuel and handling operation costs via the determination of optimal loading plans for a set of containers and pallets transported into a cargo aircraft that serves multiple airports.

- A deeper understanding of customers' itinerary choice preferences via the estimation of discrete choice models of demand that account for price endogeneity.

1.2 Airline container loading problem

1.2.1 Problem definition

The air freight industry has received less attention in the literature than the air passenger industry. Yet, air cargo transportation plays a highly significant economic role. In 2014, airlines transported 51.3 million metric tons of goods, representing more than 35% of global trade by value. This corresponds to $18.6 billion worth of goods every day. Air cargo represents thus a significant revenue source for airlines, generating on average 9% of airline revenues. This is more than twice the revenues from the first class passenger segment (IATA (2016a)).

Unfortunately, these high revenues do not automatically translate into high profits. Air freight carriers have high operating expenses. An increase in operating costs can lead to an increase in airfreight rates, which would encourage some shippers to divert freight to slower but less expensive modes of transport. Therefore, air freight carriers continually monitor costs and pursue strategies for reducing operating expenses.

The challenge for any airline (both passenger and cargo airlines) is to find new ways to cut costs but not at the expense of safety, reputation, or customer satisfaction. One of the most evident ways for an airline to save money is through fuel economy. As shown in Table 1.2, despite an important reduction of fuel spending in 2015, fuel costs remain the primary operating cost for airlines, accounting for more than a quarter of their operating expenses (28%) (IATA (2016b)).

Worldwide airline industry	**2013**	**2014**	**2015**
Fuel spend, $billion	228	226	191
% change YOY*	0.5	-1.1	-15.6
$ operating costs	33.0	32.3	28.1
Labour costs, USD $billion	135	142	150

Table 1.2: Annual fuel costs (IATA (2016b))
*YOY = Year-over-year

Reducing fuel consumption is not only cost-effective but also important in terms of air pollution emissions. The aviation industry contributes 2% of all human-induced carbon dioxide (CO_2) emissions (ATAG (2013)). According to the International Air Transport Association (IATA), airlines are collectively committed to reduce their fuel consumption and CO_2 emission by at least 25% of their 2005 levels by 2020.

The majority - but not all - of these fuel savings will come through the introduction of new technologies on existing or new aircraft. For example, winglets improve aerodynamics, composite materials reduce the weight of the aircraft, and new combustion technologies improve engine efficiency. All of these new technologies reduce fuel consumption. As a consequence, the next generation of aircraft will produce at least 70% less CO_2 emissions than aircraft manufactured 40 years ago (IATA (2016c)). Additional reductions in fuel consumption can be achieved through improvements in ground operations. This dissertation shows how cargo loading affects fuel consumption and how - by optimally loading cargo into an aircraft - fuel savings can be realized.

During the pre-loading activities, plans are developed that determine the placement of each container or pallet within the aircraft. The positioning of cargo on board is also referred to as the "Aircraft Weight and Balance Problem" because any aircraft is designed and certified to operate within certain weight and balance limitations. The aircraft's balance refers to the location of the aircraft's center of gravity (CG). The position of the center of gravity along the longitudinal axis is the primary concern. To ensure the aircraft is stable and safe to fly, the aircraft's manufacturer established a safe zone in which the longitudinal CG should be located for flight. The extremities of the range are called, respectively, the forward and aft limits. In essence the forward CG limit ensures that the aircraft always remains controllable, especially during the landing, whereas the aft CG limit is established to ensure the longitudinal stability. The aircraft's manufacturer also determines the weight limitations. The maximum allowable weight in the aircraft is stipulated as well as other weight limitations that are related to the distribution of the weight within the aircraft. These limitations are primarily established to ensure that the structural integrity of the aircraft is not compromised.

All these constraints make the elaboration of loading plans a complex and highly specialized task. Today, this activity remains a highly manual process and consequently there are potential efficiency gains that are possible. For example, the position of the CG affects the fuel consumption. More precisely, it can be shown that as the CG is moved backward, the aircraft becomes more efficient and burns less fuel. Therefore, the challenge is to find the optimal placement of cargo aboard so that the location of CG induces minimal fuel consumption.

Several publications in the literature have proposed mathematical models and algorithms to automatically solve the Aircraft Loading Problem. All the previous papers share one common assumption: they assume that the cargo aircraft has one unique destination. However, given revenue is determined by the amount of cargo carried, most airlines prefer to load as many containers as possible, even if all the loaded containers do not have the same final destination and some containers need to make one or more stops. According to Birtchnell et al. (2015), there is evidence that in order to reduce costs, cargo airlines are pursuing the concept first pioneered by major container shipping lines by replacing direct air cargo connectivity by more circuitous routing with multiple stops. In this case, frequency and speed appear to have been replaced by connectivity.

When multiple destinations are considered, pickup and delivery operations occur at intermediate airports and the loading problem becomes not only a Weight and Balance Problem but also a Sequencing Problem. As noted earlier, the load planning activity is typically done manually - or at best, with ad hoc analysis tools. As a rule, the person in charge of the development of the loading plans do not incorporate pickup and delivery operations in the planning process. As a consequence, containers that are at an intermediate airport may be unloaded and reloaded in order to access containers behind them that need to be unloaded. These unproductive shifts increase the ground time of the aircraft and increase the handling costs. In order to save time and money, airlines should take into account the pickup and delivery operations and try to avoid the unnecessary unloading operations when creating their loading plans.

In Chapter 2, we propose an exact solution approach that relies on a mixed integer linear program to find the optimal assignment of containers and pallets on board of an aircraft that serves multiple destinations.

1.2.2 Main contributions

There are four main contributions this dissertation makes relative to the Airline Container Loading Problem.

The first and most valuable contribution arises from the recognition that cargo aircraft can visit multiple destinations. To the best of our knowledge, our formulation of the Airline Container Loading Problem is the first to model cargo transport as a series of trips that consist of one or more legs with pickups and deliveries. As a consequence, two objectives are pursued concurrently: minimizing fuel consumption and minimizing handling operations. The originality of our approach is to express these two objectives in monetary terms. Our objective function measures the amount of extra costs induced by the loading plan.

The consideration of a series of trips also compels us to model the handling operations on the ground. Our second contribution is to adapt the basic constraints of the Weight and Balance Problem to the multi-leg trip. Our model is based on international standards and is valid for most commercial operators.

Our third contribution is to provide evidence that the Airline Container Loading Problem is NP-complete.

Finally, we show results based on real data from a cargo airline. These results show our approach yields better solutions, which results in substantial cost savings with respect to current typical practices.

1.3 Itinerary choice models

1.3.1 Problem definition

In any organization, a deep understanding of demand is critical to the success of that organization, in the sense that uncertainty of demand can lead to ineffective or sub-optimal decisions. Airlines are no exception. With the deregulation of the industry (which occurred first in the United States (U.S.) in 1978), airlines were thrown into a new market-driven industry, one in which they could decide where and when to fly, as well as what fares to charge. Soon after the deregulation, airlines realized the importance of integrating computerized decision support systems into areas such as scheduling, pricing, revenue management, marketing, sales, and operations. At that time computational power was much more limited than it is today and it was unthinkable for airlines to model air travel demand using disaggregate data on individual passengers.

During the 1990s and 2000s, airlines have experienced profound transformations. With the advent of the Internet, consumers were suddenly able to search themselves for information about different products and prices that were available and it was much easier for them to compare price and product offerings across different competitors. The Internet has made pricing more transparent, and increased competition among carriers. Increase pricing competition has also come from the growth of low cost carriers (LCCs).

In addition to the explosive growth of online booking channels and increased competition from low cost carriers, airlines were confronted with a series of new challenges in the first decade of the 21st century, including among others high fuel costs, terrorism threats, disease outbreaks (SARS, H1N1, Ebola, etc.), and the financial and economic crises after 9/11 and in 2007-2008. These factors have led to a new environment for airlines, one in which the computerized decision support systems developed shortly after deregulation did not function as well.

Consequently, researchers and airline practitioners began to investigate how they could develop new approaches for revenue management and network planning applications that are based on discrete choice methods. Both revenue management and network planning models forecast air passenger demand, but differ in the fundamental decisions they support. Revenue management models are used to determine how many seats to offer for sale on a given route and at a given price given current and predicted itinerary-level forecasts. In this sense, revenue management supports short-term, tactical decisions based on current market conditions. Conversely, network planning models are primarily used to support mid- to long-term decisions, such as how many aircraft to purchase, where to add new routes, where to add code share partners, etc.

Given itinerary choice models support many critical mid- to long-term decisions, the transportation community has invested significant resources in developing new discrete choice models. These new models, which belong to the Generalized Extreme Value (GEV) class, have improved forecasting accuracy by incorporating complex substitution patterns. However, these models suffer from major limitations.

One of these limitations is that they generally use average fares and do not correct for price endogeneity. Endogeneity occurs when correlation exists between an explanatory variable and the error term (or unobserved factors) in a model. In demand models, price endogeneity occurs when prices and demand are positively correlated. Left unaddressed, price endogeneity usually results in a biased (and often positive) price coefficient. One of the main causes of price endogeneity is referred to as simultaneity of supply and demand, which occurs when price influences demand, and demand influences price. When demand is anticipated to be low for certain flights, airlines react and offer lower prices on those flights with lower demand. This leads to endogeneity in the price-demand relationship: high prices can be associated with high demand and low prices with low demand. This problem is most often encountered when average prices are used but may also arise with disaggregate data when unobserved attributes of the product that affect demand are correlated with price. In the airline context, an example could be a passenger who paid a higher price for priority boarding privileges or for a seat with extra leg room because he is tall.

Endogeneity is an important issue in price-sensitivity estimation and recent research in the transportation literature has focused attention on the importance of accounting for endogeneity in demand studies. For example, Guevara (2015) notes that "endogeneity often arises in discrete-choice models, precluding the consistent estimation of the model parameters, but is habitually neglected in practical applications". In the airline literature, Mumbower et al. (2014) and Hotle et al. (2015) find strong evidence of price endogeneity and demonstrate that left unaddressed, endogeneity results in inconsistent estimators of the parameters.

However, these studies use linear regression models and, to the best of our knowledge, there is no prior research showing how to correct for price endogeneity for an itinerary choice model that is consistent with those used by industry. Yet, endogeneity also arises in discrete-choice models and leads to the same biases.

As explained by Fernandez-Antolin et al. (2014), less has been done regarding endogeneity in discrete choice models than in linear regression models and this is particularly true for transportation applications. The Control Function (CF) method is the current standard for addressing endogeneity in discrete choice models. This method can be thought of as a two stage procedure, and is the one used in this dissertation to correct for price endogeneity. For a complete description of methods used to correct for endogeneity in discrete choice models, the reader is referred to Guevara (2015).

Another limitation of almost all prior research on air travel demand estimation is that they have been based on the U.S. DOT's Origin-Destination Survey Databanks (hereafter referred to as DB1B), which represent a 10% sample of tickets sold by domestic U.S. airlines. Although these data provide information about the exact fare paid for a ticket, no information is provided about when an individual paid for the ticket, what specific flights were purchased, or when in a quarter the individual traveled. That is, only the distribution of fares at the routing level is known (e.g., fares paid to fly United routes operating from Atlanta to Seattle and connecting in Denver during the Spring 2014 quarter). Alternate sources of ticketing data exist that are able to overcome these limitations, but these sources have been unavailable to the research community.

The Airlines Reporting Corporation (ARC) is a ticketing clearinghouse that maintains financial transactions for all tickets purchased through travel agencies worldwide. This includes both online (e.g., Expedia) and brick-and-mortar agencies. Starting in 2013, multiple U.S. carriers started providing ARC with all of their tickets, including direct sales through their websites. ARC has detailed information associated with each ticket. This includes the price paid for the ticket (and associated taxes and currency), ticketing date, "high yield" and "low yield" product classifications, and detailed information about each flight associated with the ticket, e.g., departure and arrival dates/times; origin, destination, and connecting airports; total travel time; connecting times; flight numbers; equipment types and associated capacities; and operating and marketing carriers.

In Chapter 3 and 4 of this dissertation we use ARC data to formulate itinerary choice models that are consistent with those used by industry and correct for price endogeneity using a control function that uses several types of instrumental variables.

1.3.2 Main contributions

There are three main contributions related to the estimation of itinerary choice models.

Our first contribution is to estimate a baseline multinomial logit (MNL) model that controls for price endogeneity for high-yield and low-yield fare products using the control-function method. Our results show that failure to account for price endogeneity leads to over-estimation of customers' value of time, which can lead to sub-optimal business decisions.

Our second main contribution is that it is the first study to estimate highly refined departure time of day preferences. Due to the size of our analysis database, we are able to allow departure time preferences to vary according to several dimensions including the length of haul, direction of travel, number of time zones crossed, departure day of week, and itinerary type (i.e., outbound, inbound and one-way itineraries). The price elasticity and departure time of day preferences results are not restricted to itinerary choice modeling applications, and can help support evaluation of proposed airport fees and taxes, national departure and emission taxes, landing fees, and congestion pricing policies. To the best of our knowledge, these curves represent the most refined publicly-available estimates of airline passengers' time of day preferences.

Our third contribution tests the sensitivity of results from our baseline MNL to different modeling assumptions. First, we test the sensitivity of results to different time-of-day formulations by comparing three departure time formulations: one that uses a set of discrete indicator variables for each departure hour and two that use a series of sine and cosine functions to model departure times as continuous functions.

Next, we test the sensitivity of results to different price formulations by comparing models that include fare versus ln(fare) and incorporating different price sensitivities by advance purchase periods. Results show that using fare provided more intuitive value of time estimates than ln(fare) and that price sensitivities vary as a function of advance purchase periods, with those purchasing high-yield products being less price sensitive than those purchasing low-yield products (across any advance purchase periods) and those purchasing closer to departure being less price sensitive.

Finally, we estimate advanced GEV models that include 2-level nested logit (NL) models, 3-level NL models, and ordered generalized extreme value (OGEV) models. We compare our results to those reported in the literature for United Airlines which were estimated using 2000 data. Our results, based on 2013 data, are surprisingly similar to those obtained from prior researchers. Based on our estimations, we recommend using either a three-level NL model that includes upper-level nests for time of day and lower-level nests for carriers or (potentially) an OGEV model with hourly time periods in which each itinerary is allocated equally between three nests.

The results from our estimated itinerary choice models form the first step in my proposed post-doctoral research, described in the conclusions chapter, which is to further enhance these itinerary choice models by incorporating consumer characteristics and different choice set generation rules and using these models to estimate consumer welfare benefits associated with recent mergers and acquisitions.

1.4 Dissertation organization

The remainder of this dissertation contains four chapters. Chapter 2 focuses on the cargo loading optimization problem while Chapters 3 and 4 present the results of itinerary choice models that correct for endogeneity. Chapter 3 contains three appendices. The first describes the process used to generate itinerary choice sets. The second contains results for all time of day preference estimates. The third compares three departure time of day formulations. Finally, Chapter 5 summarizes overall conclusions and contributions and discusses directions for future research.

Bibliography

Air Transport Action Group (2014). Aviation benefits beyond borders. http://aviationbenefits.org/media/26786/ATAG__AviationBenefits2014_FULL_LowRes.pdf. [Online; accessed 22-February-2016].

Air Transport Action Group (2013). Facts and figures. http://www.atag.org/facts-and-figures.html. [Online; accessed 22-February-2016].

Birtchnell, T., S. Savitzky, and J. Urry (2015). *Cargomobilities: moving materials in a global age*. Routledge.

Fernandez-Antolin, A., A. Stathopoulos, and M. Bierlaire (2014). Exploratory analysis of endogeneity in discrete choice models. In *14th Swiss Transport Research Conference* (No. EPFL-CONF-202581).

Guevara, C. A. (2015). Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research Part A: Policy and Practice 82*, 240–254.

Hotle, S. L., M. Castillo, L. A. Garrow, and M. J. Higgins (2015). The impact of advance purchase deadlines on airline consumers' search and purchase behaviors. *Transportation Research Part A: Policy and Practice 82*, 1–16.

International Air Transport Association (2016a). Air cargo - enabling global trade. https://www.iata.org/whatwedo/cargo/Pages/index.aspx. [Online; accessed 22-February-2016].

International Air Transport Association (2016b). Economic performance of the airline industry. https://www.iata.org/whatwedo/Documents/economics/IATA-Economic-Performance-of-the-Industry-end-year-2015-report.pdf. [Online; accessed 22-February-2016].

International Air Transport Association (2016c). Operational fuel efficiency. https://www.iata.org/whatwedo/ops-infra/Pages/fuel-efficiency.aspx. [Online; accessed 22-February-2016].

Mumbower, S., L. A. Garrow, and M. J. Higgins (2014). Estimating flight-level price elasticities using online airline data: A first step toward integrating pricing, demand, and revenue optimization. *Transportation Research Part A: Policy and Practice 66*, 196–212.

Part I

Cargo Loading Problem

Chapter 2

The Airline Container Loading Problem with pickup and delivery

Lurkin, V. and Schyns, M. (2015). The airline container loading problem with pickup and delivery. *European Journal of Operational Research. 244(3)*, 955-965.

This chapter considers the loading optimization problem for a set of containers and pallets transported into a cargo aircraft that serves multiple airports. Because of pickup and delivery operations that occur at intermediate airports, this problem is simultaneously a Weight and Balance Problem and a Sequencing Problem. Our objective is to minimize fuel and handling operation costs. This problem is shown to be NP-hard. We resort to a mixed integer linear program. Based on real-world data from a professional partner (TNT Airways), we perform numerical experiments using a standard branch-and-cut (B&C) library. This approach yields better solutions than traditional manual planning, which results in substantial cost savings.

2.1 Introduction

In the *Airline Container Loading Problem with Pickup and Delivery* (ACLPPD), a set of containers and pallets, known as *Unit Load Devices* (ULD), must be loaded into a compartmentalized cargo aircraft. We consider that pickup & delivery operations occur at different airports during any given trip. The loading task is illustrated in Figure 2.1. We propose an exact solution approach that relies on a mixed integer linear program to find the optimal ULD assignment.

Figure 2.1: ULD loaded through main deck side cargo door (left), through nose door (middle), and through lower deck side cargo door (right)

Air cargo represents less than 1% of the world trade volume, but its value is in excess of $6.8 trillion per annum, which is over one third of the world trade value (IATA (2016)). Thus, air cargo transportation plays a highly significant economic role. Optimizing loading assignment on board is critical to airlines for several reasons.

First, correct loading conditions safety. Inappropriate loading can cause significant damage, and place the aircraft, the freight or even the crew at risk. Therefore, this study models a wide set of constraints for operators to consider daily. The proposed model applies to all aircraft and loads, and complies with international standards. Considering the same constraints as Limbourg et al. (2012), we adapt such constraints to the case of a sequence of routes, called *legs*, while considering the additional case of hazardous products and oversized ULDs.

Second, optimal loading has a positive impact on aerodynamics, thus resulting in less fuel consumption, i.e., reduced cost and environmental impact. This issue is crucial for airlines, affected by rising oil prices and increased pressure to reduce carbon dioxide emissions. This research analyzes fuel and handling operations in order to minimize costs. The management of these first two requirements is done through a proper distribution of the ULD weights within the aircraft. This part is a Weight & Balance Problem.

Third, optimal loading is important for airlines because managing operations on the ground is challenging, especially when the trip includes several legs with P&D operations. Reducing the number of handling operations reduces time, which in turn reduces labor costs per flight. Such reduction also allows shorter turnaround time, i.e., the time that elapses from the moment the plane arrives to the moment it leaves again, thus reducing airport fees.

Time saved could be used for other valuable operations.

Optimizing loading plans is also crucial and constitutes another reason to consider this problem. Indeed, loadmasters must build plans within an extremely short time, and doing so manually requires significant time. On the other hand, with an interactive computerized efficient tool, loadmasters would be able to consider different alternatives and select the best solution with respect to their experience and the real conditions faced on the ground.

In this context, the problem no longer consists merely, as in Limbourg et al. (2012), in positioning ULDs to reach a proper equilibrium, but also in defining the unloading and loading operation sequence at airports. Because there is only one path between any ULD and the exit door, this path must be free to unload ULDs. The task is to minimize, at each airport, the number of ULDs in transit to be unloaded in order to have access to the ULDs reaching their delivery point. The same question arises when pickup occurs. The problem is even more complex when several doors can be used, as occurs occasionally. The cost of these handling operations is the second element of our proposed objective function. It is important to notice that we face two conflicting objectives: optimizing board assignments for fuel and for ground operations. Our contribution is to propose an exact approach to solve simultaneously both the Weight & Balance Problem over a multi-leg trip, and the sequencing problems associated to pickups and deliveries. We resort to a mixed integer linear program where the objective is to minimize both costs.

Currently, this extremely complex problem (NP-hard) is still essentially solved manually based on best practices. Because load planners have extremely short time windows to choose assignments, they focus mainly on finding a feasible and reasonable solution. As a rule, they do not incorporate P&D operations in the planning process. A common method for managing several legs is, indeed, to plan each leg independently. Accordingly, almost the entire cargo can be unloaded at intermediate airports, and the ULDs that have not reached final destinations are reloaded subsequently, which is the worst possible scenario for ground operations. We show, based on of our first results from real data provided by industrial partners, that our approach allows significant savings.

The remainder of this chapter is organized as follows. Section 2.2 outlines the problem and the assumptions involved. Related literature and contributions are presented in Section 2.3. Section 2.4 describes the problem in more detail, and provides the mathematical formulation corresponding to the proposed model. Section 2.5 provides information on the theoretical complexity of the problem, whereas Section 2.6 illustrates the performance of the approach through numerical results. Finally, some conclusions are drawn in Section 2.7.

2.2 Problem summary and assumptions

The goal of the research is the development of a mixed integer linear program for the optimal loading of a set of unit load devices (ULDs) into a single cargo aircraft that has to serve multiple destinations under some safety, structural, economical, environmental and maneuverability constraints. As explained by Limbourg et al. (2012), a ULD is an assembly of components consisting of a container or of a pallet covered with a net, whose purpose is to provide standardized size units for individual pieces of baggage or cargo, and to allow for rapid loading and unloading.

We make the following main assumptions. A cargo aircraft has to deliver goods to several airports. The aircraft stops at one or more intermediary airports before reaching the final arrival airport. The flight plan is presumably known in advance, which means that the airports and the order in which they will be visited are known. We use the term "trip" to designate the entire journey of the aircraft, from the first departure airport to the final arrival airport. A trip is composed of multiple legs. A leg is a portion of a trip, from take-off to landing. The first departure airport is called the "depot".

We also know all the containers and pallets (ULDs) to be delivered[1]. For each ULD, we know its size, shape, weight, and respective origin and destination. We follow international standards for ULD description. Full details on the coding standards can be found in the ULD Regulations (ULDR) of the International Air Transport Association (IATA (2016a)). A cargo aircraft generally contains multiple decks with multiple position configurations for each deck. A position is simply a particular aircraft space that accommodates exactly one ULD. The location of each position and that of all ULDs that fit into such position are also known. The location of the different doors is also given. A cargo aircraft has generally one side cargo door on the main

[1]In practice, at the time of the elaboration of the loading plans, the loadmaster has often only a rough idea of what will be loaded at the intermediate airports. However, based on past flights, they can approximate the load that will be delivered between two airports.

deck and one for each of the three compartments of the lower deck. In addition, a nose door is sometimes available for the main deck. An example of cargo aircraft structure is illustrated in Figure 2.2. The focus of this research is on cargo transportation. Passenger transportation and the transportation of goods in the lower deck of passenger aircraft are beyond the scope of this research. However, the central ideas remain the same and extensions of our approach can be considered.

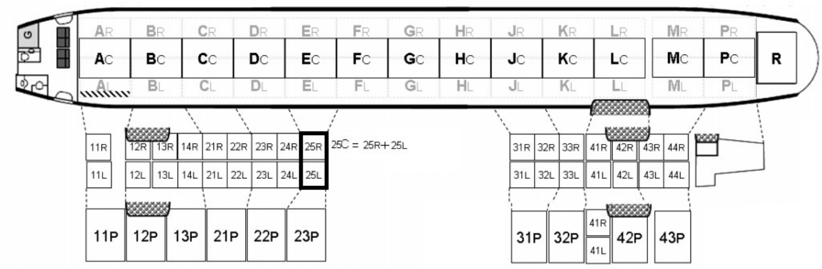


Figure 2.2: Illustration of aircraft structure

The model can be informally summarized as:

min	Fuel and loading operations costs on the entire trip
s.t.	Pickup & delivery sequences are feasible
	Customer demand is satisfied (each ULD is loaded)
	Each ULD fits in an aircraft position
	A position accepts only one ULD
	Some positions overlap and cannot be used simultaneously
	Longitudinal stability is within certified limits (ZFW, TOW, LW)
	Lateral stability is within certified limits
	Weight per position is below certified limit
	Combined weight load limits are set
	Cumulative weight load limits are set
	Regulations for hazardous goods are fulfilled
	Oversized ULDs are managed

The decision variables represent the location of each ULD within the aircraft. The function objective and the constraints will be described in detail in sections 2.4.2 and 2.4.3, respectively.

2.3 Related literature and contributions

This problem is an Assignment Problem (AP) that is referred in the literature as belonging to the family of Weight & Balance Problems. Over the past years, more attention has been paid to the problem that precedes ACLPPD by considering how to optimize freight loading within ULDs (Chan et al. (2006), Yan et al. (2008), Li et al. (2009), Wu (2010), Tang and Chang (2010), Tang (2011), Paquay et al. (2016)) independently of aircrafts. The scientific literature on aircraft cargo load planning is not extensive, but contains a number of papers.

Prior publications differ in many ways. First, the literature can be subdivided according to two approaches: bin packing and assignment. In the Bin Packing Problem (BPP) approach (for examples, see Amiouny et al. (1992), Heidelberg et al. (1998), Guèret et al. (2003), Nance et al. (2011)), the authors attempted to fill the aircraft continuously by excluding empty spaces between the items, whereas in the Assignment Problem (AP) approach (for examples, see Larsen and Mikkelsen (1980), Mongeau and Bès (2003), Limbourg et al. (2012), Vancroonenburg et al. (2014)), the authors attempted to allocate ULDs into predefined standardized positions, similar to our approach. Second, several papers treated the problem with exact approaches (for examples, see Mongeau and Bès (2003), Guèret et al. (2003), Limbourg et al. (2012), Vancroonenburg et al. (2014)), whereas others developed heuristics (for examples, see Larsen and Mikkelsen (1980), Amiouny et al. (1992), Heidelberg et al. (1998), Nance et al. (2011)). Third, some papers attempted to determine how to select the ULDs or items to be loaded in an aircraft or a fleet of aircraft (for examples see Mongeau and Bès (2003), Fok and Chun (2004), Vancroonenburg et al. (2014)), whereas others assumed that all ULDs must be loaded in the aircraft, similar to us (for examples, see Larsen and Mikkelsen (1980), Amiouny et al. (1992), Limbourg et al. (2012)). Fourth, some papers considered military applications (for examples, see Ng (1992), Heidelberg et al. (1998), Guèret et al. (2003), Kaluzny and Shaw (2009), Nance et al. (2011)), whereas others were interested in commercial applications (for examples, see Larsen and Mikkelsen (1980), Mongeau and Bès (2003), Fok and Chun (2004), Limbourg et al. (2012), Vancroonenburg et al. (2014)). Finally, among these prior works, only Larsen and Mikkelsen (1980) and Vancroonenburg et al. (2014) were interested in multi-leg flights. In Larsen and Mikkelsen (1980), the objective was, like us, to minimize the number of unloading and reloading operations at intermediate airports. However, no details were provided for how they achieved this goal. Vancroonenburg et al. (2014) proposed to assign priorities to ULDs based on their destination. However, these classifications are not exhaustive, and some papers fall within two categories. This literature also varies on at least three other dimensions: the precise definition of the objective function, the nature of the shipments, and the constraints considered.

Related problems consider container loading in trucks, trains, or ships. All such problems are combinatorial optimization problems that have to satisfy several, and sometimes similar, loading constraints. For these specific carriers, the multiple destinations case is often considered in the literature. The minimization of handling operations is then a common objective to achieve (for examples, see Altarazi (2013), Bostel and Dejax (1998), Imai et al. (2006), Øvstebø et al. (2011)). However, optimal loading for each of these carriers is a problem different from aircraft loading. In a truck, there are no predefined positions for the containers. Therefore, the Truck Loading Problem (TLP) is not AP, but it fits under the more general Container Loading Problem (CLP) (Bortfeldt and Wäscher (2013)) that includes the BPP. In a train or ship, available locations for containers are also predefined, but the carrier structure is still significantly different from that of an aircraft. Consequently, the weight constraints and loading operations are extremely specific, which makes ACLPPD peculiar and basically different from these two loading problems. Moreover, especially in the case of ships, the number of containers to be loaded leads to problems that are too large to be solved by exact methods. The literature consists mainly of heuristics, and does not present exact mathematical models for real problems.

The studies of Mongeau and Bès (2003), Vancroonenburg et al. (2014), and especially Limbourg et al. (2012), from which we started, relate most closely to our work. Indeed, these three papers consider commercial cargo aircraft with predefined positions and standardized ULDs, use exact methods, and consider the aircraft's Center of Gravity (CG). Nonetheless, we depart significantly from these works. Our contributions are multiple. The main contribution is the consideration of a trip that consists of several legs with pickups and deliveries that occur at intermediate airports. This compels us to model the handling operations on the ground, and to adapt the basic constraints of the Weight & Balance Problem to the multi-leg trip. We analyze the impact of loading operations in terms of cost. The mathematical model that we propose simultaneously considers two conflicting monetary objectives. To the best of our knowledge, this has never been done before. The integration of a larger set of realistic constraints is also a contribution.

In addition, we provide information regarding the complexity of the problem. Finally, we show performance measures based on real data.

2.4 Problem description and mathematical formulation

2.4.1 Main data and variables

Our model is built on three main sets of data.

The first is the set $\mathbb{L}$ of legs. The legs are the different parts of a trip. Each leg separates two successive airports. This model considers two legs. However, generalization to more legs is simple. A trip composed of two legs is a common case for long-range flights, whereas the consideration of more legs essentially complicates notation.

The second main set of data is the set $\mathbb{U}$ of ULDs. For each ULD i, we know its type (IATA code), weight w_i, and airports of origin and destination. We assume that the weight is uniformly distributed within the ULD. We need this assumption in order to presume that the center of gravity (CG) of the loaded ULD is at the center of the position occupied. This is important to calculate the loaded aircraft CG or to construct certain weight constraints. As explained by Paquay et al. (2016), this assumption is consistent with the control and loading manuals of airlines that stipulate that the CG of each ULD should lie in a certified area around the geometrical ULD center. By knowing the origin and destination of each ULD, three distinct ULD subsets can be established : $\mathbb{U}_1$ for ULDs loaded at the depot and unloaded at the first destination, $\mathbb{U}_2$ for those loaded at the first destination and unloaded at the second one, $\mathbb{U}_3$ for those loaded at the depot and unloaded at the second destination. For notation ease, we also define $\mathbb{U}_k^L$ as the subset of ULDs present in the aircraft for leg $k \in \mathbb{L}$. By definition, the intersection of $\mathbb{U}_1^L$ and $\mathbb{U}_2^L$ is $\mathbb{U}_3$.

The last set of data is the set $\mathbb{P}$ that defines available positions, which are the predefined spaces in the aircraft that can accommodate ULDs. The longitudinal location of a position j, denoted a_j, is called the arm and is expressed in inches from a virtual point called "datum". The datum allows accurate, and uniform, measurements to any point on the aircraft. As illustrated on Figure 2.4, the datum is generally arbitrarily set in front of the aircraft nose. Finally, each position is located on a specific deck and can be situated on the left side (L), the right side (R), or cover the entire width (C). To identify all positions on either side of the center, we use the subsets $\mathbb{P}_R$ and $\mathbb{P}_L$, respectively.

Our main variables are the binary variables x_{ijk} defined as:

$$x_{ijk} = \begin{cases} 1 & \text{if ULD } i \text{ is in position } j \text{ during leg } k, \\ 0 & \text{otherwise} \end{cases}$$
$$\forall i \in \mathbb{U}, \forall j \in \mathbb{P}, \forall k \in \mathbb{L}.$$

2.4.2 Objective function

As stated in the previous section, the objective of ACLPPD is to assign each ULD to one position while minimizing total costs, which is realized by minimizing fuel consumption on each leg and handling operations at intermediate airports. More precisely, we minimize overcost.

Fuel costs

Given that CG location impacts fuel consumption, on each leg, CG should be located at its best position. Without detailing many technical considerations, we can state that aircraft fuel consumption depends on the results of a continuous battle of forces, namely, weight, lift, thrust, and drag (see Figure 2.3). A slightly aft CG should reduce drag, and ultimately, thrust. Less thrust means less fuel consumption. Because the main impact is caused by the CG longitudinal location, we focus only on this axis in the objective function. Lateral location is maintained essentially within feasible limits (see Section 2.4.3 and Constraint 2.16). Henceforth, CG designs only the longitudinal location.

CG location is restricted within a range of certified limits defined by aircraft manufacturers. These limits are crucially important because a CG value outside such limits can cause instability that produces harmful effects. Within this range, some freedom is allowed. Airlines and pilots are aware that an aft CG usually saves fuel. Nonetheless, airlines typically define the CG target location around the center of certified ranges. In our case, we define the optimal location (denoted OCG) to be close to the aft certified limit, less an additional security margin left to the operator's appreciation. This aft certified limit is established during certification flight testing and is

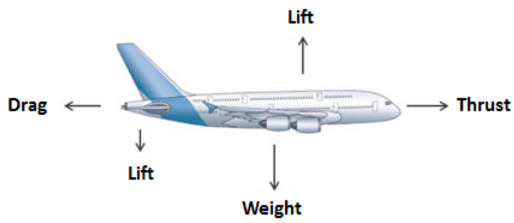


Figure 2.3: Forces acting on aircraft in flight

contained in the Pilot's Operating Handbook. The main purpose of the aft limit is to ensure controllability. We want to find loading plans for which the CG on each leg is located as close as possible to the optimal location. In mathematical terms, this gives:

$$\min \sum_{\forall k \in \mathbb{L}} \varepsilon_k$$

subject to:

$$CG_k - OCG_k - \varepsilon_k \leq 0 \qquad \forall k \in \mathbb{L} \qquad (2.1)$$

$$CG_k - OCG_k + \varepsilon_k \geq 0 \qquad \forall k \in \mathbb{L} \qquad (2.2)$$

where OCG_k is the optimal CG location and CG_k is the CG location obtained after ULD assignment. Both CG locations are expressed in inches from the datum. ε_k is the resulting deviation from the target for leg k. The datum, CG locations, and the forward and aft CG limits are illustrated in Figure 2.4.

Following the elementary principles of Aircraft Flight Manuals (AFM), the CG of a loaded aircraft, i.e., CG_k, is found by dividing the total balance moment of the aircraft by the total weight. The total balance moment of an aircraft is the algebraic sum of the balance moments from all its components (the empty aircraft, crew, ULDs, etc.). The balance moment of a component is determined by multiplying the weight of the component ($weight_k^c$) by its longitudinal distance ($distance_k^c$), in inches, to the datum. Thus, the CG position, in inches from the datum, is expressed by the following expression:

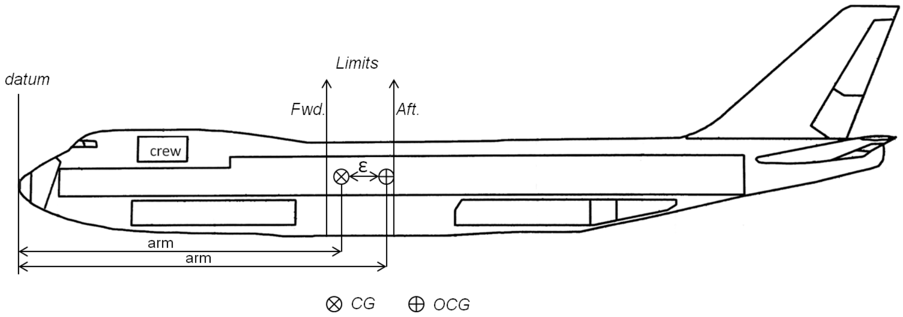


Figure 2.4: Datum, arm, CG locations and CG limits

$$CG_k = \frac{\sum_{c \in \mathbb{C}} weight_k^c \times distance_k^c}{\sum_{c \in \mathbb{C}} weight_k^c}$$

where $\mathbb{C}$ is the set of components.

One of these components is freight, i.e., all the ULDs to be loaded within the aircraft. The balance moment of a ULD is obtained by multiplying the weight of the ULD (w_i) by the arm (a_j), of its allocated position. Consequently, all ULDs to be loaded can be considered one unique large ULD with weight $weight_k^{freight} = \sum_{i \in \mathbb{U}_k^L} w_i$ and longitudinal distance ($distance_k^{freight}$) expressed as follows:

$$distance_k^{freight} = \frac{\sum_{i \in \mathbb{U}_k^L} \sum_{j \in \mathbb{P}} w_i \times a_j \times x_{ijk}}{\sum_{i \in \mathbb{U}_k^L} w_i}.$$

Note that the weight of the aircraft itself and the longitudinal distance to the datum of its unloaded CG are determined and provided by the manufacturer after physically weighing the aircraft. The location of the cabin crew and its weight is also known.

Handling operation costs

The second ACLPPD objective is to minimize the number of handling operations. All the ULDs from $\mathbb{U}_1$ (resp. $\mathbb{U}_2$) must be loaded at the origin of leg 1 (resp. 2) and unloaded at the destination of the same leg. Therefore, the number of ULD moves is equal to the number of ULDs in these sets and cannot be decreased. ULDs from $\mathbb{U}_3$ do not need to be unloaded a priori at the intermediate airport (optimal case). However, all these ULDs are typically unloaded (worst case) in order to unload the ULDs that belong to $\mathbb{U}_1$ and load those that belong to $\mathbb{U}_2$. This is because $\mathbb{U}_3$ ULDs could be situated in the path of those that must be loaded/unloaded at the intermediate airport, but also because the load plan must be adapted to the new leg while continuing to ensure a feasible and correct CG location for the next leg. If the initial load plan considers different legs, it should be possible to assign ULDs in $\mathbb{U}_3$ to positions not in the entry/exit path of other ULDs, which allows a suitable CG for the second leg. Therefore, we focus on the re-handling operations at the intermediate airport.

Moreover, an aircraft compartment can have more than one door. In this case, the number of operations can be minimized using the most accessible door. Without loss of generality, let us consider that each position can be reached from a maximum of two doors: the first in the direction of the nose (*the "nose door"*), and the second in the direction of the tail (*the "tail door"*). There can be more than two doors for the same deck, but only two are relevant for each position.

An example for the main deck of a Boeing 747 is illustrated in Figure 2.5. For position j , the *nose door* is situated exactly at the nose and the *tail door* is the lateral door. For position j', the *nose door* is the lateral door and there is no *tail door*.

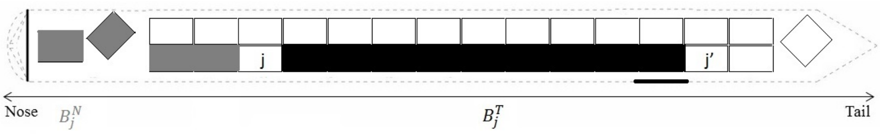


Figure 2.5: Different types of doors for Boeing 747

We can identify the set of positions in the entry/exit paths to/from position j as follows (see Figure 2.5):

- $\mathbb{B}_j^N$ is the set of all positions situated between position j and the first door in the direction of the nose. $\mathbb{B}_j^N$ is an empty set if position j is next to the nose door.

- $\mathbb{B}_j^T$ is the set of all positions situated between position j and the first door in the direction of the tail. $\mathbb{B}_j^T$ is an empty set if position j is next to the tail door.

In order to manage ULD unloading at intermediate airport, we introduce the binary variables α_j^N and α_j^T:

$$\alpha_j^N = \begin{cases} 1 & \text{if the ULD in position } j \text{ is unloaded through the } \textit{nose door,} \\ 0 & \text{otherwise.} \qquad \forall j \in \mathbb{P} \end{cases}$$

$$\alpha_j^T = \begin{cases} 1 & \text{if the ULD in position } j \text{ is unloaded through the } \textit{tail door,} \\ 0 & \text{otherwise.} \qquad \forall j \in \mathbb{P} \end{cases}$$

These variables are subject to the following constraints:

$$\alpha_j^N + \alpha_j^T \leq 1 \qquad \forall j \in \mathbb{P}, \tag{2.3}$$

$$\alpha_j^N + \alpha_j^T \geq \sum_{\forall i \in \mathbb{U}_1} x_{ij1} \qquad \forall j \in \mathbb{P}.\alpha_j^N \in 0,1 \tag{2.4}$$

Constraints (2.3) ensure that ULD unloading occurs through one door only. Constraints (2.4) guarantee that all ULDs whose final destination is the intermediate airport are unloaded well at the intermediate airport. Indeed, if there exists a ULD in $\mathbb{U}_1$ in position j ($x_{ij1} = 1$), there is necessarily an unloading from position j to either the nose door or the tail door (α_j^N or α_j^T must be equal to one). Note that Constraints (2.3) and (2.4) do not guarantee that α_j^N and α_j^T are equal to zero when a ULD does not need to be unloaded. However, the four next constraints, (2.5), (2.6), (2.7), and (2.8), in conjunction with the objective function, enforce a zero value for α_j^N and α_j^T when there is no unloading from position j. The unavailability of some doors for some positions is the subject of a specific constraint presented in section 2.4.3.2.

At the intermediate airport, we want to avoid ULD unloading in $\mathbb{U}_3$. Indeed, the final destination of these ULDs is not the intermediate airport, and these operations ought to be spared. We introduce binary variables n_j^N and n_j^T to identify unnecessarily unloading:

$$n_j^N = \begin{cases} 1 & \text{if the ULD in position } j \text{ is unnecessarily unloaded through the} \\ & \quad \text{nose door,} \\ 0 & \text{otherwise.} \qquad \forall j \in \mathbb{P} \end{cases}$$

$$n_j^T = \begin{cases} 1 & \text{if the ULD in position } j \text{ is unnecessarily unloaded through the} \\ & \quad \text{tail door,} \\ 0 & \text{otherwise.} \qquad \forall j \in \mathbb{P} \end{cases}$$

Constraints (2.5) and (2.6) stipulate that ULD $\in \mathbb{U}_3$ is unnecessarily unloaded at an intermediate airport if the ULD is in the exit path of a ULD that must be unloaded.

$$n_j^N \geq \alpha_{j'}^N + x_{ij1} - 1 \qquad \forall j \in \mathbb{P}, \quad \forall j' \in \mathbb{B}_j^T, \quad \forall i \in \mathbb{U}_3 \tag{2.5}$$
$$n_j^T \geq \alpha_{j'}^T + x_{ij1} - 1 \qquad \forall j \in \mathbb{P}, \quad \forall j' \in \mathbb{B}_j^N, \quad \forall i \in \mathbb{U}_3 \tag{2.6}$$
$$\alpha_j^N \geq n_j^N + x_{ij1} - 1 \qquad \forall j \in \mathbb{P}, \quad \forall i \in \mathbb{U}_3 \tag{2.7}$$
$$\alpha_j^T \geq n_j^T + x_{ij1} - 1 \qquad \forall j \in \mathbb{P}, \quad \forall i \in \mathbb{U}_3 \tag{2.8}$$
$$\tag{2.9}$$

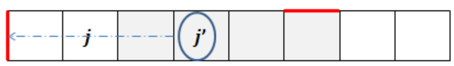


Figure 2.6: Unnecessarily unloading

As illustrated in Figure 2.6, if the ULD in position j' is unloaded through the nose door ($\alpha_{j'}^N = 1$), the ULD in position j is blocking the exit path and must also be unloaded through the nose door. If this ULD belongs to $\mathbb{U}_3$ ($x_{ij1} = 1$), it is an "unnecessary" operation. The right side of Constraints (2.5) is then strictly positive and n_j^N can only take the value one. Constraints (2.6) are equally applicable to unloading through the tail door. Given the objective function, n_j^N and n_j^N are pushed to zero when Constraints (2.5) and

(2.6) are not binding, i.e., either when there is no unloading from position j, or when the unloading from position j is necessary.

Because we want to minimize such unnecessary unloading, we obtain the objective function for the ground operations:

$$\min \sum_{\forall j \in \mathbb{P}} (n_j^N + n_j^T).$$

Finally, we want to consider both objectives together and we want to express them in monetary terms. We resort to the following multi-criteria objective function:

$$\min \sum_{\forall k \in \mathbb{L}} c_k^f \varepsilon_k + c^h \sum_{\forall j \in \mathbb{P}} (n_j^N + n_j^T). \tag{2.10}$$

where c_k^f is the monetary cost per unit deviation for fuel consumption and c^h is the cost coefficient for handling one ULD.

Both coefficients can be defined precisely when cost realities are known. For the first element, CG target location is typically expressed as a percentage of the aircraft Mean Aerodynamic Chord (MAC). MAC is defined by the airfoils and is the mean distance between the leading and trailing edges in the direction of airflow. For the Airbus 330, a CG displacement from its location of reference (28%) to a more aft CG (37%) could increase the number of air nautical miles per KG of fuel by 0.5% (Airbus Fuel Economy Material (2004)). This result can be converted in terms of fuel savings. The pilots and load planners consulted expect approximately 2.5% decrease of fuel consumption for a Boeing 777 when CG is displaced by a distance equivalent to 10% of the MAC to the aft. The exact relationship between CG location and fuel consumption depends on the aircraft type. These examples provide rough approximations for the amount of potential savings. Knowing the exact relationship between CG location and fuel savings for a specific aircraft and the absolute fuel volume required for a given trip with this optimal CG location, we can measure the amount of tons of fuel that must be added for one percent forward shift of the CG. By multiplying this figure by fuel price, we obtain the cost by unit deviation. If ε_k measures the deviation between the optimal CG location and the CG location obtained after ULD assignment and if c_k^f is the monetary cost per unit deviation for fuel consumption, then $\varepsilon_k c_k^f$ is the additional cost induced on leg k by an improper CG location (too forward). Because the first goal is to push CG to the aft, any positive value for c_k^f is effective. c_k^f can be interpreted as penalty coefficients. The main limitation of an approximate value is that it leads to an approximate value for the total cost reduction, but it has no impact on optimal assignment (at least if this value remains a reasonable approximation with respect to the second objective: cost minimization associated to loading operations).

The second cost in the objective function is associated to ground operations. The handlers assigned to this task need to operate quickly because the aircraft has a limited time window before the next departure. The total wages cost depends on the number of employees assigned to the task, which is directly proportional to the number of ULDs to be moved. Each company can measure it.

Note that if airlines and load planners prefer to select a sub-optimal CG location more within the range of values certified by the aircraft manufacturer, the target CG does not correspond anymore to the upper certified limit, and always pushing to the aft is not suitable. We decide to model it as a soft constraint. Any deviation of the CG to the aft beyond the target must also be penalized, and at least by a coefficient c_k^f, give priority to this constraint. Therefore, ε_k measures the absolute deviations with respect to the optimal CG. Alternatively, a hard constraint can be implemented easily by defining a lower maximum threshold in Constraint (2.15). When the optimal solutions exist, as is the case for most of our numerical experiments, both options are completely equivalent because the additional cost is null.

2.4.3 Constraints

This section presents the large number of constraints that characterize ACLPPD. We start with the basic constraints. More precisely, we show how we adapted the realistic weight and balance constraints presented in Limbourg et al. (2012) to the case of multiple legs. Next, we present the P&D constraints whose purpose is to consider the sequences of unloading and loading. We conclude the section with two other advanced constraints linked to hazardous and oversized products.

2.4.3.1 Adapting basic constraints

Standard assignment constraints

$$\sum_{j \in \mathbb{P}} x_{ijk} = 1 \qquad \forall i \in \mathbb{U}_k^L,\ \forall k \in \mathbb{L} \tag{2.11}$$

$$\sum_{i \in \mathbb{U}_k^L} x_{ijk} \leq 1 \qquad \forall j \in \mathbb{P},\ \forall k \in \mathbb{L} \tag{2.12}$$

Constraints (2.11) ensure that all ULDs are accommodated on board. In our problem, the commercial department of the airline company (TNT Airways) has previously negotiated contracts with clients, and the loadmaster must load all ULDs on board. The orders cannot be rejected by the optimization process at this time. Note that, as explained in Limbourg et al. (2012), there are often far fewer ULDs to load than positions in the aircraft. Constraints (2.12) state that, for a given leg, a position can accommodate one ULD only.

Possible positions for ULDs

ULDs are of many different sizes and shapes. For example, some ULDs might have a special shape to fit into the aircraft fuselage. Parameters γ_{ij} determine whether a particular ULD fits in a particular position:

$$\gamma_{ij} = \begin{cases} 1 & \text{if ULD } i \text{ fits in position } j, \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in \mathbb{U}, \forall j \in \mathbb{P}.$$

Constraints (2.13) ensure that each loaded ULD physically fits in its position. If not, the corresponding assignment is forbidden and the only allowed value for the assignment variable x_{ijk} is zero. These variables are removed during optimization.

$$x_{ijk} \leq \gamma_{ij} \qquad \forall i \in \mathbb{U},\ \forall j \in \mathbb{P},\ \forall k \in \mathbb{L} \tag{2.13}$$

For the same aircraft, different configurations of positions are possible. Some positions are larger and overlay several smaller positions (see Figure 2.2). For each larger position j, set E_j represents all smaller positions covered by j. If position j is used, all positions in set E_j must be discarded.

$$x_{ijk} + x_{i'j'k} \leq 1 \qquad \forall i, i' \in \mathbb{U}_k^L,\ \forall j \in \mathbb{P},\ \forall j' \in \mathbb{E}_j,\ \forall k \in \mathbb{L} \tag{2.14}$$

Stability constraints

$$minCG_k \leq CG_k \leq maxCG_k \qquad \forall k \in \mathbb{L} \tag{2.15}$$

$$-\bar{D} \leq \sum_{i \in \mathbb{U}_k^L} w_i \Big(\sum_{j \in \mathbb{P}_R} x_{ijk} - \sum_{j \in \mathbb{P}_L} x_{ijk}\Big) \leq \bar{D} \qquad \forall k \in \mathbb{L} \tag{2.16}$$

Aircrafts are designed and certified to operate within certain weight and balance limits. The prime concern of aircraft balancing is longitudinal balance, or the fore and aft location of the CG along the longitudinal axis. Constraints (2.15) ensure longitudinal stability by verifying that the aircraft CG on each leg is within the limits certified by the aircraft manufacturer. More precisely, the boundaries depend on aircraft weight under different scenarios: weight at departure with fuel (Take Off Weight), expected landing weight (Landing Weight), and total weight without fuel (Zero Fuel Weight). The three scenarios define three two-dimensional certified areas called feasibility envelopes. Using expected consumption and the fuel curve relationship between the three scenarios, one single lower bound $minCG$ and one upper bound $maxCG$ are determined[2]. The certified CG limits ($minCG$ and $maxCG$) are based on a certain set of design load requirements, and condition the stability and controllability of the aircraft. Consequently, exceeding these limits can be dangerous.

CG location with reference to the lateral axis is also important. Constraints (2.16) verify that the lateral balance is within reasonable limits. The difference between the weight allocated to either side of the fuselage centerline should not exceed $\bar{D}$. Laterally unbalanced conditions can cause adverse effects and lower efficiency. Indeed, the pilot might be forced to adjust the aileron trim tab or maintain constant aileron control pressure. Both measures cause more drag and hence more fuel consumption.

[2] A security margin can be added here as mentioned in the previous section.

Weight restrictions

$$w_i x_{ijk} \leq \bar{W}_j \qquad \forall i \in \mathbb{U}_k^L,\ \forall j \in \mathbb{P},\ \forall k \in \mathbb{L} \quad (2.17)$$

$$\sum_{i \in \mathbb{U}_k^L} \sum_{j \in \mathbb{P} | P_j \cap O_a^d \neq \emptyset} x_{ijk} o_{ija}^d \leq \bar{O}_a^d \qquad \forall d \in \mathbb{D},\ \forall a \in \mathbb{O}^d,\ \forall k \in \mathbb{L} \quad (2.18)$$

$$\sum_{i \in \mathbb{U}_k^L} \sum_{j \in \mathbb{P} | P_j \cap \bigcup_{c=1}^{a} F_c \neq \emptyset} \sum_{l=1}^{a} x_{ijk} f_{ijl} \leq \bar{F}_a \qquad \forall a \in \mathbb{F},\ \forall k \in \mathbb{L} \quad (2.19)$$

$$\sum_{i \in \mathbb{U}_k^L} \sum_{j \in \mathbb{P} | P_j \cap \bigcup_{c=1}^{a} T_c \neq \emptyset} \sum_{l=1}^{a} x_{ijk} t_{ijl} \leq \bar{T}_a \qquad \forall a \in \mathbb{F},\ \forall k \in \mathbb{L}. \quad (2.20)$$

For structural design reasons, different parts of the aircraft are subject to weight limits. In other words, placing excessive weight on given sections of the aircraft is not allowed. Constraints (2.17) ensure that the weight exerted on each position does not exceed the maximum weight allowed by the position (denoted by $\bar{W}_j$). There are also combined weight limits. Several compartments are defined on each deck, and on the entire aircraft. Constraints (2.18) ensure that the total weight in each compartment does not exceed the limit. $\mathbb{D}$ denotes the set of decks augmented by an artificial deck that corresponds to the entire aircraft. Let us imagine that the aircraft is cut in slices of one-inch width. There is a specific weight limit for each slice and each deck d in $\mathbb{D}$. Under some uniform distribution assumptions of the weight within the positions, the one-inch slices can be replaced by broader slices. O_a^d denotes the a^{th} slice for deck d, and $\bar{O}_a^d$ the maximum weight allowed for this area. Finally, o_{ija}^d represents the proportion of w_i in $\{P_j \cap O_a^d\}$. In the same way, there are weight limits on overlapping areas. With Constraints (2.19) (resp. Constraints (2.20)), we are interested in the cumulative weight from the nose (resp. the tail) to each of the previously defined slices. The consecutive forward and aft slices are denoted by F_a (forward) and T_a (aft). $\bar{F}_a$ (resp. $\bar{T}_a$) is the maximum cumulative weight allowed for the section that starts at the nose (resp. the tail) and ends at F_a (resp. T_a). Finally, the variables f_{ija} (resp. t_{ija}) represents the proportion of w_i in $\{P_j \cap F_a\}$ (respectively, $\{P_j \cap T_a\}$). See Limbourg et al. (2012) for details.

2.4.3.2 Pickup and delivery constraints

ULD categories

Because of their origin and destination, not all ULDs are present on each leg, which means that the corresponding x_{ijk} variables can be initialized to zero (and removed during the optimization). This is specified by Constraints (2.21).

$$x_{ijk} = 0 \qquad \forall i \notin \mathbb{U}_k^L,\ \forall j \in \mathbb{P} \tag{2.21}$$

Correct unloading sequence

Constraints (2.22) prevent the collision of ULDs unloaded through different doors. Constraints (2.22) state that the ULD in position j can only leave through the nose door ($\alpha_j^N = 1$) if there is no ULD between j and the nose door ($\mathbb{B}_j^N$) that is attempting to leave in the opposite direction, i.e., through the tail door (the sum is null). Note that Constraints (2.22) manage potential conflicting moves in both directions.

$$\alpha_{j'}^T \leq 1 - \alpha_j^N \qquad \forall j \in \mathbb{P},\quad \forall j' \in \mathbb{B}_j^N \tag{2.22}$$

No position exchange within aircraft

Any ULD assigned to a different position for the second leg must be unloaded first. Constraints (2.23) ensure that each ULD not unloaded at the intermediate destination, i.e., when $(n_j^N + n_j^T)$ is null, maintains the same position for the second leg ($x_{ij1} = x_{ij2}$).

$$x_{ij1} - x_{ij2} \leq (n_j^N + n_j^T) \qquad \forall j \in \mathbb{P},\quad \forall i \in \mathbb{U}_3 \tag{2.23}$$

Feasible loading sequence

Constraints (2.24), (2.25), and (2.26) ensure a feasible loading sequence of the ULDs that belong to $\mathbb{U}_2$ at the intermediate airport. These ULDs can only reach positions with a free path. We define binary variables β_j^N and β_j^T to ensure a feasible loading sequence:

$$\beta_j^N = \begin{cases} 1 & \text{if the path from the nose door to position } j \text{ is blocked,} \\ 0 & \text{otherwise.} \end{cases} \qquad \forall j \in \mathbb{P}$$

$$\beta_j^T = \begin{cases} 1 & \text{if the path from the tail door to position } j \text{ is blocked,} \\ 0 & \text{otherwise.} \end{cases} \qquad \forall j \in \mathbb{P}$$

Constraints (2.24) determine whether there is a free path between the nose door and position j. At the time of loading, the path can only be blocked by ULDs from $\mathbb{U}_3$ that were not unloaded ($n_{j'}^T + n_{j'}^N = 0$). If there is no free path, the left term of (2.24) is strictly positive and the new binary variable β_j^N can only obtain the value one. Similarly, Constraints (2.25) restrict β_j^T to one if the path through the tail door to position j is not free. If both paths are blocked, Constraints (2.26) state that ULD i cannot be assigned to position j.

$$x_{i'j'1} - n_{j'}^T - n_{j'}^N \leq \beta_j^N \qquad \forall j \in \mathbb{P}, \ \ \forall j' \in \mathbb{B}_j^N, \ \ \forall i' \in \mathbb{U}_3 \tag{2.24}$$

$$x_{i'j'1} - n_{j'}^T - n_{j'}^N \leq \beta_j^T \qquad \forall j \in \mathbb{P}, \ \ \forall j' \in \mathbb{B}_j^T, \ \ \forall i' \in \mathbb{U}_3 \tag{2.25}$$

$$\beta_j^T + \beta_j^N - 1 \leq (1 - x_{ij2}) \qquad \forall j \in \mathbb{P}, \ \ \forall i \in \mathbb{U}_2 \tag{2.26}$$

Inaccessibility of some doors

Different aircraft types have different door configurations. Binary parameters δ_j^N and δ_j^T are introduced to manage the availability of these doors:

$$\delta_j^N = \begin{cases} 1 & \text{if a nose door is available for pos. } j, \\ 0 & \text{otherwise} \end{cases} \qquad \forall j \in \mathbb{P}.$$

$$\delta_j^T = \begin{cases} 1 & \text{if a tail door is available for pos. } j, \\ 0 & \text{otherwise} \end{cases} \qquad \forall j \in \mathbb{P}.$$

Constraints (2.27) and (2.28) ensure that ULD unloading is not conducted through a door that does not exist. Indeed, if there is no nose door (resp. tail door) for position j, δ_j^N (resp. δ_j^T) is zero and α_j^N (resp. α_j^T) must be zero. Similarly, Constraints (2.29) and (2.30) ensure that when there is no nose (resp. tail) door associated to a position, the path from this door to position j is automatically forbidden, and consequently, $\beta_j^N = 1$ (resp. $\beta_j^T = 1$).

$$\alpha_j^N \leq \delta_j^N \qquad \forall j \in \mathbb{P} \tag{2.27}$$
$$\alpha_j^T \leq \delta_j^T \qquad \forall j \in \mathbb{P} \tag{2.28}$$
$$\beta_j^N = 1 - \delta_j^N \qquad \forall j \in \mathbb{P} \tag{2.29}$$
$$\beta_j^T = 1 - \delta_j^T \qquad \forall j \in \mathbb{P} \tag{2.30}$$

Because of ULD dimensions, some doors might not be accessible. Binary parameters μ_j^N and μ_j^T are introduced to manage door accessibility for a particular ULD.

$$\mu_i^N = \begin{cases} 1 & \text{if ULD } i \text{ pass through the nose door,} \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in \mathbb{U}.$$

$$\mu_i^T = \begin{cases} 1 & \text{if ULD } i \text{ pass through the tail door,} \\ 0 & \text{otherwise} \end{cases} \qquad \forall i \in \mathbb{U}.$$

Constraints (2.31) stipulate that if ULD i cannot pass through the nose door ($\mu_i^N = 0$), the unloading of this ULD i assigned to position j ($x_{ijk} = 1$) cannot occur through the nose door ($\alpha_j^N = 0$ must be zero). Constraints (2.32) are equally applicable to the tail door.

$$x_{ijk} + \alpha_j^N \leq 1 + \mu_i^N \qquad \forall k \in \mathbb{L} \;\; \forall i \in \mathbb{U}_k^L, \;\; \forall j \in \mathbb{P}. \tag{2.31}$$
$$x_{ijk} + \alpha_j^T \leq 1 + \mu_i^T \qquad \forall k \in \mathbb{L} \;\; \forall i \in \mathbb{U}_k^L, \;\; \forall j \in \mathbb{P}. \tag{2.32}$$

2.4.3.3 Other advanced constraints

Here are two advanced constraints linked to hazardous and oversized products[3].

Hazardous goods

Some loads can contain hazardous goods. For such cases, segregation requirements apply in order to ensure safety. In general, dangerous goods can

[3]We would like to thank T. Kleyntssens for his preliminary work on these constraints (Kleyntssens et al. (2012)).

be classified into a limited number of categories. The required segregation distance (in inches) between categories i and i' is known and denoted as $s_{ii'}$. The effective longitudinal distance (in inches) between positions j and j' is denoted as $e_{jj'}$. Based on this information, we introduce the following binary parameters $\zeta_{jj'}^{ii'}$:

$$\zeta_{jj'}^{ii'} = \begin{cases} 1 & \text{if } e_{jj} \leq s_{ii}, \\ 0 & \text{otherwise} \end{cases} \qquad \forall i, i' \in \mathbb{U},\ \forall j, j' \in \mathbb{P}.$$

Constraints (2.33) ensure that two hazardous ULDs i and i' are not assigned to positions j and j' if the distance between these positions does not exceed the requested segregation distance between the two ULDs. Indeed, if the distance between the two positions is not sufficient ($\zeta_{jj'}^{ii'} = 1$), x_{ijk} and $x_{i'j'k}$ cannot both be equal to one.

$$x_{ijk} + x_{i'j'k} \leq 2 - \zeta_{jj'}^{ii'} \qquad \forall k \in \mathbb{L},\quad \forall i, i' \in \mathbb{U}_k^L,\quad \forall j, j' \in \mathbb{P} \tag{2.33}$$

Oversized ULDs

Some oversized ULDs do not fit in one position. In this case, several positions are exactly or partially combined to form larger positions. In order to introduce these specific ULDs into our model, we divide each larger ULD into two smaller ones. However, we need to ensure that the model assigns the two parts of the larger ULD to two adjacent positions. If we denote as t_i the ULD linked to ULD i, and as $\mathbb{A}_j$ the set of positions adjacent to position j, Constraints (2.34) ensure the correct assignments of larger ULDs:

$$x_{ijk} \leq \sum_{\forall j' \in \mathbb{A}_j} x_{t_i j' k} \qquad \forall i \in \mathbb{U}_k^L,\ \forall j \in \mathbb{P},\ \forall k \in \mathbb{L} \tag{2.34}$$

2.5 Complexity

Let us now provide insight into the complexity of the problem. The first term of the objective function is to consider balanced loading and fuel consumption. We show below that the problem defined by this first part of the objective function and for one leg is already NP-hard.

Definition 2.5.1. *We define $ACLPPD_D$ as the decision version of $ACLPPD$ that asks whether the objective function (2.10) can reach a null value.*

Lemma 2.5.1. *$ACLPPD_D$ is in NP.*

Proof. Because the problem is expressed as a mixed integer linear problem, inputting the values of the decision variables into the model is sufficient for verifying in linear time all the constraints and value of the objective function. □

Now, we provide a reduction from the Set-Partition Problem (SPP) known to be NP-complete (Garey and Johnson (1979)). In SPP, the question is to determine whether it is possible to divide a set of numbers $S = \{s_1, \ldots, s_n\}$ into two disjoint subsets S_1 and S_2 so that $S = S_1 \cup S_2$, $S_1 \cap S_2 = \emptyset$, and $\sum_{i \in S_1} s_i = \sum_{i \in S_2} s_i$. For any instance of the Set-Partition problem, one can construct in polynomial time an instance $\tau(S)$ of ACLPPD$_D$ problem as follows:

$$\mathbb{U}_2 = \mathbb{U}_3 = \emptyset \qquad \text{One leg} \tag{2.35}$$

$$c^h = 0 \qquad \text{No handling costs} \tag{2.36}$$

$$\mathbb{P} = \{1 \ldots 2n\} \qquad 2n \text{ positions} \tag{2.37}$$

$$a_j = \begin{cases} OCG_1 - 1 & \forall j \in \{1 \ldots n\} \\ OCG_1 + 1 & \forall j \in \{n+1 \ldots 2n\} \end{cases} \qquad \text{Position locations} \tag{2.38}$$

$$\mathbb{E}_1 = \emptyset \qquad \text{No overlaying positions} \tag{2.39}$$

$$\mathbb{U} = \{1 \ldots n\} \qquad n \text{ ULDs} \tag{2.40}$$

$$w_i = s_i \quad \forall i \in \mathbb{U} \qquad \text{ULD weights} \tag{2.41}$$

$$\gamma_{ij} = 1 \quad \forall i \in \mathbb{U},\ \forall j \in \mathbb{P} \qquad \text{All positions are feasible} \tag{2.42}$$

$$minCG_1 = OCG_1 - 1 \qquad \text{Longitudinal stability} \tag{2.43}$$

$$maxCG_1 = OCG_1 + 1 \qquad \text{Longitudinal stability} \tag{2.44}$$

$$\mathbb{P}_R = \mathbb{P}_L = \emptyset \qquad \text{Only central positions} \tag{2.45}$$

$$\bar{W}_j = \max s_i \quad \forall j \in \mathbb{P} \qquad \text{No weight restriction} \tag{2.46}$$

$$O_a^d = \sum_{i \in S} s_i \quad \forall d \in \mathbb{D},\ \forall a \in \mathbb{O}^d \qquad \text{No weight restriction} \tag{2.47}$$

$$\bar{F}_a = \bar{T}_a = \sum_{i \in S} s_i \quad \forall a \in \mathbb{F} \qquad \text{No weight restriction} \tag{2.48}$$

Let us interpret each element of S as one standard ULD. The values $\{s_1, \ldots, s_n\}$ denote the ULD weights (Equations (2.40)-(2.41)). We build an aircraft with $2n$ positions (Equation (2.37)) divided into two subsets. The first (resp. last) n positions are located at one unit of distance to the front (resp. to the aft) with respect to the ideal location of the center of gravity OCG (Equation (2.38)). Let us denote as $\mathbb{P}_1$ (resp. $\mathbb{P}_2$) the set of forward (resp. aft) positions. Because the positions are located at only two different locations, this implies that the containers are stackable. Alternatively or simultaneously, we could consider each position to be on a different deck. In this sense, the positions are not defined as overlaying exclusive positions (Equation (2.39)). The corresponding ACLPPD$_D$ is less constrained than the general problem could be. All ULDs that can fit in any positions (Constraints (2.13) are always satisfied because of Equations (2.42). Longitudinal stability is guaranteed (Equations (2.43)-(2.44)) for Constraints (2.15)). All positions are located centrally (Equation 2.45), which ensures lateral stability (Constraints (2.16)). There is no weight restriction (Equations (2.46)-(2.48) vs. Constraints (2.17)-(2.20)). Finally, we assume that the aircraft possesses at least one deck and one door per deck.

The transformation τ is obviously polynomial. Let us now show that it transforms yes-instances of Set-Partition into yes-instances of ACLPP$_D$.

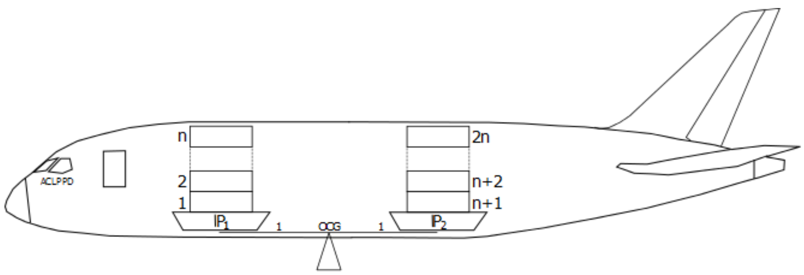


Figure 2.7: Set partition ACLPPD problem

Lemma 2.5.2. *Any feasible solution* $\{S_1, S_2\}$ *of the Set-Partition Problem can be converted in polynomial time into a solution of* $\tau(S)$ *with value zero.*

Proof. This Weight & Balance Problem is depicted in Figure 2.7. Let us apply transformation $\tau(S)$ and load randomly all ULDs associated to elements S_1 into positions in $\mathbb{P}_1$. By definition, $\mathbb{P}_1$ is sufficiently large to contain all these elements. The same is applicable for S_2 and $\mathbb{P}_2$. This is a feasible solution for ACLPPD. Because $\{S_1, S_2\}$ is a solution of the Set-Partition Problem, $\sum_{i \in S_1} w_i = \sum_{i \in S_2} w_i$ and exactly the same weight is loaded on each side of the requested center of gravity OCG. Because the distance between OCG and the two loading locations of loading are the same, this implies that CG is located exactly at OCG and that $\varepsilon_1 = 0$. Therefore, this is a solution with optimal value zero for ACLPPD.

□

Lemma 2.5.3. *Any solution with value zero for $\tau(S)$ can be converted in polynomial time into a solution of the Set-Partition Problem.*

Proof. The smallest value for ACLPPD objective value is zero when CG is located at OCG. Because the distance between the two locations and OCG is the same, this can occur only when exactly the same weight is loaded in the forward and aft areas. It suffices to list in S_1 (resp. S_2) the weights of the ULDs in $\mathbb{P}_1$ (resp. $\mathbb{P}_2$) to obtain a solution for the Set-Partition Problem.

□

Theorem 2.5.1. *$ACLPPD_D$ is NP-complete.*

Proof. This follows directly from Lemmata (2.5.1), (2.5.2), and (2.5.3). In reality, it also shows that the simpler Weight & Balance Problem without handling cost operations and only one leg is already NP-complete.

□

It is important to note that even if the Weight & Balance Problem for one leg and without handling operations costs is NP-hard, in a few minutes, we can solve all the realistic instances we tested. This might be because problem size is strongly limited by aircraft capacity. However, when there is more than one leg and the handling operations have to be optimized, the problem becomes far more difficult.

The complexity of defining assignments over two legs cannot be divided into two independent balanced loading problems. Considering two destinations implies solving simultaneously two (NP-hard) related instances of the balanced loading problem: one for each leg. The minimization of the number of operations (at the intermediate airport) is the link that makes these two problems dependent and that explains why it is far more complex to solve the problem with an intermediate airport than to independently solve two instances of the same size. Indeed, the optimization inputs over the second leg, i.e., mainly the positions that are available, depend on the set of feasible solutions for the optimization problem restricted to the first leg only. For each feasible solution of the latter, there exists one instance to optimize over the second leg. This drastically increases the global domain space.

Finally, another originality of our work is that unloading can occur through several doors. For each ULD, when settled in a given position, the question is as follows: through which door should the ULD be unloaded in order to minimize handling operations? This raises the question of complexity associated with the introduction of α_j^T and α_j^N in the model, and more operationally, of the optimal unloading sequence. This is not the most complex part of the problem. A naïve algorithm with complexity $O(n^2)$, where n is the number of ULDs, already answers this question. Indeed, because the ULDs cannot collide once unloaded, there exists, for each deck and lane, at least one position between two doors of the plane for which all ULDs close to the tail door leave by this door (or stay on board), and all ULDs close to the nose door leave by the nose door (or stay on board). For each lane and deck with several doors, each position from the nose door to the tail door can be successively considered as a candidate for partitioning. Globally n operations are required. Starting from this position, the next step is to return to the nose door and stop at the first ULD to be delivered at the intermediate airport (i.e., the first ULD that must be unloaded). From there, the next step is to move to the nose and count the number of ULDs to be delivered at the final airport, i.e., the sum of n_j^N. The same steps are applicable from the pivot position to the tail door to compute $\sum n_j^T$. Again, this is performed in n operations. The best partitioning position, which defines all α_j^T and α_j^N, is that which minimizes the sum of n_j^N and n_j^T, and the complexity of the entire algorithm is $O(n^2)$.

2.6 Implementation and results

Our mathematical model was tested on a set of real-world instances provided by TNT Airways, a wholly owned subsidiary of TNT Express. Their main activity is to provide TNT Express with an airfreight network that connects daily all TNT Express locations throughout the world, and more specifically, in Europe. TNT Express is one of the leading delivery integrators in Europe. The model was implemented in Java and relies on the IBM ILOG CPLEX 12 library (default parameters). Thanks to a graphical interface, it is possible to visualize, for each leg, the loading plans, CG position, and different weight distributions. The tests were performed on a personal computer (Windows 7, Intel Core i5-2450M, 2.50GHz, and 8.00 GB of RAM).

2.6.1 The case of a Boeing 777

First, we present the detailed results of a historical flight. This is an intercontinental flight with a first leg of approximately 2,740 nautical miles and a second leg of 3,200 nautical miles. The aircraft is a Boeing 777 Freighter (B777F). The B777F is the successor of the famous B747F, one of the most used freighters in the world. As illustrated in Figure 2.2, this Boeing 777 has four doors (one per compartment) and a rather large number, 98, of predefined positions. A total of 40 predefined positions are on the main deck. This corresponds to 13 large positions that overlap 26 small positions (defined on the right and left sides of the aircraft) and one last central position at the rear. On the lower deck, ten large positions (P) overlap 32 small positions (R and L). There is one more central position (C) for each pair of right and left positions.

The B777F was loaded at full capacity. The ULD sets of ULDs are listed in Table 2.1. For this shipment, we computed and set the optimal CG location for the first leg (resp. second leg) at 39.3% MAC (resp. 38.3% MAC), but we reduced it by 1% as a security margin. The tanks for each leg were filled with 48 tons of fuel. The cost c^h of an unnecessary operation is fixed at 40USD. Assuming an increase of 2% fuel consumption for a 10% MAC shift of the CG location, and a cost of 1 USD per fuel liter, we obtain an approximate fuel cost coefficient c_k^f of 100USD.

Set	Origin	Destination	# ULDs
$\mathbb{U}_1$	Liège	DES1	19
$\mathbb{U}_3$	Liège	DES2	24 (1 large one)
$\mathbb{U}_2$	DES1	DES2	9
$\mathbb{U}_1^L$	Leg 1 (Liège-DES1)		43 (1 large one)
$\mathbb{U}_2^L$	Leg 2 (DES1-DES2)		33 (1 large one)

Table 2.1: Data for our main case

The optimal solution found by our software is depicted in Figures 2.8 and 2.9, for the first and second leg, respectively. All positions are represented with boxes. When a ULD is assigned to a position, its representative box is colored, and the type and weight of the ULD is indicated on the box. In Figure 2.8, the ULDs in light (resp. dark) gray belong to $\mathbb{U}_1$ (resp. $\mathbb{U}_3$). This solution meets all requirements. The aircraft has a lateral weight im-

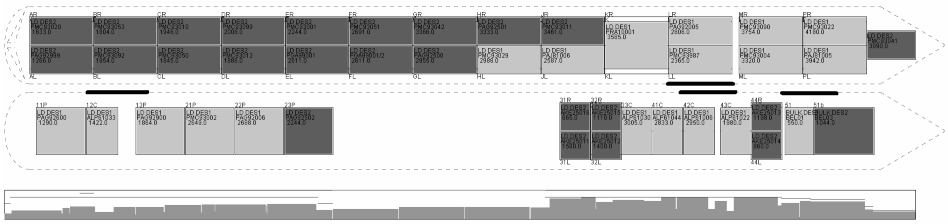


Figure 2.8: Optimal loading plan obtained by our model for the first leg

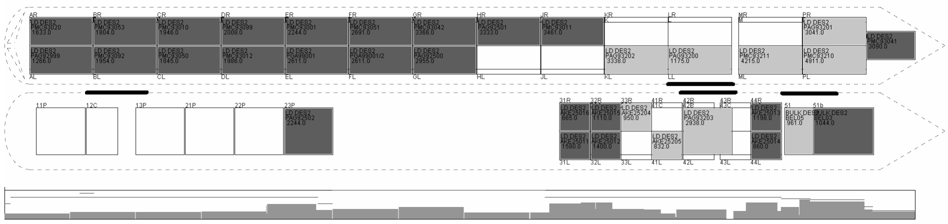


Figure 2.9: Optimal loading plan obtained by our model for the second leg

balance below the thresholds: 1,929Kg for the first leg and 4,089Kg for the second leg. In the same figures, the two rectangles situated below each loading plan provide the level of combined weights. The lines situated above are the thresholds. The cumulative load limits are also in compliance. For each leg, the three CGs (ZFW-TOW-LW) lie within their respective feasibility envelopes, as depicted in Figure 2.10.

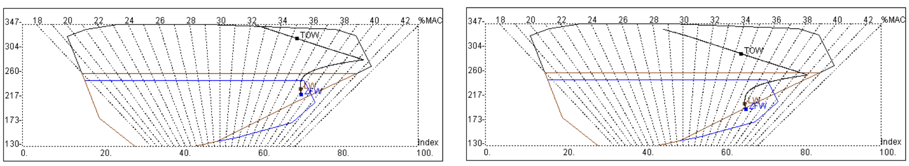


Figure 2.10: Feasibility envelopes as defined by the aircraft manufacturer

First, we measure the solution quality by observing the two terms of the objective function, i.e., the CG location and the number of re-handling operations. The solution found by our software is the best we could expect.

The requested CG location, minimizing fuel, is reached for each of the legs ($\varepsilon_k \simeq 0$), and no additional re-handling operations are required at the intermediate airport. This last fact appears clearly in Figures 2.8 and 2.9 because all the ULDs that must be unloaded at the intermediate airport are assigned positions close to the doors, and the other ULDs maintain the same positions over the two legs.

The second performance measure is computation time. We limit the computation precision to 1 USD. The CPLEX optimizer stops immediately after finding the optimal solution, or a feasible solution with an objective value, i.e., the excess cost, of no more than 1 USD. We know that the problem is NP-hard and that solving it could be time consuming. Indeed, we required 54 minutes to obtain a near-optimal solution (the excess cost is less than 1 USD) for this instance. However, this seems to be an extreme case and we usually obtain results in a few minutes (see the additional results in the next section). We believe that at least two reasons can explain why this specific instance remains difficult. First, the aircraft is loaded at full capacity on the first leg, which is rather uncommon. Second, the number of positions to consider is extremely large because of the number of different ULD types that are involved. Simply removing one ULD leads to an optimal solution achieved in only 4 min. It is also worth noting that the Branch and Cut process found a good feasible solution quickly but considerable time is required to validate the optimal solution. Therefore, we also provide the best result obtained with a time limit of 10 min. Finally, another interesting insight is that optimizing independently over the two legs, as done in Limbourg et al. (2012), only requires a few seconds. Considering several legs with P&D operations significantly increases the complexity of the problem.

	Load Master	Our solution	With time limit
%MAC ZFW (leg1)	24.4	38.3	38.3
%MAC ZFW (leg2)	31.1	37.3	37.3
# unloadings at DES1	43	19	23
# loadings at DES1	33	9	13
# ULDs $\in \mathbb{U}_3$ unloaded at DES1	24	0	4
Computation Time (in minutes)	20	54	10

Table 2.2: Loading into a Boeing 777 aircraft: main results

The last performance measure is to approximate potential savings. For that, we compare our solution with that obtained by the loadmaster who works manually. Loadmasters from different companies follow similar procedures, as verified through interviews. When time permits, the loadmasters attempt to load cargo in such a way that the observed CG lies close to a predefined value in the feasibility envelope, e.g., at 28% or 29% of the MAC range. The main results are summarized in Table 2.2.

To save fuel, our approach pushes the leg CGs as much as possible to the aft, well beyond the values usually considered. Using our initial assumptions, this implies a fuel savings of 2,010 USD. Moreover, no ULD in transit has to be unloaded at the first airport, whereas it is common to plan the two legs independently and to be required to unload all or a large part of the cargo, as observed in many companies. With respect to the worst case, this saves 24 unnecessary operations, i.e., approximately 960 USD. The total savings for this single trip adds to a maximum of 2,970 USD. When we limit computation time to only 10 min, we still reach an extremely good feasible solution. The total savings in this case adds to a maximum of 2,810 USD. Let us assume that this flight operates under the same conditions three times per week, almost every week, and in both directions; then, the savings can add to a maximum of 650,000 USD per year. Moreover, the same optimization process can be applied to all other aircrafts in the fleet.

2.6.2 Additional cases

The question now is to verify whether the results presented for the specific case in the previous section are representative. To provide a partial answer, we solved hundreds of other real-world cases. We considered the same aircraft, but with different loads and different sets of ULDs. We provide the results for eight configurations in Table 2.3. All other simulations lead to similar results. Cases (A) and (B) imply Pickup and Delivery Operations. No delivery and only pickups occur with cases (C), (D), and (E), and the opposite is true for cases (F), (G), and (H). We set a computation time limit of 10 min. The optimal CG location is reached for all instances, implying fuel savings. Case (H) is the only case that was stopped after 10 min of computation. CG is was at the optimal location, but we observed eight re-handlings of ULDs in transit. This remains a much better solution than that observed in practice, in which 23 ULDs in $\mathbb{U}_3$ were moved at the first airport.

We also generated some hazardous goods among the loaded ULDs ran-

	P&D		Pickup			Delivery		
	A	B	C	D	E	F	G	H
$\lvert\mathbb{U}_1\rvert$	22	20	0	0	0	24	18	19
$\lvert\mathbb{U}_3\rvert$	15	21	23	22	14	14	23	23
$\lvert\mathbb{U}_2\rvert$	14	10	14	20	22	0	0	0
Status	O	O	O	O	O	O	O	F
ε %MAC	0	0	0	0	0	0	0	0
# unloadings	0	0	0	0	0	0	0	8
Comput. time	1'15"	6'58"	2'04"	5'15"	54"	3'26"	5'58"	10'

Table 2.3: Additional results for different cases of ULDs to load

domly and we considered some cases in which a nose door was available. Because these experiments did not provide different results or insights, we decided not to include them in this paper.

2.7 Conclusion

In this chapter, we analyzed ACLPPD, which is a crucial problem encountered daily by airlines. We considered trips of several legs at the end of which P&D operations might occur. We proposed a new mixed integer linear model.

Our contributions are multiple. First, the model is based on international standards and is valid for most commercial operators. We integrated, and adapted to the multi-leg context, a large set of the constraints they encounter. Most operators should be able to use this approach in real life, and if needed, to extend it to any of their specifications. Second, we showed how to consider unloading and loading sequences when Pickup and Delivery arises. In this context, we again attempted to maintain reality by considering aircraft with several doors. Third, we showed that the Weight & Balance Problem is NP-hard. Considering several legs and several doors causes the problem to become even more complex to manage. Finally, another originality of our approach is the focus on cost. We analyzed two important costs directly linked to ULD loading: impact on fuel consumption and handling operations cost. We showed that locating the CG closer to the aft should be performed in order to decrease fuel consumption.

Our approach was tested on real data, and we conducted hundreds of experiments. We demonstrated that it is possible to quickly find near optimal or excellent feasible solutions, and that our approach leads to substantial savings with respect to current typical practices.

There are several ways in which future research can overcome the limitations mentioned above. These extensions are explained in details in Chapter 5.

Bibliography

Airbus Fuel Economy Material (2004). Getting to grips with fuel economy. `http://www.cockpitseeker.com/wp-content/uploads/goodies/ac/a320/pdf/data/FuelEconomy.pdf`. [Online; accessed 22-February-2016].

Altarazi, S. (2013). A new prioritizing-stacking heuristic algorithm for the inner-city truck loading problem. *International Journal of Business and Management 8*(13), 137–143.

Amiouny, S.-V., J.-J. Bartholdi, J.-H. Vande Vate, and J. Zhang (1992). Balanced loading. *Operations Research 40*(2), 238–246.

Bortfeldt, A. and G. Wäscher (2013). Constraints in container loading — a state-of-the-art review. *European Journal of Operational Research 229*, 1–20.

Bostel, N. and P. Dejax (1998). Models and algorithms for container allocation problems on trains in a rapid transshipment shunting yard. *Transportation Science 32*(4), 370–379.

Chan, F.-T.-S., R. Bhagwat, N. Kumar, M.-K. Tiwari, and P. Lam (2006). Development of a decision support system for air-cargo pallets loading problem: A case study. *Expert Systems with Applications 31*, 472–485.

Fok, K. and A. Chun (2004). Optimizing air cargo load planning and analysis. In *Proceeding of the International Conference on Computing, Communications and Control Technologies*, Austin, Texas, USA.

Garey, M. and D. Johnson (1979). *Computers and Intractability; A Guide to the Theory of NP-Completeness.* W. H. Freeman and Company.

Guèret, G., N. Jussien, O. Lhomme, C. Pavageau, and C. Prins (2003). Loading aircraft for military operations. *Journal of the Operational Research Society 54*, 458–465.

Heidelberg, K.-R., G.-S. Parnell, and J.-E. Ames (1998). Automated air load planning. *Naval Research Logistics 45*(8), 751–768.

International Air Transport Association (2016). Air cargo - enabling global trade. `https://www.iata.org/whatwedo/cargo/Pages/index.aspx`. [Online; accessed 22-February-2016].

International Air Transport Association (2016a). ULD Regulations (ULDR). `http://www.iata.org/publications/Pages/uld-regulations.aspx`. [Online; accessed 22-February-2016].

Imai, A., K. Sasaki, E. Nishimura, and S. Papadimitriou (2006). Multi-objective simultaneous stowage and load planning for a container ship with container rehandle in yard stacks. *European Journal of Operational Research 171*, 373–389.

Kaluzny, B.-L. and D.-R. Shaw (2009). Optimal aircraft load balancing. *International Transactions in Operational Research 16*, 767–787.

Kleyntssens, T., Limbourg, S., and M. Schyns (2012). Automatic Cargo Load Planning: Special shipments. In *Proceedings of the 4rd International Conference on Information Systems, Logistics and Supply Chain*, ILS, Quebec, Canada.

Larsen, O. and G. Mikkelsen (1980). An interactive system for the loading of cargo aircraft. *European Journal of Operational Research 4*, 367–373.

Li, Y., Y. Tao, and F. Wang (2009). A compromised large-scale neighborhood search heuristic for capacitated air cargo loading planning. *European Journal of Operational Research 199*, 553–560.

Limbourg, S., M. Schyns, and G. Laporte (2012). Automatic aircraft cargo load planning. *Journal of the Operational Research Society 63*, 1271–1283.

Mongeau, M. and C. Bès (2003). Optimization of aircraft container loading. *IEEE Transactions on Aerospace and Electronic Systems 00*, 1–27.

Nance, R.-L., A.-G. Roesener, and J.-T. Moore (2011). An advanced tabu search for solving the mixed payload airlift loading problem. *Journal of the Operational Research Society 62*, 337–347.

Ng, K.-Y.-K. (1992). A multi-criteria optimization approach to aircraft loading. *Operations Research 40*(6), 1200–1205.

Øvstebø, B.-O., L.-M. Hvattum, and K. Fagerholt (2011). Optimization of stowage plans for RoRo ships. *Computers & Operations Research 38*, 1425–1434.

Paquay, C., M. Schyns, and S. Limbourg (2016). A mixed integer programming formulation for the three-dimensional bin packing problem deriving from an air cargo application. *International Transactions in Operational Research 63*, 1271–1283.

Tang, C.-H. (2011). A scenario decomposition-genetic algorithm method for solving stochastic air cargo container loading problems. *Transportation Research Part E: Logistic and Transportation Review 47*(4), 520–531.

Tang, C.-H. and H.-W. Chang (2010). Optimization of stochastic cargo container loading plans for air express delivery. In *IEEE Second International Conference on Computer and Network Technology*, Bangkok, pp. 416–420.

Vancroonenburg, W., J. Verstichel, K. Tavernier, and G. Vanden Berghe (2014). Automatic air cargo selection and weight balancing: A mixed integer programming approach. *Transportation Research Part E 65*, 70–83.

Wu, Y. (2010). A dual-response forwarding approach for containerizing air cargoes under uncertainty, based on stochastic mixed 0-1 programming. *European Journal of Operational Research 207*, 152–164.

Yan, S., Y.-L. Shih, and F.-Y. Shiao (2008). Optimal cargo container loading plans under stochastic demands for air express carriers. *Transportation Research Part E 44*, 555–575.

Part II

Itinerary Choice Models

Chapter 3

Accounting for Price Endogeneity in Airline Itinerary Choice Models: An Application to Continental U.S. Markets

Lurkin, V., Garrow, L., Higgins, M., Newman, J., and Schyns, M. (2016). Accounting for price endogeneity in airline itinerary choice models: An application to continental U.S. markets. *Under review in Transportation Research Part A: Policy and Practice.*

Network planning models, which forecast the profitability of airline schedules, support many critical decisions, including equipment purchase decisions. Network planning models include an itinerary choice model which is used to allocate air total demand in a city pair to different itineraries. Multinomial logit (MNL) models are commonly used in practice and capture how individuals make trade-offs among different itinerary attributes; however, none that we are aware of account for price endogeneity. This chapter formulates an itinerary choice model that is consistent with those used by industry and corrects for price endogeneity using a control function that uses several types of instrumental variables. We estimate our model using a database of more than 3 million tickets provided by the Airlines Reporting Corporation. Results based on Continental U.S. markets for May 2013 departures show that models that fail to account for price endogeneity overestimate customers' value of time and result in biased price estimates and incorrect pricing recommendations. The size and comprehensiveness of our database allows us to estimate highly refined departure time of day preference curves that account for distance, direction of travel, the number of time zones traversed, departure day of week and itinerary type (outbound, inbound or one-way). These time of day preference curves can be used by airlines, researchers, and government organizations in the evaluation of different policies such as congestion pricing.

3.1 Introduction and motivation

Network planning models, which are used to forecast the profitability of airline schedules, support many important long- and intermediate-term decisions. For example, they aid airlines in performing merger and acquisition scenarios, route schedule analysis, code-share scenarios, minimum connection time studies, price-elasticity studies, hub location and hub buildup studies, and equipment purchasing decisions (Garrow et al. (2010)).

Network planning models forecast schedule profitability by determining the number of passengers who travel in an origin destination (OD) pair, allocating these passengers to specific itineraries, and calculating expected costs and revenues. The passenger allocation model is often referred to as an itinerary choice model because it represents how individuals make choices among itineraries. Many airlines use discrete choice models to capture how individuals make trade-offs among different itinerary characteristics, e.g., departure times, elapsed times, the number of connections, equipment types, carriers, and prices (see Garrow et al. (2010), Jacobs et al. (2012) for reviews of itinerary choice models used in practice and Coldren et al. (2003), Koppelman et al. (2008) for specific studies conducted for United Airlines and Boeing, respectively). However, to the best of our knowledge, none of the itinerary choice models used in practice account for price endogeneity. As explained in Chapter 1, endogeneity occurs when correlation exists between an explanatory variable and the error term (or unobserved factors) in a model. In demand models, price endogeneity occurs when prices are influenced by demand, i.e., higher prices are observed when demand is high and lower prices are observed when demand is low. Failure to correct for price endogeneity is critical, as it will result in biased estimates and incorrect profitability calculations. Recent work has focused attention on the importance of accounting for endogeneity in demand studies. For example, Guevara (2015) notes that "endogeneity often arises in discrete-choice models, precluding the consistent estimation the model parameters, but is habitually neglected in practical applications". Guevara (2015) provides several examples from the mode choice, residential location, and intercity travel demand literatures that provide evidence of endogeneity due to omission of attributes and reviews approaches researchers have been using to account for this endogeneity. These studies include those by Wardman and Whelan (2011) and Tirachini et al. (2013) for mode choice applications, Guevara and Ben-Akiva (2006, 2012) for residential location applications, and Mumbower et al. (2014) for intercity applications.

Earlier publications in air travel demand modeling have found strong evidence of price endogeneity. In Mumbower et al. (2014), they model flight-level price elasticities in four markets using linear regression models and find striking differences in price elasticity estimates between a model that ignores and a model that accounts for price endogeneity. The model that ignores price endogeneity produces inelastic results (-0.58) whereas the model that accounts for price endogeneity using a two-stage least squares (2SLS) approach produces elastic (-1.32) results. In Hotle et al. (2015), they investigate the impact of airlines' advance purchase deadlines on individuals' online search and purchase behaviors for 60 markets. Their model, which is also based on a 2SLS method, finds strong evidence of price endogeneity.

This study builds on prior research by showing how to correct for price endogeneity for an itinerary choice model that is consistent with those used by industry. Unlike the previous applications, our model includes "all" continental U.S. markets and is based on discrete choice versus linear regression methods. Specifically, we follow the approach of Coldren et al. (2003) described for United Airlines and use a multinomial logit (MNL) to model itinerary choice for Continental U.S. markets. Results demonstrate the importance of accounting for price endogeneity; failure to do so results in value of time estimates that are too high, biased price estimates, and incorrect pricing recommendations. The results are intuitive and validation tests indicate that the corrected model outperforms the uncorrected specification.

Our study is distinct from the majority of prior studies reported in the literature in that we use a large database of individual tickets from multiple carriers for our analysis. Specifically, we estimate our model using an analysis database of 3 million tickets provided by the Airlines Reporting Corporation (ARC). We are uniquely positioned to examine the potential of using the ARC ticketing database for itinerary choice modeling applications as we are able to work with detailed price data whereas airlines cannot due to antitrust regulations. Our research contributes to the literature in three key ways. First, we demonstrate the ability to use the ARC ticketing database (in spite of its limitations) to replicate itinerary choice models representative of those used in practice. Second, we find a valid set of instruments to correct for price endogeneity for Continental U.S. markets. Third, due to the size of our analysis database, we are able to estimate detailed departure time of day preference curves that are segmented by distance, direction of travel, number of time zones traveled, day of week, and itinerary type (outbound, inbound or one-way). To the best of our knowledge, these curves represent the most refined publicly-available estimates of airline passengers' time of day preferences.

The remaining sections are organized as follows. Section 3.2 describes the data processing assumption we used to create our analysis database and the variables used in our study. Section 3.3 presents our methodology, with a particular focus on how we addressed price endogeneity. Empirical results are presented in Section 3.4. We conclude by highlighting how our model contributes to the literature and offering directions for future research, many of which are based on the data limitations commonly faced by industry when estimating discrete choice models for itinerary choice applications.

3.2 Data

This section describes the data and variables we used, explains the process we used to generate choice sets, and assesses the representativeness of our analysis database.

3.2.1 Airlines reporting corporation ticketing database

The Airlines Reporting Corporation provided the data for our analysis. As explained in Chapter 1, ARC is a ticketing clearinghouse that maintains financial transactions for all tickets purchased through travel agencies worldwide. Some carriers, most notably Southwest, are underrepresented in the database because the majority of their ticket sales are through direct sales channels (e.g., Southwest.com) that are not reported to ARC.

ARC has detailed information associated with each ticket. This includes the price paid for the ticket (and associated taxes and currency), ticketing date, booking class, and detailed information about each flight associated with the ticket, e.g., departure and arrival dates/times; origin, destination, and connecting airports; total travel time; connecting times; flight numbers; equipment types and associated capacities; and operating and marketing carriers. ARC classifies tickets into five product categories: First, Business, Unrestricted Coach, Restricted Coach, and Other/Unknown. This product classification is based on tables provided by the International Air Transport Association (IATA) that associates booking classes for each carrier with these five product categories.

The ticketing database provided by ARC contains tickets that have at least one leg that departed in May of 2013. May was selected because it is a month with average demand that falls between off-peak and peak seasons. Given the majority of these tickets are for travel that originates and terminates within the continental U.S., we restrict our analysis to these markets. Only tickets with six or fewer legs representing simple one-way or round-trip journeys were included in the analysis. More than 93% of all tickets in the ARC database can be classified as simple one-way and round-trip tickets. A simple one-way ticket does not contain any stops. A stop occurs when the time between any two consecutive flights is more than six hours. A simple round-trip itinerary represents a journey in which the individual starts and ends the journey in the same city and makes at most one stop in a different city. Round-trip itineraries can include multiple airports that belong to the same city, e.g., an individual who flies round-trip from San Francisco to Chicago can fly from San Francisco (SFO) to Chicago O'Hare (ORD), make a stop in Chicago, and then fly from Chicago Midway (MDW) to Oakland (OAK).We excluded tickets that had directional fares of less than $50 to eliminate tickets that were (likely) purchased using miles or by airline employees. We also calculated the 99.9th fare percentile for four product classes (First, Business, Unrestricted Coach, Restricted Coach/Other). For each product class we eliminated 0.1% of observations that had the highest prices. This process, which is consistent with that used by ARC, was done to eliminate tickets that were (likely) charter flights.

Our final database used for model estimation contains 3,265,545 directional itineraries, representing 10,034,935 passenger trips.

3.2.2 Variables definitions

The dependent variable in our itinerary choice models is a choice indicator defined as follows:

$$y_{ni} = \begin{cases} 1 & \text{if individual } n \text{ chooses itinerary } i, \\ 0 & \text{otherwise} \end{cases}$$

Table 3.1 defines and describes the independent variables included in our final itinerary choice models. Several additional variables related to carrier presence[1] were also included in the analysis but were not significant and excluded from the final model specification. Among those variables included in our models, the definitions and descriptions for elapsed time, number of connections, equipment type, and carrier preference (also referred to as carrier-specific constants) are straight-forward to interpret. Variables used to define direct flights, departure time of day, price, and marketing relationships merit additional discussion.

Variable	**Definition**
Travel Time, Number of Connections, Connection, and Equipment Attributes	
Elapsed time	Elapsed time is defined as the difference between the arrival time at the itinerary destination and the departure time at the itinerary origin (in minutes). All arrival and departure times are reported in Coordinated Universal Time (UTC), which accounts for time zone differences.
Number of connections	Number of itinerary connections. A value of zero indicates a nonstop itinerary and a value of one (two) indicates a single (double) connection.
Direct flight	A "direct flight" is one that has two flight legs. The operating carrier and operating flight number of the two flight legs are the same. A value of one indicates a direct flight. Those flights that are not direct include nonstop, single connection, and double connection itineraries. A direct flight is defined to have zero connections.

[1]Several studies have found that increased carrier presence in a market leads to increased market share for that carrier (Algers and Beser (2001), Benkard et al. (2008), Cornia et al. (2012), Gayle (2008), Nako (1992), Proussaloglou and Koppelman (1999), Suzuki et al. (2001)).

Wide- or narrow-body Regional jet or propeller	Equipment types include two categories. The first includes wide-body and narrow-body aircraft (no regional jets). The variable is equal to one if the equipment type is a wide-body or narrow-body aircraft and zero otherwise. The second includes narrow-body regional jets and propellers. The variable is equal to one if the equipment type is a narrow-body regional jet or a an aircraft propeller and zero otherwise. For itineraries with more than one leg, the smallest equipment type is used.
Departure Time of Day	
$sin2pi\_DOW_i\_TripType_j$ ... $cos6pi\_DOW_I\_TripType_J$	Itinerary departure time is expressed in local time as minutes past midnight (range: [0,1339]). Departure time are modeled using 1260 terms. Three sine (sin2pi, sin4pi, sin6pi) and three cosine functions (cos2pi, cos4pi, cos6pi) apply to each departure day of week i=1,2,...,7 and three trip types j=outbound, inbound, one-way.
Price	
Average high yield fare Average low yield fare	Fares are expressed in USD. We calculate separate prices for high yield and low yield fare products. We include First, Business, and Unrestricted Coach products as high yield fares and the Restricted Coach and Other/Unknown products as low yield fares. We calculate average high yield and average low yield fares for each itinerary origin, destination, carrier, and level of service (nonstop/direct, single connection, and double connection).
Marketing Relationships	
Online	An online itinerary is one that has the same marketing and same operating carrier for all legs. The variable is equal to one if it is an online itinerary and zero otherwise.

Codeshare	A codeshare itinerary is one that has the same marketing carrier for all legs, but different operating carriers. The variable is equal to one if it is an codeshare itinerary and zero otherwise.
Interline	An interline itinerary is one that has different marketing carriers. Only itineraries with two or more legs can be interline itineraries. The variable is equal to one if it is an interline itinerary and zero otherwise.
Carrier Preference	
Carrier_1 Carrier_2 ... Carrier_9	For k=1,...,9, the indicator variable Carrier_k =1 if the itinerary operating carrier associated with an itinerary is carrier k and 0 otherwise. The itinerary operating carrier is defined as the carrier that operates the longest flight leg. The first eight terms represent carriers that each have more than 1% market share in the estimation data. All other carriers are combined into the Carrier_9 term. Carrier names are suppressed to maintain confidentiality.

Table 3.1: Independent variables and definitions

Direct itineraries

We include nonstop, single connection, double connection and direct itineraries in our analysis. Nonstop, single connection and double connection itineraries have zero, one, and two stops, respectively. Similar to a single connection, a direct itinerary also has one stop. For a single connection itinerary, the flight numbers and aircraft used for each leg differ whereas for a direct itinerary, the flight numbers for each leg are identical and the aircraft used for each leg is (typically) the same[2]. Direct itineraries

[2]For example, Southwest distinguishes between nonstop, direct, and single connection itineraries on its website. Direct itineraries have a single flight number and are indicated as a "1 Stop No Plane Change" whereas connection itineraries have two flight numbers listed and are indicated as "1 Stop Change Planes XXX" where XXX is the airport code for the connection city (Southwest Airlines (2016)).

are more attractive than single connection itineraries, as the passenger typically "stays with the same aircraft" throughout the journey. For these reasons, we follow the approach used by other researchers (e.g., see Coldren et al. (2003), Coldren and Koppelman (2005a,b), Koppelman et al. (2008)) and distinguish between single connection and direct itineraries.

Departure time of day preferences

There are multiple approaches that can be used to model departure time preferences. The first approach uses a set of categorical variables to represent nonoverlapping departure time periods, e.g., one variable for each departure hour. However, the use of categorical variables can be problematic for forecasting applications when the difference in coefficients associated with two consecutive time periods is large (e.g., for the departure periods 9:00-9:59 AM and 10:00-10:59AM). In this case, moving a flight by a few minutes (e.g., from 9:58 AM to 10:02 AM) can result in unrealistic changes in demand predictions. The second approach overcomes this limitation by using a continuous specification that combines sine and cosine functions. We model time of day preferences using a continuous time of day formulation and follow the approach originally proposed by Abou-Zeid et al. (2006) for intracity travel and adapted by Koppelman et al. (2008) for itinerary choice models[3] by including three sine and three cosine functions representing frequencies of 2π, 4π, and 6π. For example, the sin 2π term is given as:

$$\sin 2\pi = \sin\left(\frac{2\pi \times \text{departure time}}{1440}\right) \tag{3.1}$$

where departure time is expressed as minutes past midnight and 1440 is the number of minutes in the day. Similar logic applies to the $\sin 4\pi$, $\sin 6\pi$, $\cos 2\pi$, $\cos 4\pi$, and $\cos 6\pi$ terms. One of the main contributions of our research (which is possible due to the size of our analysis database) is that we allow departure time preferences to vary according to several dimensions including the length of haul, direction of travel, number of time zones crossed, departure day of week, and itinerary type (i.e., outbound, inbound and one-way itineraries). More precisely, we create ten segments based on the length of haul, direction of travel and number of time zones crossed. For each segment, we estimate separate time of day preferences for departure day of week

[3] Carrier (2008) uses four sine and four cosine functions to model departure time preferences for European markets.

and itinerary type. Thus, our model includes 1260 departure time preference variables[4].

Note that our time of day formulation assumes a 24-hour cycle; however, some researchers (most notably Carrier (2008)) have suggested that truncated cycle lengths are more appropriate. Appendix A3 compares different time of day formulations and shows that our assumption of using a 24-hour cycle is appropriate.

In developing a model for United Airlines, Coldren et al. (2003) estimated 16 separate MNL models for Continental U.S. markets, one for each time zone pair (e.g., itineraries that start and end in the Eastern time zone (EE), itineraries that start and end in the Central time zone (CC), etc.). The authors note that, aside from time of day preferences, the estimated coefficients for other itinerary characteristics were similar across these 16 segments. We modify the segmentation approach proposed by Coldren et al. (2003) to: (1) distinguish between short and long distances within the same time zone; and, (2) combine time zone pairs that correspond to the same direction of travel and number of time zones. Descriptive statistics for our ten segments are shown in Table 3.2. The table provides information about the total number of city pairs, itineraries, and passengers associated with each segments. The mean, minimum and maximum distance traveled in each segment are also shown. This detailed segmentation allows us to estimate time of day preferences that vary as a function of distance[5], direction of travel, and the number of time zones traveled (in addition to the itinerary type (outbound, inbound, or one-way) and the departure day of week).

Price

The ARC ticketing database contains ticket-level price information linked to specific itineraries and the time of purchase. This price included on the ticket includes only the base fare (which corresponds to the revenues the airline receives) and does not include information on additional ancillary fees (such as fees for checking baggage or reserving a seat). Information about

[4]Note that 6 sine and cosine functions $\times 10$ segments $\times 7$ days of week $\times$ 3 itinerary types=1260 variables.

[5]For the itineraries whose the origin and the destination are in the same time zone or separated by only one time zone, we distinguish between short haul itineraries (≤ 600 miles) and medium haul (> 600 miles).

Segment	City pairs	Itins	Pax	Distance Min	Distance Mean	Distance Max
Same TZ, distance $\leq$ 600 miles	4,703	711,282	2,219,511	31	415.0	600
Same TZ, distance > 600 miles	3,524	520,481	1,848,742	601	839.3	1,534
One TZ WB, distance $\leq$ 600 miles	859	112,615	306,119	84	463.6	600
One TZ WB, distance > 600 miles	3,864	498,999	1,466,815	601	993.9	1,925
One TZ EB, distance $\leq$ 600 miles	863	115,187	312,265	84	462.0	600
One TZ EB, distance > 600 miles	3,898	501,345	1,446,807	601	993.7	1,925
Two TZ WB	1,860	239,936	681,666	643	1,576.4	2,451
Two TZ EB	1,823	233,113	684,627	643	1,571.4	2,451
Three TZ WB	1,121	165,428	509,346	1,578	2,203,3	2,774
Three TZ EB	1,091	167,159	559,037	1,578	2,210.9	2,774
TOTAL	23,606	3,265,545	10,034,935			

Table 3.2: Descriptive statistics by segment
Key: TZ = Time Zone, WB = Westbound, EB = Eastbound

taxes and fees applied to the base fare are included in the ARC ticketing database. In the U.S., domestic air travel taxes and fees include four main categories: a passenger ticket tax (7.5 percent of the base fare); a flight segment tax ($3.90 a flight segment); a passenger facility charge (up to $4.50 a flight segment); and a federal security fee, also called the Sept. 11 fee ($2.50 a segment). These taxes and fees are not revenues the airline receives. The first two taxes go to the Airport and Airway Trust Fund, which finances the Federal Aviation Administration. Passenger facility charges are passed on to airports and security fees finance the Transportation Security Administration.

Our discussions with industry practitioners revealed differing (and often strong) opinions as to whether the "price variable" included in itinerary choice models should include or exclude these taxes and fees. We discovered that multiple U.S. airlines and aviation consulting firms do not include these taxes and fees in their "price variable". Two primary reasons were offered for this practice: (1) these firms believed models that included taxes and fees provided results similar to those that excluded taxes and fees; and, (2) these firms noted that airlines receive revenues only from the base fare. Conversely, those firms that did include taxes and fees in their "price variable(s)" noted that: (1) including taxes is critical for international itineraries, as the taxes and fees can be quite large and exceed the base fare; and, (2) customers do not see the base fare, but rather the "total" price of the itinerary, thus models that represent the "price variable" as the sum of the base fare, taxes, and fees better reflect customer behavior.

As part of our modeling exercise, we estimated models that included taxes and fees and compared them to models that excluded taxes and fees. Results were similar for the two price formulations; however, the model that included taxes and fees fit the data slightly better. We include a price variable that includes the base fare as well as taxes and fees in our specifications as this variable better reflects the prices considered by consumers.

There are several other assumptions we used to create our price variable. Although we have detailed, ticket-level data in our analysis database, it is important to note that due to antitrust concerns, airlines do not have access to this same information for their competitors. For example, the U.S. Department of Transportation's Origin and Destination Survey Databank 1A/1B (DOT (2016)) provides a 10% sample of route-level prices, i.e., the actual price paid for a ticket is known but it is not linked to the time of purchase (number of days in advance of flight departure) or specific itineraries (e.g., flight numbers and departure times). Given our focus on demonstrating how we can address price endogeneity in itinerary choice models representative of those used in practice, we include an "average" price variable that is similar to that used by industry. Our price variable represents the average price paid by consumers for a specific itinerary origin, destination, carrier, level of service (i.e., nonstop/direct, single connection, double connection), and product type (i.e., high-yield (business) or low-yield (leisure)[6]). Also, consistent with industry practice, for round-trip itineraries, we assume the price associated with an outbound or inbound itinerary is the ticket price/2.

Marketing relationships

A codeshare is a marketing relationship between two airlines in which the operating airline allows its flight to be sold by a different carrier. Codeshare relationships can be determined from the ARC ticketing database using information about marketing and operating carriers. Each flight leg in the ARC ticketing database has a marketing carrier, marketing flight number, operating carrier and operating flight number. The marketing carrier is the carrier that sold the flight. The operating carrier is the airline that physically operated the flight. A codeshare itinerary is one that has the same marketing carrier for all legs, but different operating carriers. As an example, consider a ticket purchased from U.S. Airways for travel from Seattle (SEA) to Dallas (DFW) through Phoenix (PHX); the first leg is sold as US flight 102 and is operated by U.S. Airways (as US102) and the second leg is sold as US flight 5998 and is operated by American Airlines (as AA1840). In this example, the marketing carrier for each leg is the same because two US Airways flight numbers are used to sell the ticket - US102 and US5998, i.e., American and US Airways have established a marketing agreement that allows US Airways to sell tickets on AA1840.

Individuals can also purchase an itinerary that has two operating carriers that do not have a marketing relationship. We define an interline itinerary as one that has different marketing carriers. An interline itinerary is less attractive than a codeshare itinerary because there is no coordination - or

[6]We recognize that not all passengers who purchase a high-yield product that includes First, Business, or Unrestricted Coach tickets will be traveling for "business" and that not all passengers who purchase a low-yield product that includes Restricted Coach or Other tickets will be "leisure" passengers. Consistent with industry practice, we use the terms "business" and "leisure" interchangeable with higher-yield (predominately business) travel and lower-yield (predominately leisure) travel.

joint responsibility - between the two operating carriers. For example, if a bag is checked, the passenger will need to exit security at the connecting airport, retrieve the bag, and re-check it on the airline operating the second leg. Unlike a codeshare, if the first leg is delayed, the airline operating the second leg has no obligation to accommodate the passenger on a later flight.

An itinerary that is neither a codeshare or interline itinerary is an online itinerary. An online itinerary is one that has the same marketing and operating carrier for all legs of the itinerary.

3.2.3 Construction of choice sets

The ARC database provides information on the itinerary that was purchased by an individual; however, in order to model itinerary choices using discrete choice models, we also need to know what other alternatives were available and not chosen by the individual. We construct choice sets for each OD city pair that departs on day of week d using the revealed preferences from the ARC ticketing database.[7] We assume that any alternative purchased on day of week d_a, $a = \{\text{Monday,Tuesday,...,Sunday}\}$ was also available for purchase for all a days in the month, e.g., if an itinerary was purchased on the first Monday in May 2013, we assume that the itinerary was available on all Mondays in that month. We need to select a representative Monday that we can use to populate schedule attributes (except for marketing relationships). We follow the convention of United Airlines (Garrow (2004)) and define the representative week as the week beginning the Monday after the ninth of the month. This corresponds to May 13 - May 19, 2013 in our data. If an itinerary was not purchased during the representative week, we populate itinerary attributes (except for marketing relationships) based on the first day of the week in the month the itinerary was purchased. In our MNL model, the number of passengers who chose an itinerary represents the total number of passengers who traveled on day of week d_a in May 2013 on that itinerary.

Formally, we define a unique itinerary as follows: Given m legs, a unique itinerary departing on day d_a is defined by the $\{leg_m$ origin airport, leg_m destination airport, leg_m operating carrier, and leg_m operating flight number for

[7]We recognize that using the observed choices to create the choice sets is not ideal and is a concern that could bias the parameter estimates. Theoretically, choice sets should be generated using an exogenous process using leg-level schedule files as an input. These schedule files were not available at the time of the analysis.

$m = 1, 2, 3\}$. We assume that if any of the itineraries meeting this definition was sold as a codeshare during the month, the unique itinerary is a codeshare.

We performed a sensitivity analysis on each variable in the utility function to ensure the assumptions we used to populate schedule attributes were reasonable and did not result in large measurement errors due to using a representative week. The percentage of itineraries in our analysis database that have a measurement error is small (we estimated these errors to be less than 2 percent for any given schedule attribute). An example of the process we used to construct choice sets is included in Appendix A1.

Finally, to improve computational efficiency, we only included in our analysis OD pairs that had more than 30 passengers. We performed a sensitivity analysis on our MNL model to ensure this assumption was innocuous. Specifically, we estimated a MNL model based on itineraries with an origin in the Eastern time zone and a destination in the Western time zone with all OD pairs and compared it to one that only included OD pairs with more than 30 passengers. Excluding intercept terms, the parameter estimates between these two models differed by at most 5 percent and did not impact behavioral interpretations.

The table 3.3 provides information about the total number of choice sets associated with each segments. The mean, minimum and maximum number of alternatives by choice sets are also shown.

3.2.4 Representativeness of data

The ARC ticketing database is non-representative of the U.S. market as it does not contain tickets purchase from some distribution channels, most notably direct sales channels such as Southwest.com. This can be seen in Table 3.4, which compares carrier market shares between the ARC and DB1B databases[8]. The ARC database contains proportionately more tickets from major carriers, and fewer tickets from low costs carriers.

[8]In addition to direct sales, there are other differences between the ARC and DB1B databases that can influence market share calculations. ARC data contains tickets for travel in May 2013 whereas the DB1B data contains tickets for travel in 2Q of 2013. ARC data represents the last known ticket status provided to the travel agency. Ticketing changes that occur when a passenger calls the airline (and not the travel agency) are not reflected in the ARC data. The DB1B data are tickets that were ultimately used.

Segment	Choice sets	Choice sets Min Alts	Mean Alts	Max Alts
Same TZ, distance ≤ 600 miles	30,943	2	20.6	95
Same TZ, distance > 600 miles	22,861	2	27.6	105
One TZ WB, distance ≤ 600 miles	5,617	2	20.3	64
One TZ WB, distance > 600 miles	24,820	2	29.1	127
One TZ EB, distance ≤ 600 miles	5,630	2	19.6	63
One TZ EB, distance > 600 miles.	25,062	2	28.0	137
Two TZ WB	11,505	2	33.1	133
Two TZ EB	11,267	2	29.7	93
Three TZ WB	6,732	2	41.5	156
Three TZ WB	6,619	2	37.4	138
TOTAL	151,056			

Table 3.3: Choice sets by segment
Key: TZ = Time Zone, WB = Westbound, EB = Eastbound

Carrier	ARC Mkt Share	DB1B Mkt Share
Delta Air Lines (DL)	29.5%	23.4%
United Airlines (UA)	22.9%	17.1%
US Airways (US)	18.4%	10.0%
American Airlines (AA)	17.5%	19.0%
Alaska Airlines (AS)	3.3%	4.2%
JetBlue Airways (B6)	3.2%	3.0%
Frontier Airlines (F9)	2.2%	1.7%
AirTran Airways (FL)	1.4%	2.8%
Virgin America (VX)	1.3%	0.9%
Sun Country Airlines (SY)	0.3%	0.2%
Southwest Airlines (WN)	0.0%	17.7%
Total	100%	100%

Table 3.4: Airline market shares in ARC and DB1B Data

Although the sample is not representative of the population in every way, this is less of a concern when the purpose of the sample is to uncover relationships among variables (as it is here) than when it is purely to describe a population (Babbie (1998), Groves (2004)). For example, if we were using the sample to estimate the true share of various carriers in the population, it would be problematic, but a model based on the sample can properly predict itinerary choice given distribution channel. In particular, when the model is a MNL model, Manski and Lerman (1977) showed that under certain conditions, the MNL parameter estimates obtained from a stratified sample would be consistent and unbiased relative to the MNL estimates obtained from a simple random sample. Thus, we do not expect that parameter estimates for the variables shown in Table 3.1 will be impacted by the non-representativeness of our estimation database.

3.3 Methodology

This section reviews the MNL model, describes how we used a control function to account for price endogeneity.

3.3.1 Multinomial logit model

We model the itinerary choice y_ni that individual n chooses alternative i:

$$y_{ni} = \begin{cases} 1 & \text{if individual } n \text{ chooses itinerary } i, \\ 0 & \text{otherwise} \end{cases}$$

Each choice set is modeled as the set of all directional itineraries between each city-pair for each day of the week d. For example, all Monday itineraries between Atlanta and Chicago constitute a choice set. This choice is as a function of itinerary, carrier, and product characteristics. We exclude socioeconomic information as we have no information about the individual who purchased the ticket.

For cases where y_ni represents discrete outcome, as in the current situation, it is natural to model the probability that y takes on a given value, using a discrete choice model such as the MNL (McFadden (1974)). The majority of prior studies have used MNL models for itinerary choice applications, including those that describe models used in practice (e.g., see Coldren et al. (2003)). Given the focus of our study is on determining how we can correct for price endogeneity and include price for representative itinerary choice models used in practice, we thus follow this convention and use MNL models[9]. In the MNL, the utility U for individual n in choosing alternative

[9]We recognize that our assumption that customer preferences and competition among alternatives can be represented using a MNL model is questionable. However MNL model remains the most common model used in practice. More advanced discrete choice models

i from choice set $\mathbb{J}_n$ is a linear function of $\mathbf{x}_{ni}$:

$$U_{ni} = \boldsymbol{\beta}_i' \mathbf{x}_{ni} + \varepsilon_{ni} \tag{3.2}$$

where $\mathbf{x}_{ni}$ comprises the itinerary, carrier and product variables described in Table 3.1 and $\boldsymbol{\beta}_i'$ is the transpose of the vector of coefficients associated with all variables. If ε_{ni} is distributed independently and identically with a Gumbel (or extreme value type I) distribution, the probability of individual n choosing alternative i is given as:

$$P(y_n = i|\mathbf{x}_{ni}) = \frac{e^{\boldsymbol{\beta}_i' \mathbf{x}_{ni}}}{\sum_{j \in \mathbb{J}_n} e^{\boldsymbol{\beta}_j' \mathbf{x}_{nj}}} \tag{3.3}$$

3.3.2 Price endogeneity

Many prior studies of airline demand have failed to properly address price endogeneity and have assumed that prices are exogenous. As already said, endogeneity occurs when correlation exists between an explanatory variable and the error term (or unobserved factors) in a model. This correlation means that the conditional expectation of the error term on the endogenous explanatory variable will not equal zero, which violates a main assumption required to ensure estimator consistency for most models (Greene (2003)).

In demand models, prices are endogenous because they are influenced by demand, which is influenced by prices (often referred to as simultaneity of supply and demand). Many empirical demand studies have shown that price coefficients are underestimated if endogeneity is not corrected, including recent studies that estimate: demand for high speed rail travel (Pekgün et al. (2013)), household choice of television reception options (Goolsbee and Petrin (2004), Petrin and Train (2010)), household choice of residential location (Guevara and Ben-Akiva (2006), Guevara (2010)), choice of yogurt and ketchup brands (Villas-Boas and Winer (1999)), consumer-level choice of and aggregate product demand for the make and model of a new vehicle (Berry et al. (1995, 2004), Train and Winston (2007)), and brand-level demand for hypertension drugs in the U.S. (Branstetter et al. (2011)).

There are multiple methods that can be used to correct for price endogeneity, including the two-stage control-function (2SCF) method that accounts for endogeneity using instruments (Guevara (2015)). An instrument

that allow for different substitution patterns across product dimensions are clearly desirable and are estimated in Chapter 4.

is a variable that does not belong in the demand equation, but is correlated with the endogenous price variable. Instruments that satisfy the following two conditions will generate consistent estimates of the parameters, subject to the model being correctly specified: (1) instruments should be correlated with the endogenous variable, and (2) they should be independent of the error term in the model (Rivers and Vuong (1988), Villas-Boas and Winer (1999)). Therefore, we need to find instruments that are correlated with airfares but not correlated with a customer's purchase or choice of an itinerary. Validity tests are used to statistically determine whether the instruments are correlated with airfares, but not correlated with the error term of the demand model (i.e., customers' purchase or choice of a flight).

Mumbower et al. (2014) review instruments that have been or could potentially be used in airline applications and classify these instruments into four main categories: (1) cost-shifting instruments; (2) Stern-type measures of competition and market power; (3) Hausman-type price instruments; and, (4) BLP-type measures of non-price characteristics of other products. Cost-shifting instruments help explain why costs differ across geographic areas and/or product characteristics. Stern-type measures of competition and market power focus on the number of products in the market and also the time since a product (and/or firm) was introduced into the market (Stern (1996)). Hausman-type price instruments are based on prices of the same airline in other geographic contexts (Hausman et al. (1994), Hausman (1996)). BLP instruments, introduced by Berry et al. (1995), are based on the average non-price characteristics of other products.

We use two instruments to correct for endogeneity: the first is a Hausman-type price instrument, the other a Stern-type competition instrument. The Hausman-type instrument is calculated for itinerary i as the cube of the average price of similar itineraries in other similar markets. Itineraries are considered to be similar if they have the same carrier and level of service (i.e., nonstop/connection(s)). We assume that markets are similar if they have the same level of competition (i.e., presence of a low-cost carrier or not). For Stern-type competition instrument, we use a measure of capacity, i.e., the square of monthly seats flown in market by carrier and product type (i.e., the business or leisure).

The first-stage of our two-stage control-function (2SCF) model is an ordinary least-square (OLS) regression, Equation (3.4), that uses price as the dependent variable. As noted by Guevara and Ben-Akiva (2006), the purpose of the price equation is not to make a precise forecast of the price but to correct for endogeneity. Explanatory variables include the set of instruments along with all other exogenous regressors (except for price) used in the discrete choice model. The residual, defined as the difference between the actual and predicted price $\widehat{\delta}_{ni} = p_{ni} - \widehat{p}_{ni}$ from the first stage regression is introduced in the second-stage discrete choice model regression, Equation (3.5). The first-stage regression model and second-stage discrete choice model are formulated as follows:

- Stage 1: Estimate price by ordinary-least-square (OLS)

$$p_{ni} = \alpha_1 IV_{ni}^1 + ... + \alpha_k IV_{ni}^k + \boldsymbol{\gamma}_i' \mathbf{x}_{ni} + \delta_{ni} \tag{3.4}$$

- Stage 2: Estimate the choice model using the residuals δ from Stage 1

$$U_{ni} = \beta_{\widehat{\delta}} \widehat{\delta}_{ni} + \beta_p p_{ni} + \boldsymbol{\beta}_i' \mathbf{x}_{ni} + \varepsilon_{ni} \tag{3.5}$$

where

- p_{ni} is the price associated with alternative i for individual n,
- IV_{ni}^k the k^{th} instrumental variables included in the price equation for alternative i for individual n,
- α_k is the coefficient associated with the k^{th} instrumental variable,
- $\boldsymbol{\gamma}_i'$ is the transpose of the vector of coefficients associated with all exogenous regressors, excluding price, from Stage 1,
- $\widehat{\delta}_{ni}$ is the difference between actual and predicted prices from Stage 1, $p_{ni} - \widehat{p}_{ni}$,
- β_p is the coefficient associated with price from Stage 2,
- $\beta_{\widehat{\delta}}$ is the coefficient associated with the difference between actual and predicted prices,
- $\boldsymbol{\beta}_i'$ is the transpose of the vector of coefficients associated with all other exogenous regressors, excluding price, from Stage 2.

We performed two diagnostic tests: an endogeneity test of endogenous regressors and a test for instrument validity. The first tests the null hypothesis that price can be treated as an exogenous regression using the t-statistic associated with the residual from Equation 3.4. If the t-statistic is significant at the 0.05 level the null hypothesis is rejected, indicating that price should be treated as endogenous. We test the null hypothesis that the set of instruments are valid (uncorrelated with the error term) and correctly excluded from the demand model using the Direct Test for discrete choice models proposed by Guevara (Guevara and Ben-Akiva (2006), Guevara (2010)). To use the Direct Test, an additional (or auxiliary) discrete choice model is estimated; this auxiliary model is identical to the one used in Equation 3.5 but includes $k-1$ instruments. The log-likelihood (LL) values between these two models is small, the null hypothesis is rejected, indicating the instruments are valid. The intuition behind this test is as follows: if the instruments are correlated with price but not demand, then the inclusion of any instrument as an additional variable into the corrected Stage 2 model, Equation 3.5, should produce a nonsignificant increase in the log-likelihood variable. Due to identification restrictions, only $k-1$ of the k instruments can be included in the auxiliary discrete choice model. Formally, given k instruments,

$$S_{Direct} = -2(LL_{CF} - LL_{auxiliary}) \sim \chi^2_{NR,0.05} \tag{3.6}$$

where the number of restrictions (NR) is equal to $k-1$ and the significance level of 0.05 is used. Given two instruments, the difference in log-likelihood values between the two discrete choice models can be at most 3.84.

3.4 Model results

Model results[10] for our MNL models are shown in Table 3.5. Coefficients for carrier preference are suppressed for confidentiality reasons and coefficients for time of day preferences are suppressed for presentation purposes. The table includes two MNL models: the first does not account for price endogeneity whereas the second model does. Our presentation of results is organized into two sections. The first section provides behavioral interpretations for non-price attributes and the second focuses on pricing results.

3.4.1 Interpretation of non-price estimates

The results of the MNL itinerary choice models shown in Table 3.5 are intuitive, and interpretation of coefficients for non-price estimates is similar between the two models. Individuals strongly prefer nonstop itineraries

[10] All estimations in this dissertation were performed in Larch (Newman (2016)) on a Dell PowerEdge R710 Dual Intel Xeon E5520 with 2.4GHz CPUs and 288GB of memory.

	Base Model	Control Function	
Variable	**Parameter**	**Parameter**	Δ
Average high yield fare ($)	**-0.00254***	**-0.00332***	**-30.80%**
Average low yield fare ($)	**-0.00492***	**-0.00657***	**-33.50%**
Elapsed time (min)	**-0.00403***	**-0.00387***	3.90%
Number of connections	**-2.59***	**-2.66***	-3.10%
Direct flight	**-2.08***	**-2.14***	-2.80%
Regional jet or propeller (ref.)	**0**	**0**	
Wide- or narrow-body	**0.411***	**0.377***	-8.30%
Online (ref.)	**0**	**0**	0.00%
Codeshare	**0.272***	**0.284***	4.30%
Interline	**-0.248***	**-0.185***	25.50%
$\widehat{\delta}$ (residuals)	**-**	**0.00161***	
LL(0)	-32,652,846.05	-32,652,846.05	
Final LL	**-26,258,983**	**-26,253,933**	
Adj. ρ^2	0.196	0.196	

Table 3.5: Model results

* Significant at 0.01 level, LL= log likelihood, Adj. ρ^2 = 1 - (Final LL - #Attributes) / LL(0)

and have a slight preference for direct itineraries compared to connecting itineraries. In terms of equipment type, individuals prefer larger aircraft over regional jets and propeller aircraft. The marketing relationship variables are also intuitive and reveal the benefits of codeshare agreements. Itineraries sold by multiple carriers via code share agreements are more likely to be purchased than itineraries sold by a single airline (or as an online itinerary). In this sense the marketing relationships are capturing a level of advertising presence. As expected, interline itineraries are the least preferred type of itinerary (as these involve the lowest level of coordination in baggage, ticketing, and other services across flight legs that are operated by different carriers.

Departure times of day preferences are also intuitive. Figure 3.1 shows the results of the departure times of day preferences for one (out of the ten) segments, specifically for itineraries less than 600 miles that travel westbound and cross one time zone. Model results for all 10 segments are shown in Appendix A2. The curves for Monday to Friday departures show distinct morning and evening peak preferences. These peaks differ depending on itinerary type. For example, the morning peak is strongest for outbound departures (particularly for those on Monday, Tuesday, and Wednesday). The afternoon peak is strongest for inbound itineraries (particularly for the Wednesday and Thursday departures). These preferences are consistent with people who travel for business (who can depart early in the morning, gain one hour after traveling westbound, and arrive to a meeting early in the day and then return home later in the week). Departure time preferences for Saturday are similar with a strong morning peak for outbound departures (likely corresponding to the start of leisure trips). Departure time preferences for Sunday are the weakest, but show a slight preference for Sunday evening departures (likely corresponding to the return of leisure trips and/or the beginning of a weekly business trip). Finally, the time of day preferences for one-way itineraries are not as strong as those for outbound and inbound itineraries (and typically fall between the two curves). This is expected, as the one-way itineraries may represent either the outbound or inbound portion of a trip (but is unknown to the researcher). Similar patterns are observed for different segments, although the exact interpretation and peak periods differ depending on the segment.

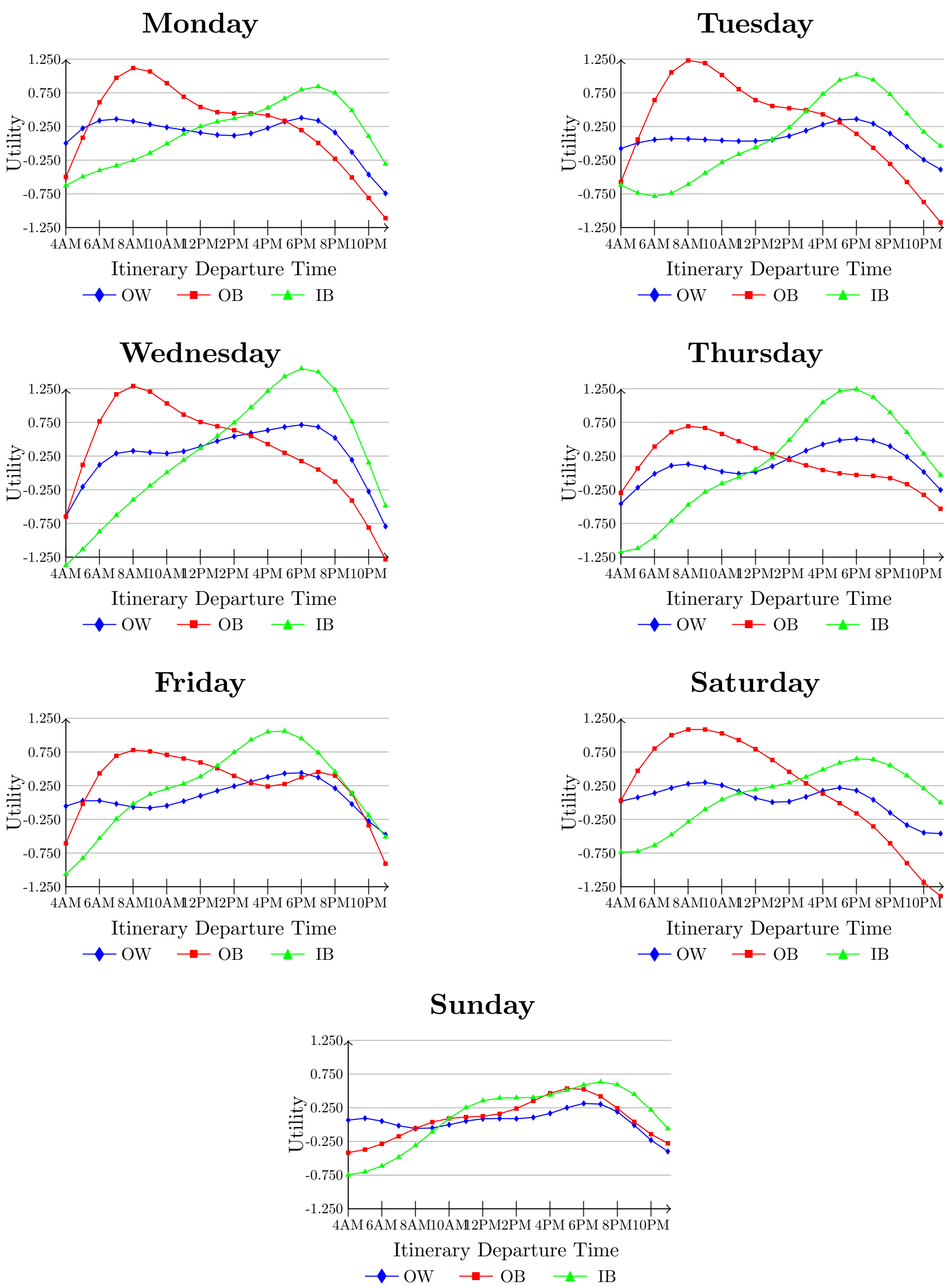


Figure 3.1: Departure time preferences: One TZ WB, distances ≤ 600 mi. Key: TZ = Time Zone, WB = Westbound, OW = One-way, OB = Outbound, IB = Inbound

3.4.2 Interpretation of price estimates

Table 3.5 compares a base MNL model that does not correct with price endogeneity with a MNL that controls for price endogeneity using a control function. As described in the methodology, we performed several statistical tests to ensure our instruments are valid (i.e., correlated with price) and strong (i.e., not correlated with itinerary choices).

The results of the first-stage OLS regression for the two price instruments we used to control for endogeneity indicated that the parameter estimates associated with both instruments are significantly different from zero at the 99% confidence interval level (p-value < 0.001). In addition, the first-stage F-statistic of is well above the critical value of 10, recommended as a rule of thumb by Staiger and Stock (1997)[11]. Finally, the R^2 of the regression is equal to 0.3979. We conclude from these statistical tests that both instruments are relevant.

Next, the residuals (δ) from Equation (3.4) are retained and included without transformation as an additional variable in the utility function of the itinerary choice model. As shown in Table 3.5 (under the "Control Function" model), the parameter estimate associated with this residual is statically significant at the 99% confidence level (p-value <0.01), which confirms the presence of endogeneity, and specifically that our instruments are correlated with price and are thus valid.

As a final test, we need to show that our instruments are exogenous, i.e., that they are uncorrelated with the decision to buy an itinerary. Using the Direct Test proposed by Guevara (Guevara and Ben-Akiva (2006), Guevara (2010)), we estimate a third choice model including the residuals and one of the instruments (the Hausman-type price instrument in our case) as additional variables and retrieve the log-likelihood of this auxiliary model. We then compare the difference in loglikelihood values between the corrected model and the auxiliary model. We find that:

$$S_{Direct} = -2(-26,253,932.77 - (-26,253,932.22)) = 1.1 < \chi^2_{1,0.05} = 3.84. \quad (3.7)$$

We therefore conclude that our instruments are exogenous, or are valid

[11] Staiger and Stock (1997) have focused on the 2SLS method but Guevara-Cue and Navarro (2013) suggest that similar thresholds are applicable in the case of the CF in logit models.

instruments.

Tables 3.6-3.8 demonstrate the importance of correcting for price endogeneity. Table 3.6 shows the value of times associated with the base model and the model that corrects for price endogeneity using a control function. The values of time are overestimated in the uncorrected model: \$95.30/hr for high-yield and \$49.10/hr for low-yield products compared to \$70.00/hr and \$35.40/hr, respectively in the corrected model. This overestimate can lead to sub-optimal business decisions. For example, a carrier that uses the uncorrected itinerary choice model would overestimate customers' willingness to pay for a new aircraft that reduces flight times. This could, in turn, lead to overinvestment in capital expenditures in new aircraft.

Tables 3.7 and 3.8 show price elasticity estimates for the high yield and low yield products, respectively. These differences are economically important. In Table 3.7, the segments for which elasticity flips from inelastic (greater than -1.0) to elastic (less than -1.0), can lead to completely opposite effects than which were intended by firms. For example, in the "Same Time Zone, distance > 600 miles." segment, we report a mean fare of \$207.96 and a base model elasticity of -0.7712. Using simple first principles and basic economic theory, given this inelasticity a firm could raise price, quantity demanded would decline (as would total costs) but total revenues would in fact increase. Such a move, with certainty, would increase economic profits. However, results from the control function indicate that low-yield products on that segment are in fact elastic. As such, a price increase would cause total revenues to decline; quantity demand would also decline (as would total cost). The resulting impact on economic profits is now uncertain.

Value of Time	**Base Model**	**Control Function**
High-yield (\$/hr)	95	70
Low-yield (\$/hr)	49	35

Table 3.6: Value of time results

Again, using first principles, managers should never lower price on inelastic consumers, as this will, with certainty, lead to lower revenues. In contrast, lowering price on elastic consumers will result in increased revenues. In Table 3.7 eight segments are incorrectly identified by the basic model as being inelastic when in reality low-yield products are elastic. For these segments, managers would incorrectly assume that they should not decrease price in

	Low-Yield Products		
Segment	Mean fare	Base model elasticities	Control function elasticities
Same TZ, distance ≤ 600 mi.	221.42	-0.8205	**-1.095**
Same TZ, distance > 600 mi.	207.96	-0.7712	**-1.029**
One TZ WB, distance ≤ 600 mi.	221.32	-0.8084	**-1.079**
One TZ WB, distance > 600 mi.	248.76	-0.9374	**-1.251**
One TZ EB, distance ≤ 600 mi.	219.15	-0.8216	**-1.096**
One TZ EB, distance > 600 mi.	251.80	-0.9534	**-1.272**
Two TZ WB	265.62	-0.9959	**-1.328**
Two TZ EB	263.82	-0.9923	**-1.323**
Three TZ WB	289.58	-1.111	**-1.481**
Three TZ EB	290.41	-1.155	**-1.541**
Average for All Segments	240.20	-0.8938	**-1.193**

Table 3.7: Price elasticities for low-yield products
Key: TZ = Time Zone, WB = Westbound, IB = Inbound

	High-Yield Products		
Segment	Mean fare	Base model elasticities	Control function elasticities
Same TZ, distance ≤ 600 mi.	290.73	-0.5265	**-0.6855**
Same TZ, distance > 600 mi.	293.41	-0.4951	**-0.6419**
One TZ WB, distance ≤ 600 mi.	320.53	-0.5742	**-0.7461**
One TZ WB, distance > 600 mi.	329.03	-0.5725	**-0.7434**
One TZ EB, distance ≤ 600 mi.	315.33	-0.5755	**-0.7464**
One TZ EB, distance > 600 mi.	344.59	-0.5925	**-0.7717**
Two TZ WB	364.98	-0.6239	**-0.8153**
Two TZ EB	347.66	-0.5761	**-0.7477**
Three TZ WB	503.74	-0.8412	**-1.115**
Three TZ EB	498.00	-0.8649	**-1.145**
Average for all Segments	343.60	-0.5867	**-0.7652**

Table 3.8: Price elasticities for high-yield products
Key: TZ = Time Zone, WB = Westbound, IB = Inbound

those markets. We see similar trends in Table 3.8 with the "3 Time Zone Westbound" and "3 Time Zone Eastbound" segments. In general, the results in Table 3.8 demonstrate that high-yield products are not as inelastic as predicted by the base model. In other words, consumers are more price sensitive. Again, these results are economically meaningful. For example, the base model reports an average elasticity of -0.5867; a 10% increase in price will lead to a 5.87% decline in quantity demanded. In contrast, our control function results report an elasticity of -0.7652; a 10% increase in price will lead to a 7.65% decline in quantity demanded. This suggests that managers or revenue models would underestimate the impact on quantity demanded by approximately 1.8%.

3.5 Limitations, contributions, and future research

To the best of our knowledge, this is the first study to control for price endogeneity for an itinerary choice model that is representative of those currently used in practice. Our model suffers from the same data limitations faced by industry. Our sample is non-representative in the sense that low cost carriers are under-represented. We are therefore implicitly assuming that those customers who purchase tickets on low-cost carriers have similar itinerary preferences as those who purchase on major carriers. Our ticketing database provides no information about the customers who purchased the ticket, preventing us from examining differences in preference based on trip purpose and socio-economic factors. The lack of information about customers also prevents us from modeling schedule delay, defined as the difference between an individual's preferred departure time and the scheduled departure time of an itinerary. We also assume that customer preferences and competition among alternatives can be represented using a MNL model (which is the most common model used in practice); however more advanced discrete choice models that allow for random coefficients and different substitution patterns across product dimensions are clearly desirable[12].

Nonetheless, our analysis provides an important contribution by demonstrating how models representative of those currently used in practice can be enhanced to correct for price endogeneity. Our results show that failure to account for price endogeneity leads to over-estimation of customers' value

[12]Linking socioeconomic data with consumer preferences derived from survey data into a random coefficient model is the focus of my post-doctoral research project, described in Chapter 5.

of time. This can lead to sub-optimal business decisions, e.g., a carrier that uses the uncorrected itinerary choice model would over-estimate customers' willingness to pay for a new aircraft that reduces flight times (and potentially over-invest in new aircraft). A second main contribution is that it is the first study to estimate highly refined departure time of day preferences. The price elasticity and departure time of day preferences results are not restricted to itinerary choice modeling applications, and can help support evaluation of proposed airport fees and taxes, national departure and emission taxes, landing fees, and congestion pricing policies.

There are several research extensions. As part of our analysis, we used an average price variable similar to that used by industry. However, prior research has shown that customers' price sensitivities vary as a function of how far in advance a ticket is purchased. Prior research (e.g., Coldren and Koppelman (2005b)) has also shown that there are potentially many layers of correlation within and across product attributes, with relationships extending across airline, time of day, level of service (e.g., nonstop versus connecting), and potentially other dimensions. Replacing the MNL with a simple nested logit choice model or a more complex but flexible generalized extreme value (GEV) model is another potential research direction. In particular, it would be interesting to compare if the substitution patterns observed by Coldren and Koppelman (2005b) using data from 2000 are also observed on more recent data. It would also be interesting to determine if other product dimensions they did not consider for nesting (such as high-yield versus low-yield product distinctions) or that they found to be insignificant (such as level of service) are important to incorporate for models based on more recent data. These extensions are the subject of the next chapter.

Finally, it would be interesting to compare the results from the "baseline" model we developed in this research that corrects for price endogeneity to one that incorporates advanced modeling techniques found in the economic welfare estimation literature. For example, Armantier and Richard (2008) propose a method to account for the non-random nature of data available for estimating airline itinerary choice models using distributions from publicly available data such as DB1B (U.S. Department of Transportation (2016)). As always, much remains to be done.

Appendices

A1 Example of choice set generation process

An example illustrating the process we used to generate choice sets is shown in Tables 3.9 and 3.10. Table 3.9 contains five unique itineraries from ATL to SEA for Tuesday departures in May of 2013. The final choice set, shown in Table 3.10, contains five itineraries. The rows from Table 3.9 that were used to populate schedule attributes (with the exception of marketing relationships and passenger counts) are highlighted. For itineraries 1, 2, and 4 the date falling in the representative week (May 14) is used to populate schedule attributes whereas for itineraries 3 and 5 the first date that itinerary was purchased is used since there are no purchases that occurred on May 14.

The number of passenger and marketing type associated with itinerary q in the final choice set are calculated using information from all rows in Table 3.9 associated with itinerary q. For example, the total number of passengers who purchase itinerary 1 is 23. The marketing type for itinerary 2 is online because the marketing carriers and operating carriers are always the same for all rows associated with itinerary 2. The marketing type associated with itinerary 1 in the final choice set is a codeshare, because two tickets for travel on May 28 for Alaska operated flight 938 were sold by AA. The marketing type for itinerary 4 is an interline because the marketing carriers for leg1 and leg2 differ.

A2 Continuous departure time of day preferences for continental U.S. airline markets segmented by distance, direction of travel, number of time zones, day of week and itinerary type

The size and comprehensiveness of our database allowed us to estimate highly refined continuous departure time of day preference curves that account for distance, direction of travel, the number of time zones traversed, departure day of week and itinerary type (outbound, inbound or one-way). This appendix contains the results of all model coefficients (including the 1260 time of day parameters) and summarize results in a series of ten figures. An accompanying Excel spreadsheet is located at http://garrowlab.ce.gatech.edu. These highly-refined time of day preference curves can be used by airlines, researchers, and government organizations in

Itin			Leg 1							Leg 2					
#	# Pax	Mkt Type	Org	Dst	Op Carr	Mkt Carr	Op Flt	Dept Date	Dept Time	Org	Dst	Op Carr	Mkt Carr	Op Flt	Dept Time
1	6	Online	ATL	SEA	AS	AS	938	05-Jul	8:16						
1	3	Online	ATL	SEA	AS	AS	938	May-14	8:16						
1	6	Online	ATL	SEA	AS	AS	938	May-21	8:16						
1	2	Online	ATL	SEA	AS	AS	938	May-28	8:16						
1	6	CShare	ATL	SEA	AS	AA	938	May-28	8:16						
2	8	Online	ATL	SEA	DL	DL	319	Jul-05	10:10						
2	5	Online	ATL	SEA	DL	DL	319	May-14	10:15						
2	3	Online	ATL	SEA	DL	DL	319	May-21	10:10						
3	1	Online	ATL	JFK	DL	DL	688	Jul-05	8:05	JFK	SEA	DL	DL	417	11:23
4	2	ILine	ATL	PHX	DL	DL	545	May-21	9:15	PHX	SEA	WN	WN	2849	13:30
4	1	ILine	ATL	PHX	DL	DL	545	May-28	9:20	PHX	SEA	WN	WN	2849	13:30
5	2	Online	ATL	SLC	DL	DL	1278	Jul-05	12:10	SLC	SEA	DL	DL	784	15:25
5	1	Online	ATL	SLC	DL	AF	1278	May-14	12:20	SLC	SEA	DL	AF	784	15:25
5	1	CShare	ATL	SLC	DL	KL	1278	May-21	12:20	SLC	SEA	DL	KL	784	15:25
5	1	CShare	ATL	SLC	DL	DL	1278	May-28	12:10	SLC	SEA	DL	DL	784	15:25

Table 3.9: Example of itineraries departing on tuesdays from ATL-SEA
Key: pax = passengers, Mkt Type = type of marketing relationship, Org = origin, Dst = destination, Op Carr = operating carrier, Mkt Carr = marketing carrier, Op Flt = operating flight number, Dept Date = departure date, Dept Time = departure time, ILine = interline, CShare = codeshare

Itin			Leg 1							Leg 2					
#	# Pax	Mkt Type	Org	Dst	Op Carr	Mkt Carr	Op Flt	Dept Date	Dept Time	Org	Dst	Op Carr	Mkt Carr	Op Flt	Dept Time
1	23	CShare	ATL	SEA	AS	AS	938	May-14	8:16						
2	16	Online	ATL	SEA	DL	DL	319	May-14	10:15						
3	1	Online	ATL	JFK	DL	DL	688	Jul-05	8:05	JFK	SEA	DL	DL	417	11:23
4	3	ILine	ATL	PHX	DL	DL	545	May-21	9:15	PHX	SEA	WN	WN	2849	13:30
5	5	CShare	ATL	SLC	DL	AF	1278	May-14	12:20	SLC	SEA	DL	AF	784	15:25

Table 3.10: Example of choice set for itineraries departing on tuesdays from ATL-SEA
Key: pax = passengers, Mkt Type = type of marketing relationship, Org = origin, Dst = destination, Op Carr = operating carrier, Mkt Carr = marketing carrier, Op Flt = operating flight number, Dept Date = departure date, Dept Time = departure time, ILine = interline, CShare = codeshare

the evaluation of demand-management and other policies.

Figures 3.2- 3.11 show the results of the departure time of day preferences for each segments. The accompanying Excel spreadsheet contains the parameter estimates and t-statistics for the 1260 time of day coefficients (10 segments $\times$ 6 sine/cosine terms $\times$ 3 itinerary types $\times$ 7 departure days = 1260 parameters).

Results are intuitive and show that departure time of day preferences vary across many dimensions. For example, Figure 3.2 shows the results of the departure time of day preferences for those itineraries in the same time zone with distances $\leq$ 600 miles. The curves show distinct departure time preferences by day of week and itinerary type. There is a strong preference for morning flights for outbound itineraries departing Monday, Tuesday and Wednesday and a strong preference for evening flights for inbound itineraries returning Wednesday or Thursday and to a lesser extent on Friday. These patterns likely correspond to people traveling for business who depart early in the week on morning flights and return later in the week on evening flights. Departure time preferences are not as pronounced on Saturday, although Saturday does exhibit a preference for morning departures for outbound itineraries, likely corresponding to the start of leisure trips. Departure time preferences are least pronounced for Sunday. The time of day preferences for oneway itineraries are not as strong as those for outbound and inbound itineraries (and typically fall between the two curves). This is expected, as the one-way itineraries may represent either the outbound or inbound portion of a trip (but is unknown to the researcher).

Similar patterns are observed for different segments, although the exact interpretation and peak periods differ depending on the segment. In general, the shorter the itinerary and the fewer time zones crossed, the stronger the departure time of day preferences are (note the scale of the y-axis is the same for all figures). For example, departure time of day preferences are not as strong for itineraries traveling two or three time zones. This is likely due to increased travel times (combined with the loss of hours due to time zone changes traveling eastbound) that make it difficult to depart in the morning and arrive for an early morning meeting (outbound) or return at the end of the day and arrive home at a reasonable hour (inbound). A similar pattern is observed for itineraries that are in the same time zone or one time zone apart in the sense that those itineraries that correspond to shorter distances ($\leq$ 600 miles) have stronger departure time of day preferences than longer

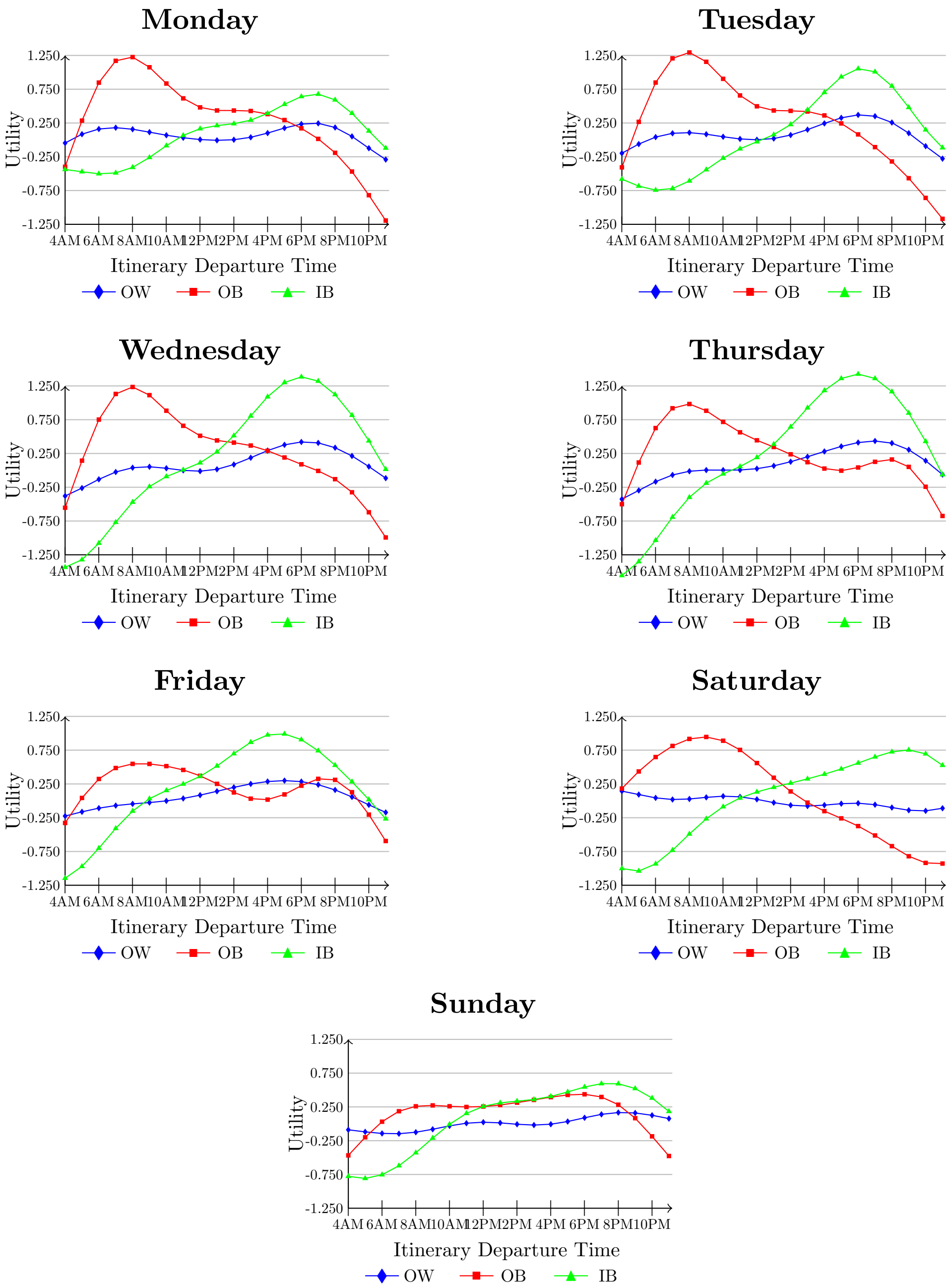


Figure 3.2: Departure time preferences: Same TZ , distances $\leq$ 600 mi.
Key: TZ = Time Zone, OW = One-way, OB = Outbound, IB = Inbound

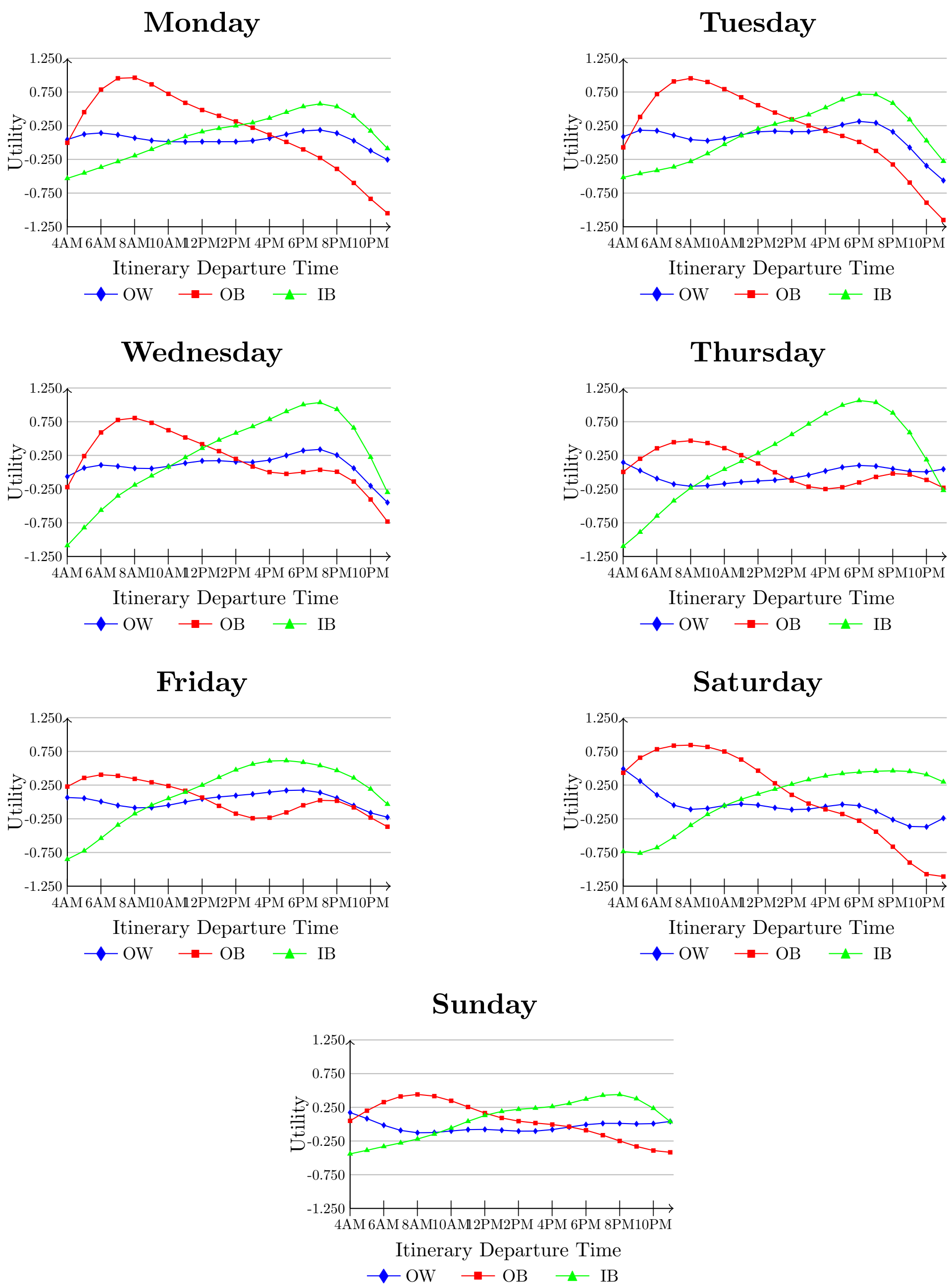


Figure 3.3: Departure time preferences: Same TZ , distances > 600 mi.
Key: TZ = Time Zone, OW = One-way, OB = Outbound, IB = Inbound

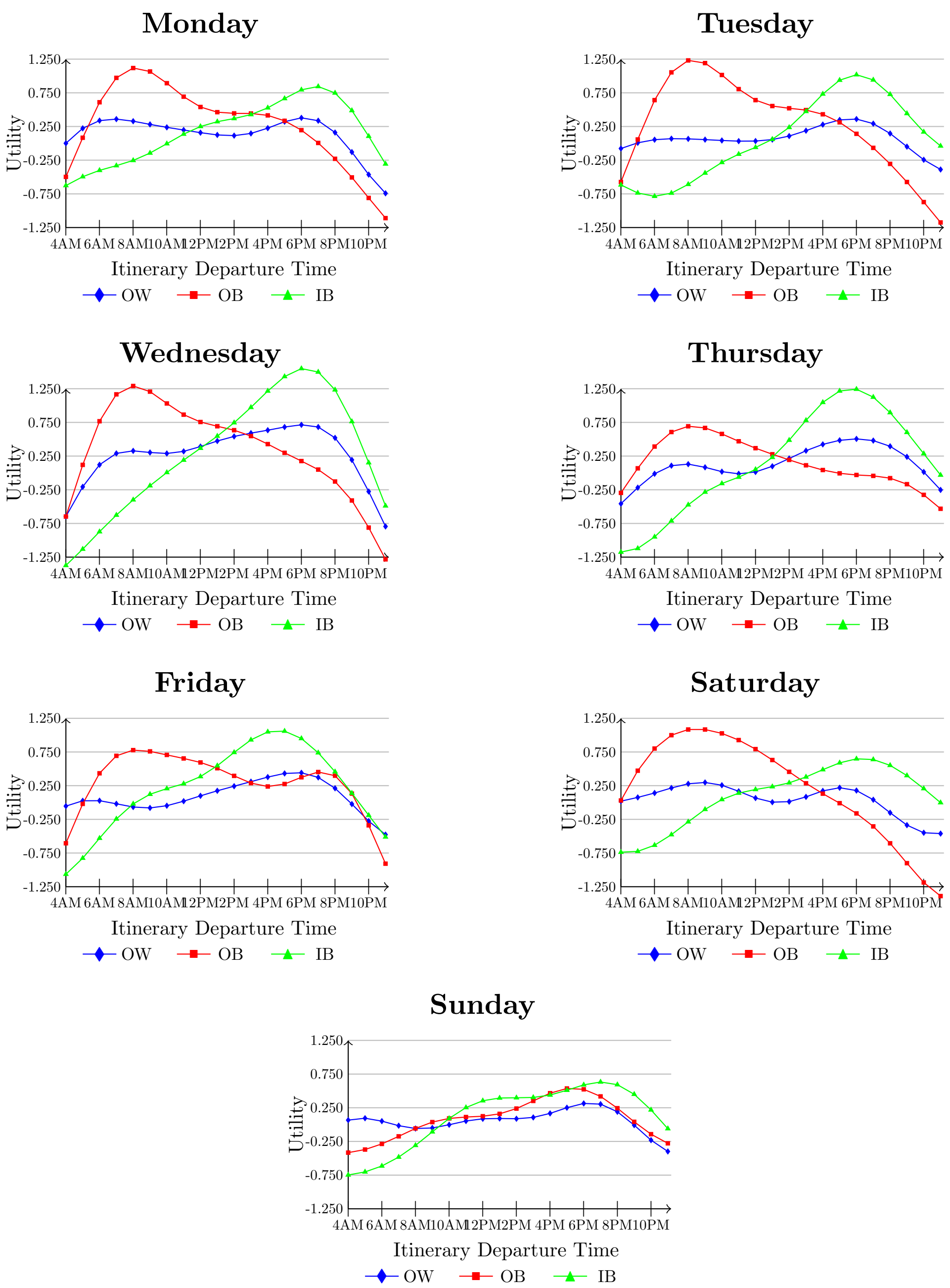


Figure 3.4: Departure time preferences: One TZ WB , distances $\leq$ 600 mi. Key: TZ = Time Zone, WB = Westbound, OW = One-way, OB = Outbound, IB = Inbound

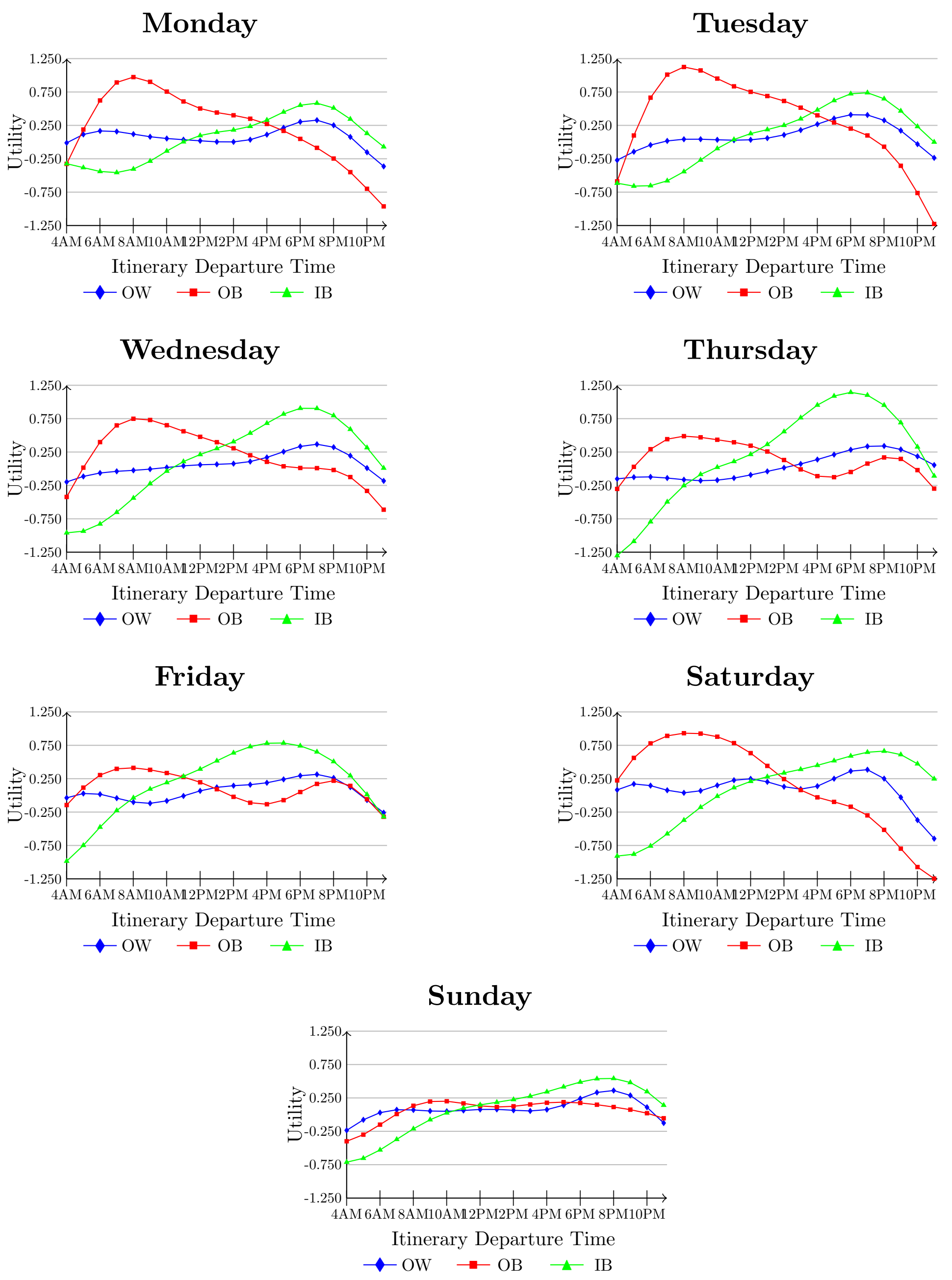


Figure 3.5: Departure time preferences: One TZ WB , distances > 600 mi. Key: TZ = Time Zone, WB = Westbound, OW = One-way, OB = Outbound, IB = Inbound

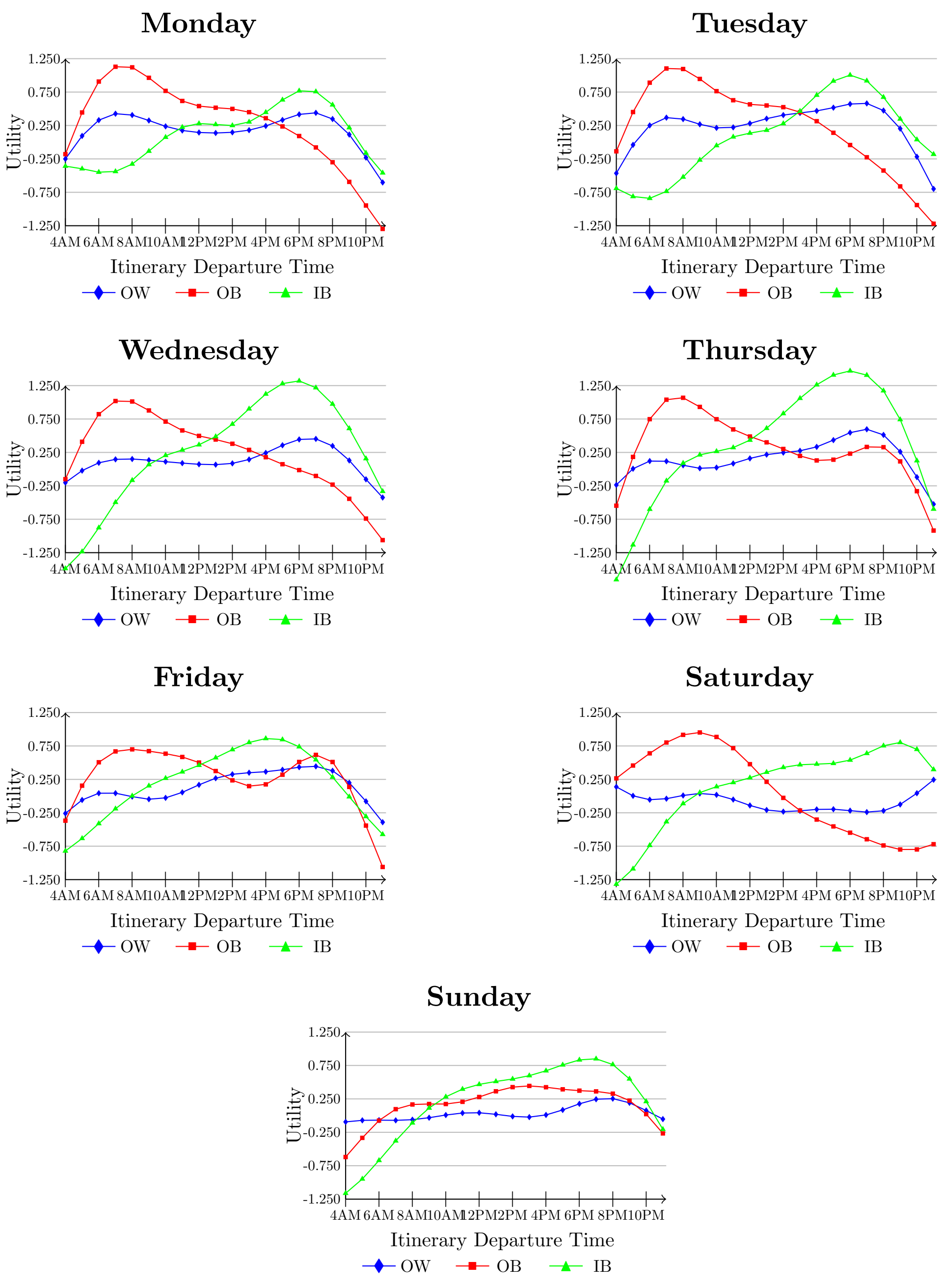


Figure 3.6: Departure time preferences: One TZ EB , distances ≤ 600 mi. Key: TZ = Time Zone, EB = Eastbound, OW = One-way, OB = Outbound, IB = Inbound

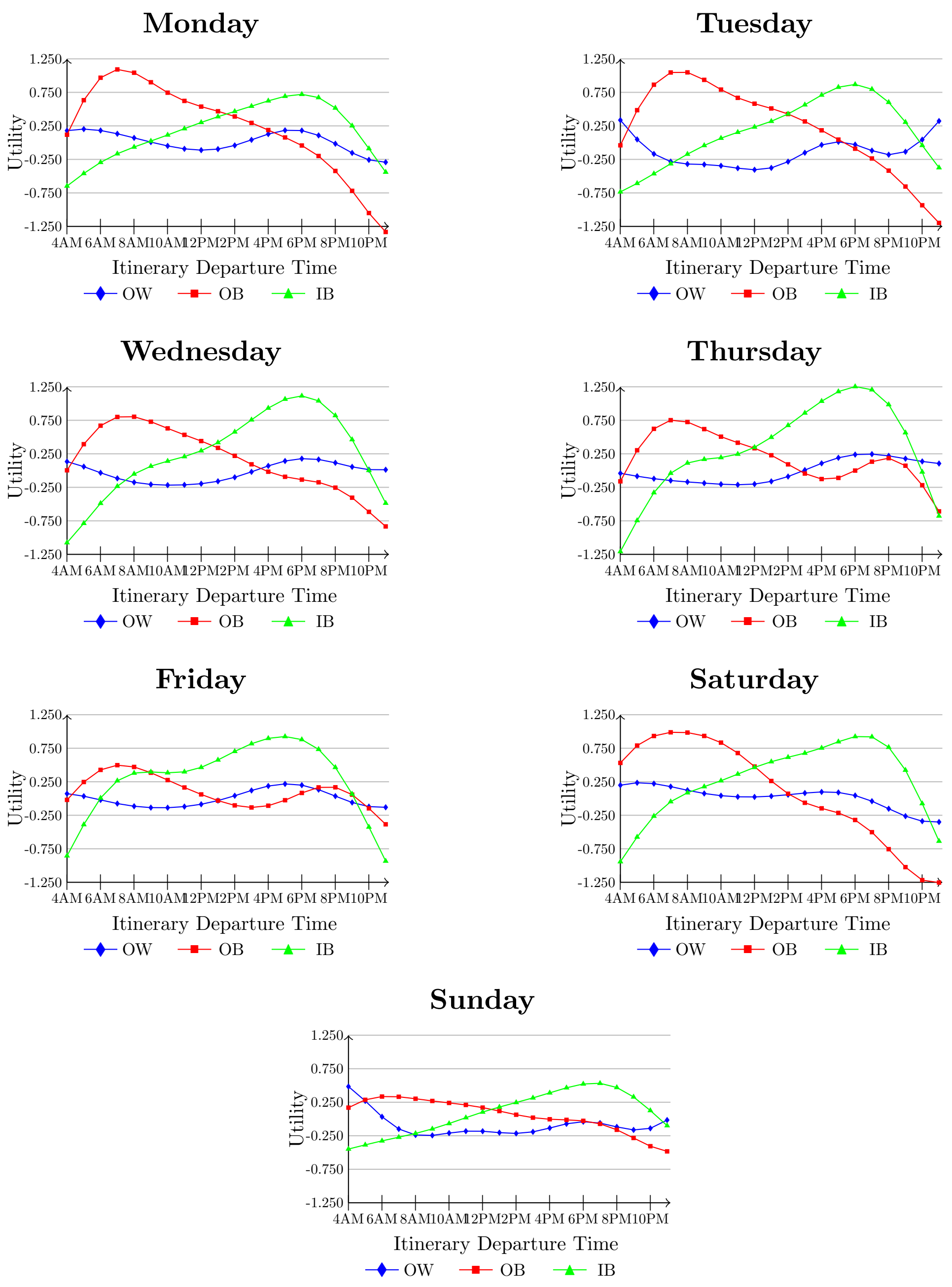


Figure 3.7: Departure time preferences: One TZ EB , distances > 600 mi. Key: TZ = Time Zone, EB = Eastbound, OW = One-way, OB = Outbound, IB = Inbound

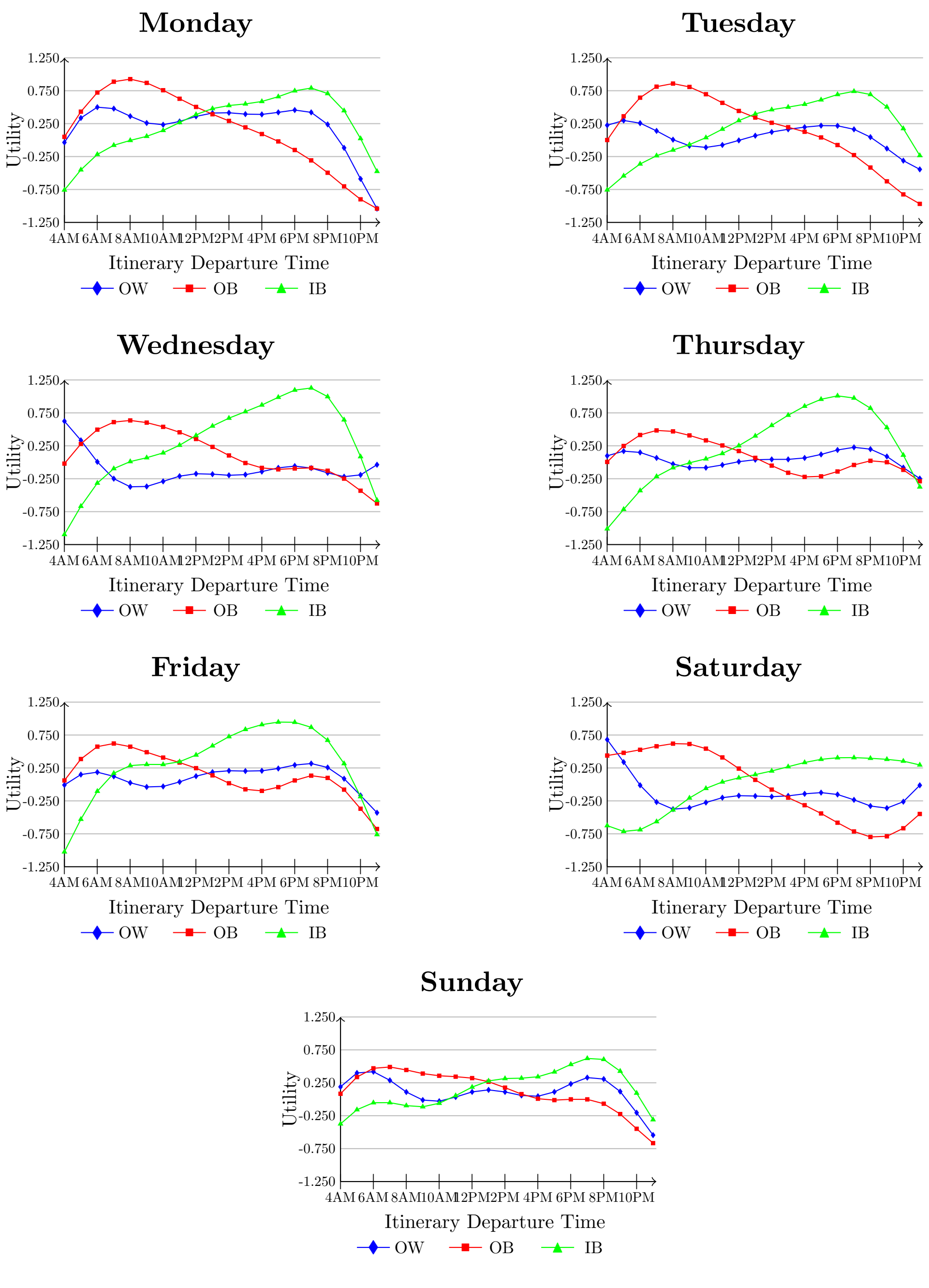


Figure 3.8: Departure time preferences: Two TZ WB
Key: TZ = Time Zone, WB = Westbound, OW = One-way, OB = Outbound, IB = Inbound

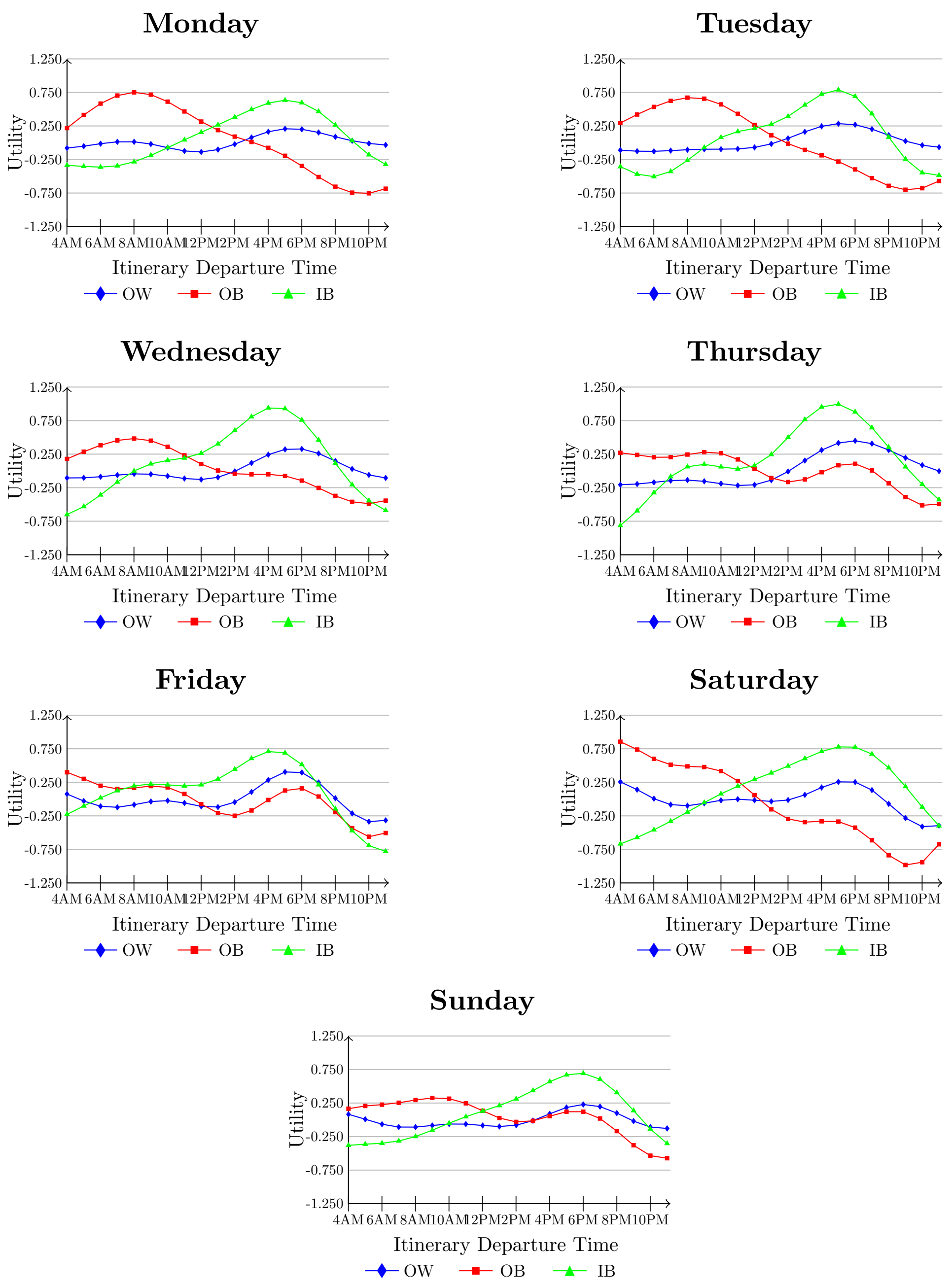


Figure 3.9: Departure time preferences: Two TZ EB
Key: TZ = Time Zone, EB = Eastbound, OW = One-way, OB = Outbound, IB = Inbound

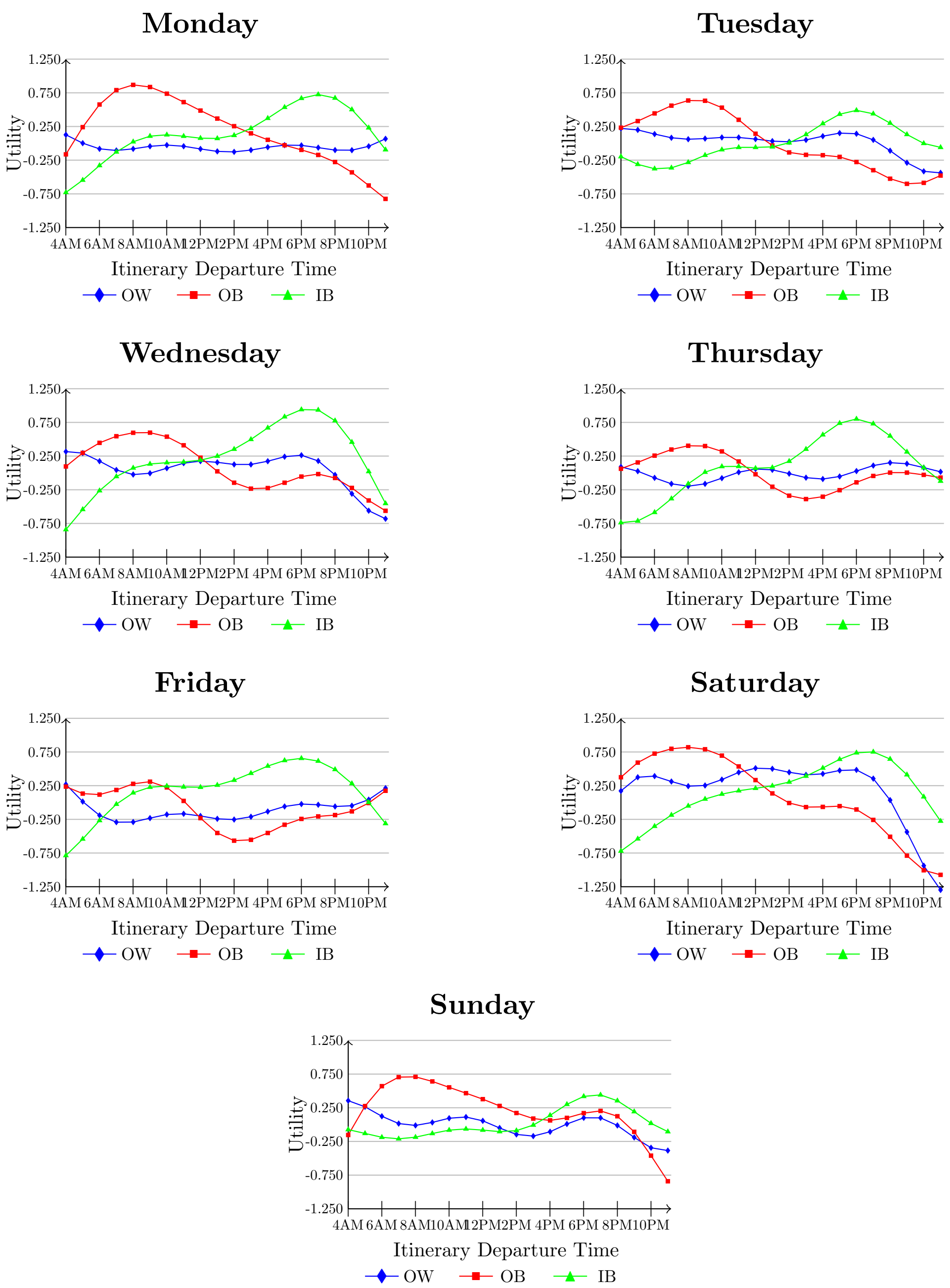


Figure 3.10: Departure time preferences: Three TZ WB
Key: TZ = Time Zone, WB = Westbound, OW = One-way, OB = Outbound, IB = Inbound

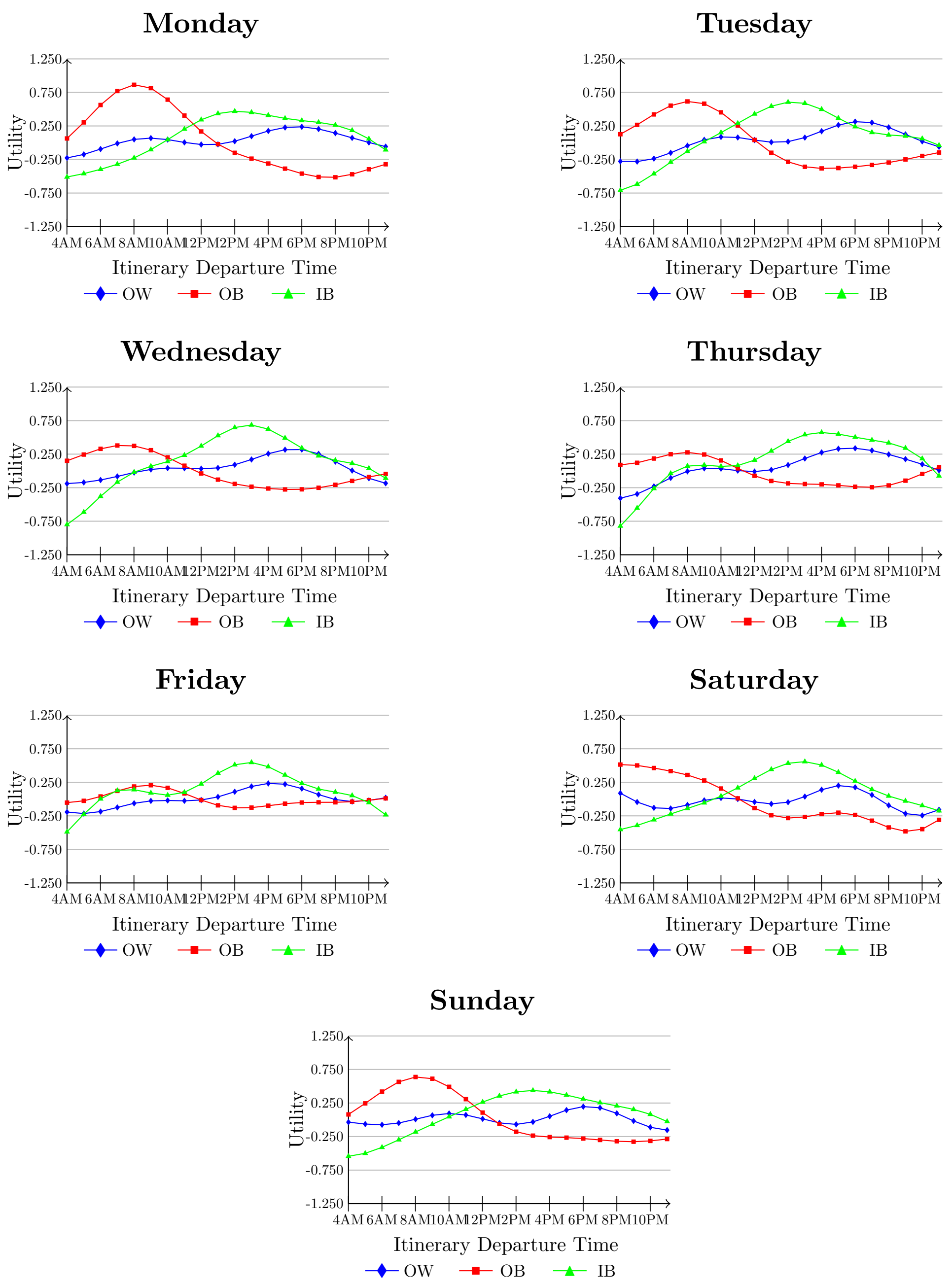


Figure 3.11: Departure time preferences: Three TZ EB
Key: TZ = Time Zone, EB = Eastbound, OW = One-way, OB = Outbound, IB = Inbound

itineraries.

A3 A comparison of departure time of day formulations

Lurkin V., Garrow A.L., Higgins M.J., Newman J.P, and Schyns M. (2016). Departure Time of Day Formulations. *Under review in Journal of Air Transport Management.*

This appendix compares three different methods that have been used to model departure time of day preferences. The first is a discrete formulation that uses indicator variables to represent the hour of departure. The next two methods are based on a continuous formulation that uses a series of sine and cosine functions. One assumes departure time preferences over a 24-hour cycle and the other uses shorter cycle lengths that account for fewer departures during certain hours of the day. We compare models using itineraries in the Continental U.S. that are separated by two time zones. Although the discrete formulation fits the data better, the two continuous time of day formulations are preferred as they provide more intuitive predictions and require fewer parameters. Results between the two continuous time of day formulations are similar but differ in how strongly they weight itineraries that depart very early or very late in the day. Based on empirical results, we recommend testing both 24-hour and less than 24-hour cycle lengths for a particular dataset.

The discrete formulation approach uses a set of categorical variables to represent non-overlapping departure time periods, e.g., one variable for each departure hour. However, the use of categorical variables can be problematic for forecasting applications when the difference in coefficients associated with two consecutive time periods is large. In these situations, moving a flight by a few minutes from one category to another can result in unrealistic changes in demand predictions. The second approach overcomes this limitation by modeling departure times using the initial part of a Fourier sine and cosine series, a formulation first proposed for urban travel by Abou-Zeid et al. (2006) and applied to airline itinerary choice by Koppelman et al. (2008).

$$\mu_i = ... + \beta_1 \sin\left(\frac{2\pi t_i}{1440}\right) + \beta_2 \cos\left(\frac{2\pi t_i}{1440}\right) + \beta_3 \sin\left(\frac{4\pi t_i}{1440}\right) + \beta_4 \cos\left(\frac{4\pi t_i}{1440}\right) + ... \quad (3.8)$$

where μ_i is the utility associated with itinerary i, t_i is the departure time of the first flight in itinerary i expressed as minutes past midnight, β_1 is the estimated model coefficient, and 1440 is the number of minutes in the day, ensuring that the time of day components repeat cyclically each day. In practice, four to six parameters are typically used. This formulation also avoids undesirable boundary jumping changes in utility associated with shifting a departure time by one minute into or out of a discrete departure time period.

However, to account for the low attractiveness of some periods of the day and/or periods in which little to no flights departure, Carrier (2008) proposed the use of a continuous formulation that has a cycle length d that is less than 24 hours (1440 minutes) and has a start time of s as:

$$\begin{aligned}\mu_i = ... + \beta_1 \sin\left(\frac{2\pi(t_i - s)}{d}\right) + \beta_2 \cos\left(\frac{2\pi(t_i - s)}{d}\right) + \beta_3 \sin\left(\frac{4\pi(t_i - s)}{d}\right) \\ + \beta_4 \cos\left(\frac{4\pi(t_i - s)}{d}\right) + ...\end{aligned} \tag{3.9}$$

where
$l - e \leq d \leq 1440$
$0 \leq s \leq e$

with l and e as the latest and earliest departure times (in minutes) of the itineraries in the estimation dataset. In estimating this function, Carrier (2008) notes that the starting time s is arbitrary (i.e., changing the starting time does not impact the relative difference in utilities across alternatives, and therefore does not change the behavioral interpretation or model fit). The problem of interest then becomes finding the cycle length d that best fits the data. Due to the fact that Equation (3.9) is nonlinear-in-parameters, many software packages for discrete choice models cannot be used to estimate this function. In these situations, it is common to use a trial-and-error approach in which the analyst iteratively sets the cycle length d and compares the log likelihood values (a measure of model fit) of these different models to find the cycle length that fits the data the best.

The number of passengers departing each hour of the day for these various segments is shown in Table 3.11. The distribution of departure times varies by segment; however, note that almost all segments have no flights between midnight and 5 AM.

For the first part of analysis, we compared departure time of day formulations using the "Base MNL model" specification for the "2 time zones eastbound (2 TZ EB)" and "2 time zones westbound (2 TZ WB)" segments.

	Same TZ ≤ 600	Same TZ >600	1TZ WB ≤ 600	1TZ WB >600	1TZ EB ≤ 600	1TZ EB >600	2TZ WB	2TZ EB	3TZ WB	3TZ EB
12:00-12:59 AM	0	0	0	0	0	230*	0	8,883	0	10,145
1:00-1:59 AM	0	0	0	0	0	144*	0	3,666	0	2,921
2:00-2:59 AM	0	0	0	0	0	0	0	0	0	0
3:00-3:59 AM	0	0	0	0	0	0	0	0	0	0
4:00-4:59 AM	0	0	0	0	0	0	0	0	0	0
5:00-5:59 AM	32,649	39,631	3,605	22,466	3,594	41,449	10,283	15,799	11,033	8,376
6:00-6:59 AM	138,733	155,181	16,840	118,279	26,690	148,392	59,216	82,651	42,404	77,634
7:00-7:59 AM	159,939	153,296	24,212	109,520	19,929	140,325	58,587	52,742	54,852	52,913
8:00-8:59 AM	134,329	142,565	22,316	105,586	17,772	88,735	57,943	65,276	49,383	48,384
9:00-9:59 AM	134,057	95,699	16,164	79,624	17,099	76,844	49,846	41,735	30,274	35,295
10:00-10:59 AM	127,492	124,695	23,324	76,792	24,831	119,580	31,705	60,291	20,235	33,128
11:00-11:59 AM	136,847	136,301	14,414	108,346	18,810	96,985	33,686	48,791	24,825	55,579
12:00-12:59 PM	128,622	123,586	17,325	100,362	21,342	87,767	36,762	60,992	22,731	38,486
1:00-1:59 PM	132,940	105,130	24,645	85,833	21,305	104,197	32,281	49,922	18,546	43,663
2:00-2:59 PM	135,134	117,016	13,341	101,955	12,573	78,562	38,493	39,445	22,984	27,908
3:00-3:59 PM	153,224	111,482	23,385	90,446	16,827	94,074	48,208	52,433	33,305	15,524
4:00-4:59 PM	144,133	84,680	24,983	93,039	33,995	95,462	37,015	28,793	35,923	17,828
5:00-5:59 PM	213,747	149,351	24,804	129,568	24,294	99,918	67,240	30,557	55,500	3,388
6:00-6:59 PM	154,355	122,645	19,878	98,338	21,176	78,646	52,138	16,949	41,608	1,497
7:00-7:59 PM	155,735	98,785	15,927	86,739	22,167	69,699	37,567	8,026	23,110	4,776
8:00-8:59 PM	77,819	58,214	14,971	36,794	3,651	19,140	17,114	2,057	17,564	4,852
9:00-9:59 PM	30,247	22,671	3,526	15,697	5,252	4,437	11,753	1,739	3,826	17,031
10:00-10:59 PM	28,361	7,812	2,309	7,431	957	1,895	1,829	3,671	1,243	33,983
11:00-11:59 PM	1,148	2*	150*	0	1*	326*	0	10,209	0	25,726
Total	**2,219,511**	**1,848,742**	**306,119**	**1,466,815**	**312,265**	**1,446,807**	**681,666**	**684,627**	**509,346**	**559,037**

Table 3.11: Frequency distribution of number of passengers departing per hour in May 2013
Key: TZ=time zone, EB=eastbound WB=westbound. Truncated cycle lengths corresponding to models reported in Table 3.13 are shown in shaded cells.

These segments differ in how many hours of the day there are no flight departure; the 2 TZ EB model has no flights from 2AM to 5AM and the 2 TZ WB model has no flights for a longer period of time, from 11PM to 5AM. Results are shown in Table 3.12.

Model	Final LL	LL(0)	# TOD parameters	Adj. ρ^2
Westbound (WB)				
Discrete	-1,924,237.44	-2,359,997.71	18	0.185
24 hour cycle	-1,924,792.03	-2,359,997.71	6	0.184
5 AM - 11 PM cycle	-1,925,085.45	-2,359,997.71	6	0.184
5 AM - 10 PM cycle	-1,925,176.72	-2,359,997.71	6	0.184
6 AM - 9 PM cycle	-1,925,379.04	-2,359,997.71	6	0.184
6 AM - 8 PM cycle	-1,925,599.80	-2,359,997.71	6	0.184
Eastbound (EB)				
Discrete	-1,851,635.64	-2,307,797.31	21	0.198
24 hour cycle	-1,851,833.78	-2,307,797.31	6	0.198
5 AM - 11 PM cycle	-1,851,874.18	-2,307,797.31	6	0.198

Table 3.12: Model fit statistics for two time zone EB and WB models
Key: Final LL = log likelihood at convergence, LL(0) = log likelihood at null parameters, TOD = time-of-day

The discrete formulation fits the data the best (as it has the smallest log likelihood model value) in both segments, but as shown in Figure 3.12 for the 2 TZ EB model, results in counter-intuitive changes in utility across hours. For example, moving the departure time of an itinerary from 7:59PM to 8:00PM would result in a change of utility from 0.2133 to 0.2033, leading to (a potentially large) decrease in the number of passengers predicted to choose the itinerary. The continuous time of day formulations provide a smoother transition in utilities and demand changes. As shown in Figure 3.12, the two continuous time of day formulations for the 2 TZ EB model are practically identical, with negligible differences observed in utilities associated with the itineraries departing earliest and latest in the day.

As shown in Table 3.12 and Figure 3.13, results for the 2 TZ WB model are similar to those observed for the 2 TZ EB model. For this this segment, departure time preferences are not as strong (note the scale of the y-axes differ across figures) and the shorter cycle length has shifted utilities downward. However, as noted by Carrier (2008), a shift upwards or downwards in utility curves makes no difference to model predictions or behavioral inter-

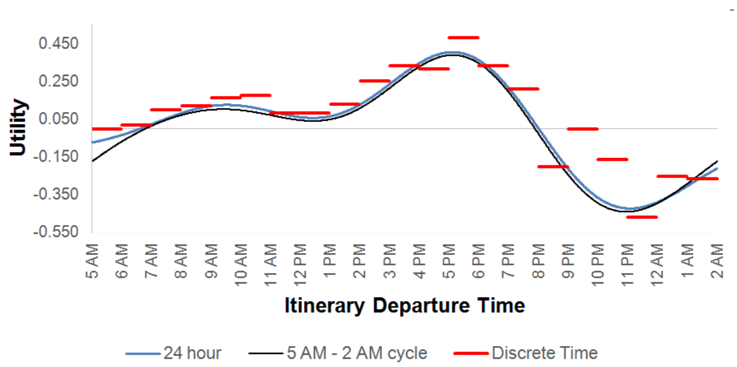


Figure 3.12: Time of day preferences for two time zone eastbound model

pretations. The shorter cycle length gives a higher preference for itineraries departing later in the day (9 PM and after). Overall, the 24-hour cycle results in the better model fit for this segment.

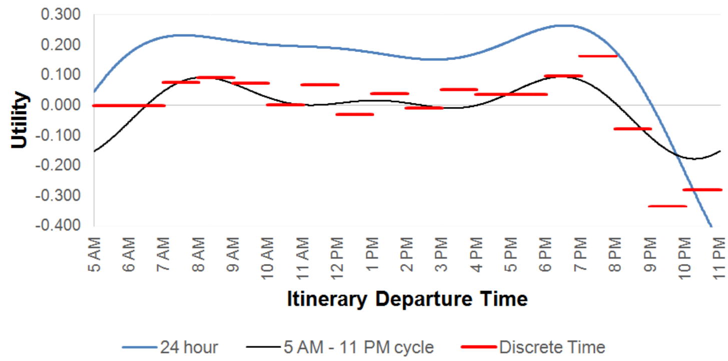


Figure 3.13: Time of day preferences for two time zone westbound model

For the second part of the analysis, we compared two continuous time of day formulations for each segment. The first was based on a 24-hour cycle while the second defined the cycle length to include only those time periods that had 500 or more observations (as shown in Table 3.10). In order to compare log-likelihood values between these two models, we needed to include all itineraries in the same model. Thus, we assigned itineraries corresponding to time periods with fewer than 500 observations to the nearest time period that had more than 500 itineraries. This means we assigned a departure time of 10:59PM for itineraries departing from 11PM to 1:59AM. Model results are show in in Table 3.13.

Segment	LL 24 hour cycle	LL truncated cycle	Truncated cycle length
Same TZ $\leq$ 600	-5,569,465.03	-5,569,210.85	5 AM-midnight
Same TZ $>$ 600	-4,811,564.18	-4,812,128.71	5 AM - 11 PM
One TZ WB $\leq$ 600	-748,951.53	-748,985.81	5 AM - 11 PM
One TZ WB $>$ 600	-4,015,854.88	-4,015,920.68	5 AM - 11 PM
One TZ EB $\leq$ 600	-764,993.36	-764,898.25	5 AM - 11 PM
One TZ EB $>$ 600	-3,933,165.37	-3,933,500.59	5 AM - 11 PM
2 TZ WB	-1,924,792.03	-1,925,085.45	5 AM - 11 PM
2 TZ EB	-1,851,833.78	-1,851,874.18	5 AM - 2 AM
3 TZ WB	-1,492,853.58	-1,492,935.29	5 AM - 11 PM
3 TZ EB	-1,611,863.53	-1,611,807.19	5 AM - 2 AM

Table 3.13: Model fit statistics for 24 hour and truncated cycle for all segments

Key: TZ=time zone, EB=eastbound WB=westbound. LL = log likelihood

Results based on log likelihood values show that the two continuous time of day formulations provide similar model fits, but that the 24-hour cycle fits the data better for seven of the ten market segments.

Bibliography

Abou-Zeid, M., T. Rossi, and B. Gardner (2006). Modeling time-of-day choice in context of tour-and activity-based models. *Transportation Research Record: Journal of the Transportation Research Board* (1981), 42–49.

Algers, S. and M. Beser (2001). Modelling choice of flight and booking class-a study using stated preference and revealed preference data. *International Journal of Services Technology and Management 2*(1-2), 28–45.

Armantier, O. and O. Richard (2008). Domestic airline alliances and consumer welfare. *The RAND Journal of Economics 39*(3), 875–904.

Babbie, E. R. (1998). *The Practice of Social Research*, Volume 112. Wadsworth Publishing Company Belmont, CA.

Benkard, C., A. Bodoh-Creed, J. Lazarev, et al. (2008). The long run effects of US airline mergers. Working Paper, Yale University.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Econometrica: Journal of Political Economy 112*(1), 68–105.

Branstetter, L. G., C. Chatterjee, and M. Higgins (2011). Regulation and welfare: Evidence from paragraph iv generic entry in the pharmaceutical industry. Technical report, National Bureau of Economic Research.

Carrier, E. (2008). *Modeling the Choice of an Airline Itinerary and Fare Product Using Booking and Seat Availability Data*. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Coldren, G. and F. Koppelman (2005a). Modeling the proximate covariance property of air travel itineraries along the time-of-day dimension. *Transportation Research Record: Journal of the Transportation Research Board* (1915), 112–123.

Coldren, G. M. and F. S. Koppelman (2005b). Modeling the competition among air-travel itinerary shares: GEV model development. *Transportation Research Part A: Policy and Practice 39*(4), 345–365.

Coldren, G. M., F. S. Koppelman, K. Kasturirangan, and A. Mukherjee (2003). Modeling aggregate air-travel itinerary shares: Logit model development at a major US airline. *Journal of Air Transport Management 9*(6), 361–369.

Cornia, M., K. S. Gerardi, and A. H. Shapiro (2012). Price dispersion over the business cycle: Evidence from the airline industry. *The Journal of Industrial Economics 60*(3), 347–373.

Daly, A. and M. Bierlaire (2006). A general and operational representation of generalised extreme value models. *Transportation Research Part B: Methodological 40*(4), 285–305.

Garrow, L. A. (2004). *Comparison of Choice Models Representing Correlation and Random Taste Variation: An Application to Airline Passengers' Rescheduling Behavior*. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL.

Garrow, L. A., G. M. Coldren, and F. S. Koppelman (2010). MNL, NL, and OGEV models of itinerary choice. In *Discrete Choice Modelling and Air Travel Demand: Theory and Applications*, Chapter 7, pp. 203–252. Ashgate Publishing, Ltd.

Gayle, P. G. (2008). An empirical analysis of the competitive effects of the delta/continental/northwest code-share alliance. *Journal of Law and Economics 51*(4), 743–766.

Goolsbee, A. and A. Petrin (2004). The consumer gains from direct broadcast satellites and the competition with cable tv. *Econometrica 72*(2), 351–381.

Greene, W. H. (2003). *Econometric Analysis*. Pearson Education India.

Groves, R. M. (2004). *Survey Errors and Survey Costs*, Volume 536. John Wiley & Sons.

Guevara, C. and M. Ben-Akiva (2006). Endogeneity in residential location choice models. *Transportation Research Record: Journal of the Transportation Research Board* (1977), 60–66.

Guevara, C. A. (2015). Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research Part A: Policy and Practice 82*, 240–254.

Guevara, C. A. and M. E. Ben-Akiva (2012). Change of scale and forecasting with the control-function method in logit models. *Transportation Science 46*(3), 425–437.

Guevara, C. A. (2010). *Endogeneity and Sampling of Alternatives in Spatial Choice Models*. Ph.D. Dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Guevara, C. A. and P. Navarro (2013). Control-function correction with weak instruments in logit models. Working Paper, Universidad de los Andes, Chile.

Hausman, J., G. Leonard, and J. D. Zona (1994). Competitive analysis with differenciated products. *Annales d'Economie et de Statistique*, 159–180.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The Economics of New Goods*, pp. 207–248. University of Chicago Press.

Hotle, S. L., M. Castillo, L. A. Garrow, and M. J. Higgins (2015). The impact of advance purchase deadlines on airline consumers' search and purchase behaviors. *Transportation Research Part A: Policy and Practice 82*, 1–16.

Jacobs, T. L., L. A. Garrow, M. Lohatepanont, F. S. Koppelman, G. M. Coldren, and H. Purnomo (2012). Airline planning and schedule development. In *Quantitative Problem Solving Methods in the Airline Industry: A Modeling Methodology Handbook.* Part of the Fred Hillier International Series on Operations Research and Management Sciences. Vol. 169. Eds. C. Barnhart and B. Smith. New York: Springer. pp. 35-100.

Koppelman, F. S., G. M. Coldren, and R. A. Parker (2008). Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B: Methodological 42*(3), 263–273.

Manski, C. F. and S. R. Lerman (1977). The estimation of choice probabilities from choice based samples. *Econometrica: Journal of the Econometric Society*, 1977–1988.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In: Zarembka,P. (Ed.), *Frontiers in Econometrics*. Academic Press, New York, pp. 105–142.

Mumbower, S., L. A. Garrow, and M. J. Higgins (2014). Estimating flight-level price elasticities using online airline data: A first step toward integrating pricing, demand, and revenue optimization. *Transportation Research Part A: Policy and Practice 66*, 196–212.

Nako, S. M. (1992). Frequent flyer programs and business travellers: An empirical investigation. *Logistics and Transportation Review 28*(4), 395.

Newman, J. P. (2016). Larch documentation. Discrete choice modeling software available at `http://larch.readthedocs.org/en/latest/model.html`.

Pekgün, P., P. M. Griffin, and P. Keskinocak (2013). An empirical study for estimating price elasticities in the travel industry. Technical report, Working Paper, University of South Carolina.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research 47*(1), 3–13.

Proussaloglou, K. and F. S. Koppelman (1999). The choice of air carrier, flight, and fare class. *Journal of Air Transport Management 5*(4), 193–201.

Rivers, D. and Q. H. Vuong (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of econometrics 39*(3), 347–366.

Southwest Airlines (2016). Southwest airlines website, search for itineraries between DAL and PHX for travel on april 1, 2016. `https://www.southwest.com/`. [Online; accessed 22-January-2016].

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica: Journal of the Econometric Society 65*, 557–586.

Stern, S. (1996). Market definition and the returns to innovation: Substitution patterns in pharmaceutical markets. Technical report, Working Paper, Sloan School of Management, Massachusetts Institute of Technology.

Suzuki, Y., J. E. Tyworth, and R. A. Novack (2001). Airline market share and customer service quality: a reference-dependent model. *Transportation Research Part A: Policy and Practice 35*(9), 773–788.

Tirachini, A., D. A. Hensher, and J. M. Rose (2013). Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transportation research part A: policy and practice 53*, 36–52.

Train, K. E. and C. Winston (2007). Vehicle choice behavior and the declining market share of us automakers. *International Economic Review 48*(4), 1469–1496.

U.S. Department of Transportation (2016). Bureau of transportation statistics (quarter 3, 2013) origin and destination survey databank 1A/1B. `https://www.transportation.gov/`. [Online; accessed 18-January-2016].

Villas-Boas, J. M. and R. S. Winer (1999). Endogeneity in brand choice models. *Management Science 45*(10), 1324–1338.

Wardman, M. and G. Whelan (2011). Twenty years of rail crowding valuation studies: Evidence and lessons from British experience. *Transport Reviews 31*(3), 379–398.

Chapter 4

Modeling the Competition among Airline Itineraries: A Comparison of 2000 and 2013 data

Lurkin, V., Garrow, L., Higgins, M., Newman, J., and Schyns, M. Modeling the competition among airline itineraries: A comparison of 2000 and 2013 data. *Working paper to be submitted to Transportation Research Part A: Policy and Practice.*

Discrete choice models are commonly used to forecast the probability an airline passenger chooses a specific itinerary. Prior work conducted for United Airlines based on January 2000 data used nested, ordered generalized extreme value, cross-nested, and network generalized extreme value discrete choice models to examine inter-itinerary competition along three dimensions: nonstop versus connecting level of service, carrier, and time of day. This chapter models competition among itineraries for these three dimensions using data for May 2013 departures. Despite the many structural changes that occurred in the airline industry over the past 15 years (including industry consolidation, the increased use of online distribution channels, and product de-bundling trends), our results are strikingly similar to those observed based on 2000 data. In contrast to estimation databases available to airlines, our database contains information about both booking and departure dates. Consequently, our price formulation is the first within the itinerary choice literature to estimate different price sensitivities as a function of advance purchase periods. This is important, as the airline industry can use our results to incorporate different customer segments as revealed through high-yield and low-yield booking curves when evaluating the profitability of airline schedules.

4.1 Introduction and motivation

Understanding demand for products is an integral part of many fields, including aviation. Airlines, airports, local and regional planning organizations, national government agencies, aircraft manufacturers, and other private and public organizations need to understand demand for air travel in order to evaluate and implement different business and policy options. For example, many airports and planning commissions have conducted studies to better understand why airline passengers choose to use a particular airport in multi-airport regions and how potential investments in public transit and other transportation infrastructure will impact demand at these airports (see *ACRP Report 98: Understanding Airline and Passenger Choice in Multi-Airport Regions* (ACRP (2013)) for a review of dozens of studies published from 2005-2011, and Nesset and Helgesen (2014) and Seelhorst and Liu (2015) for examples of more recent studies). Policy makers need to understand passengers' willingness to pay to travel during different times of the day when evaluating demand management strategies, such as airport slot restrictions. Understanding how individuals make choices among itineraries is also a critical part of network planning models, which airlines and aircraft manufacturers use to forecast the profitability of airline schedules.

Network planning models forecast schedule profitability by determining the number of passengers who travel in an origin-destination pair, allocating these passengers to specific itineraries, and calculating expected costs and revenues. The passenger allocation model is often referred to as an itinerary choice model because it represents how individuals make choices among itineraries. Many airlines use discrete choice models to capture how individuals make trade-offs among different itinerary characteristics, e.g., departure times, elapsed times, the number of connections, equipment types, carriers, and prices. These network planning models support many important long- and intermediate-term decisions. For example, they aid airlines in performing merger and acquisition scenarios, route schedule analysis, codeshare scenarios, minimum connection time studies, price-elasticity studies, hub location and hub buildup studies, and equipment purchasing decisions (Garrow et al. (2010a)).

Despite the importance of these itinerary choice models, there are few published studies describing how these models are estimated and used in practice. A notable exception is work done in the early 2000s by Greg Coldren. His dissertation (Coldren (2005)) describes itinerary choice models that were estimated and implemented by United Airlines[1]. Several papers and book chapters resulted from his dissertation work (Coldren et al. (2003), Coldren and Koppelman (2005a,b), Koppelman et al. (2008), Garrow et al. (2010a), Jacobs et al. (2012)).

As part of our prior work described in Chapter 3, we estimated a multinomial logit itinerary choice model for continental U.S. markets that accounted for price endogeneity using multiple instruments. Our estimations were based on a ticketing database of more than 10 million passenger itineraries that departed in May of 2013. Results showed the importance of accounting for price endogeneity in itinerary choice models; failure to do so resulted in value of time estimates that were high, biased price estimates, and incorrect pricing recommendations.

This chapter builds on prior research by extending the analysis to discrete choice models that incorporate inter-alternative competition. Prior work conducted for United Airlines based on 2000 data used nested, ordered generalized extreme value, cross-nested, and network generalized extreme value discrete choice models to examine inter-itinerary competition along three dimensions: nonstop versus connecting level of service, carrier, and time of day. Our work also models competition among itineraries for these three dimensions, but using data for May 2013 departures. *A priori*, we expected to see differences between our models and those estimated by Coldren as there have been numerous structural changes that occurred in the U.S. airline industry over the past 15 years including industry consolidation, increased tickets sales on the internet, de-bundled pricing trends, and other factors.

For example, from 2005-2013 ten major U.S. carriers went through mergers/acquisitions (America West and US Airways in 2005; Delta and Northwest in 2008; Continental and United in 2010; Southwest and AirTran in 2011; American and US Airways in 2013). From 1999 to 2002, the U.S. General Accounting Office (GAO) estimated that the percentage of airline tickets booked online (through direct airline-owned websites and online travel agencies) increased from 7 to 30 percent (GAO (2003)). In 2013, more than 35 percent of bookings were made through airline websites, and this percentage

[1]United Airlines implemented Coldren's multinomial logit models, described in Coldren et al. (2003), and saw significant improvements in forecasting accuracy. The mean absolute percent deviation in the predicted number of onboard passengers measured at the flight level decreased from 17.91 to 16.23.

is expected to grow to 59% by 2017[2] (Atmosphere Research Group (2012)). Finally, after the economic crisis of 2008, U.S. airlines increased their ancillary sources of revenue by charging for products and services previously included in the price of an airline ticket (such as imposing fees for the first checked bag) and increasing existing fees (such as ticket exchange fees). IdeaWorksCompany (2015) estimates that worldwide, these and other sources of ancillary revenue grew from \$22.6B (representing 4.8 percent of global airline revenues) from 2010 to \$59.2B (or 7.8 percent) in 2015.

Our work contributes to the literature in two key ways. First, we estimate nested logit and ordered generalized extreme value itinerary choice models that incorporates inter-alternative competition along three dimensions: nonstop versus connecting level of service, carrier, and departure time of day. Despite the significant changes in the U.S. airline industry that have occurred over the past 15 years, our model results are strikingly similar to those obtained by Coldren. This is important, as it suggests that customer preferences[3] have been stable over time and that customer behavior is similar across distribution channels. Second, we estimate multiple multinomial logit models that differ in how they represent price in the utility function. That is, we take advantage of the fact that in contrast to estimation databases available to airlines, our database contains information about both booking and departure dates. Consequently, our price formulation is the first within the itinerary choice literature to estimate different price sensitivities as a function of advance purchase periods. This is important, as the airline industry can use our results to incorporate different customer segments as revealed through high-yield and low-yield booking curves when evaluating the profitability of airline schedules. Our study provides empirical evidence on the value of including booking date information on publicly-available ticketing data sources, most notably the Bureau of Transportation Statistic's Airline Origin and Destination Survey, commonly referred to as the DB1B database (BTS (2016)).

The remaining sections are organized as follows. Section 4.2 provides an overview of the key data and modeling differences used by Coldren and in our

[2]This excludes additional bookings from online travel agencies, so the overall percentage of bookings made online will be even higher.

[3]This applies for the demand models supporting intermediate and long-term decisions discussed in the paper. Customer preferences modeled as part of more short-term and tactical models for revenue management, online conversion, and similar applications may have changed over the past 15 years.

analysis. Section 4.3 provides an overview of discrete choice models we used to incorporate inter-alternative correlation. Empirical results are presented in Section 4.4. We conclude in Section 4.5 by highlighting how our model contributes to the literature.

4.2 Review of industry itinerary choice models

As noted in the introduction, there have been very few studies published in the literature that describe itinerary choice models used by industry. The dissertation of Coldren (Coldren (2005)) provides one of the most comprehensive treatments and includes estimation results based on data from January 2000. We used Coldren's study to inform the models we estimated to account for inter-alternative correlations as part of our study. However, there are several important differences, shown in Table 4.2, which could potentially influence our estimation results.

Both studies used data on booking or ticketing sales. Coldren used "CRS data [that] are commercially available and compiled from several computer reservation systems including Apollo, Sabre, Galileo, and WorldSpan as well as Internet travel sites such as Orbitz, Travelocity, Expedia, and Priceline" (Coldren et al. (2003)). Although Coldren's data did not include direct sales (e.g., through Southwest.com), it did capture approximately 90 percent of domestic U.S. air-travel bookings during the study year period. Similar to Coldren, we used a database that included sales information made through brick-and-mortar and online travel agencies. Our database, provided by the Airlines Reporting Corporation (ARC), does include direct sales for some carriers, but like Coldren's data is underrepresented in the number of transactions from low cost carriers. ARC estimates that its ticketing database captures approximately 40-50 percent of domestic U.S. travel (Howard (2016)).

There are two key differences between the sales data used by Coldren and in our study, First, Coldren's data is based on bookings whereas our data is based on tickets. The difference between booking and ticketing data relates to whether the passenger has paid for a reservation. A reservation, or booking request, that has been paid for appears in both booking and ticketing databases, whereas a booking request that has not yet been paid for appears only in a booking database. Second, Coldren's data are more aggregate than our data. CRS data do not provide information about individual tickets, such as when the ticket was purchased or the amount paid. CRS data mask this information to satisfy anti-trust concerns. In contrast, we are not lim-

	Coldren's Model	Our Model
Data	January 2000 CRS booking data	May 2013 ARC ticketing data
Bookings	Approximately 90% (few Southwest bookings and no direct bookings)	Approximately 40-50% (few bookings by low cost carriers; includes direct bookings for some carriers and both online and brick-and-mortar travel agency bookings)
Segments	Itineraries that depart in eastern time zone and land in western time zone in U.S.	All continental U.S. Markets
# Reporting Carriers*	85	40
Choice set generation	Alternatives created from schedule file and itinerary generation rules	Alternatives created from set of purchased alternatives
Code share alternatives	For a two leg itinerary, given a carriers that can market leg 1, b carriers for leg 2 and c carriers for leg 3, $a \times b \times c$ itineraries are generated, one for each possible combination	Single alternative generated with an indicator variable it can be sold as a code share
Price	Ratio of average carrier fare divided by industry average fare for an airport pair	Average high yield and average low yield fare by operating carrier and level of service for an airport pair

Table 4.1: Comparison of data and modeling assumptions in our work and Coldren's work

*Source: BTS Origin and Destination Survey (DB1B). The number of reporting carriers in Q1 2000 and the number of reporting carriers in Q2 2013 are reported in the tables (BTS 2016).

ited by these regulations and worked with ARC to obtain detailed ticketing information. This allows us to incorporate more detailed price specifications in our utility function.

Although Coldren had data for all continental U.S. markets, he only reported models results using his entire analysis database for multinomial logit models. For models that incorporated inter-alternative competition, he used data for directional itineraries that departed in the eastern time zone and arrived in the western time zone. For our study, we estimate our models using one month of ticketing data for all continental U.S. markets[4].

There are also subtle differences in how the itineraries were generated for the choice sets. We generated our alternatives using information on purchased itineraries (see Appendix A1 for a detailed description of our choice set generation process). In contrast, Coldren generated alternatives from a schedule file and used a proprietary itinerary generation algorithm to combine legs into different itineraries. We also modeled code shares differently than Coldren. A codeshare is an itinerary that can be sold by more than one carrier. In our choice set generation process, we created a single itinerary representing the carrier(s) that operated each flight leg. If any of the legs could be sold by a different airline, we set a code share indicator variable to 1. In contrast, Coldren generated multiple itineraries representing each possible marketing relationship. For a three leg itinerary, given a carriers that can market leg 1, b carriers for leg 2, and c carriers for leg 3, $a \times b \times c$ itineraries are generated in Coldren's process, one for each possible combination. We prefer our approach, as the number of possible code share itineraries can become quite large using Coldren's approach (particularly today as the number of code share partners on each leg has expanded) and many of these itineraries will have little to no demand.

The studies also model fare differently. Coldren uses a ratio of the average carrier fare divided by the industry average fare for an airport pair; no distinction is made for high yield and low yield products. We also represent fare using an average (as carriers do not have access to the same individual-level ticket prices we do). However, our average fare distinguishes between high yield and low yield products and is averaged for each operating carrier and level of service (nonstop, direct, single connection and double connection) for an airport pair. For the advanced models reported in this chapter,

[4]All estimations were performed in Larch (Newman (2016)) on a Dell PowerEdge R710 Dual Intel Xeon E5520 with 2.4GHz CPUs and 288GB of memory

we do not account for price endogeneity; however, extensions to account for endogeneity are currently in progress.

As a final note, the different time periods used for the analysis represent different market conditions. The U.S. airline industry was less concentrated in Q1 of 2000 versus Q2 of 2013 as seen by the fact that 85 carriers reported ticketing data to the Bureau of Transportation Statistics (BTS) in Q1 2000 versus only 40 in Q2 2013 (BTS (2016)).

A priori, we assumed that these differences could explain differences observed in results between the two studies; however, as discussed in Section 4.4, the general patterns observed in inter-alternative correlation between the two studies are remarkably similar. With a few exceptions, the parameter estimates that capture these inter-alternative correlations are also similar.

4.3 Generalized extreme value (GEV) models

This section describes the multinomial logit (MNL), two-level nested logit (NL), three-level nested logit, OGEV models, cross-nested logit (CNL), and network GEV (NetGEV) models. All of these discrete choice models have been previously used to estimate itinerary choice models; as part of our analysis we estimate NL and OGEV models and, based on a comparison with prior models reported in the literature, qualitatively assess the need to estimate more complex CNL and NetGEV models. Tables 4.2 and 4.3, provided at the end of this section, summarize the probability functions and direct- and cross-elasticities for multiple models presented in this section. This section draws heavily on Garrow (2010b).

4.3.1 Multinomial logit models

We model the itinerary choice y_i that individual i makes among the set of directional itineraries in an origin-destination pair that depart on day of week d as a function of itinerary, carrier, and product characteristics. We exclude socioeconomic information as we have no information about the individual who purchased the ticket.

For cases where y_i represents discrete outcome, as in the current situation, it is natural to model the *probability* that y_i takes on a given value, using a discrete choice model. The utility U for individual n in choosing alternative i from choice set $\mathbb{J}_n$ is a linear-in-parameters function of $\mathbf{x}_{ni}, U_{ni} = V_{ni} + \varepsilon_{ni} = \boldsymbol{\beta}_i'\mathbf{x}_{ni} + \varepsilon_{ni}$, where U_{ni} is the true utility, V_{ni} is the estimated utility, $\mathbf{x}_{ni}$ comprises the itinerary, carrier and product variables described above and β is a vector of estimated coefficients. If ε_{ni} is distributed independently and identically (iid) with a Gumbel (or extreme value type I) distribution, the discrete choice model becomes a multinomial logit (MNL) model (McFadden (1974)) and the probability of individual n choosing alternative i is given as:

$$P(y_n = i|\mathbf{x}_{ni}) = \frac{e^{V_i}}{\sum_{j \in \mathbb{J}_n} e^{V_j}} = \frac{e^{\boldsymbol{\beta}_i'\mathbf{x}_{ni}}}{\sum_{j \in \mathbb{J}_n} e^{\boldsymbol{\beta}_j'\mathbf{x}_{nj}}} \tag{4.1}$$

In the MNL model, the assumption that the error terms are iid Gumbel is advantageous in the sense that the choice probability takes on a closed-form expression that is computationally simple. However, the same assumption leads to the independence of irrelevant alternatives (IIA), a property which states that the ratio of choice probabilities P_{ni}/P_{nj} for $i, j \in \mathbb{J}_n$ is independent of the attributes of any other alternative. In terms of substitution patterns, this means a change or improvement in the utility of one alternative will draw share proportionately from all other alternatives. This can be seen mathematically in the direct- and cross-elasticities shown in Table 4.3, which measure the change in market share for alternative i and alternative j, respectively, due to making a change in the k^{th} variable associated with alternative i, x_{ik}.

4.3.2 Two-Level nested logit models

In many applications, the IIA property may not be realistic. For example, in itinerary choice model applications, we expect itineraries operated by the same carrier to compete more with each other than itineraries operated by other carriers or nonstop itineraries to compete more with other nonstop itineraries than connecting itineraries. Other models that belong to the GEV class relax the independence assumption by including covariance terms that are created through allocating an alternative to one or more nests while maintaining the assumption that total variance is identically distributed across alternatives. From a theoretical perspective, these two requirements (more general covariance structure combined with equality of total variance) relax the IIA property while maintaining closed-form expressions for probabilities.

For example, the nested logit (NL) model (Williams (1977), McFadden et al. (1978)), relaxes the assumption that errors are independently distributed by grouping alternatives into M nests, i.e., $, m = 1, 2, ..., M$. An alternative belongs to one and only one nest. The NL utility function can be expressed as follows (suppressing the index for individual n for notational convenience):

$$U_{im} = V_i + \varepsilon_m + \varepsilon_i \tag{4.2}$$

That is, the total variance associated with each alternative in nest m is decomposed into a common error component, ε_m , and an independent error term, ε_i. Alternatives that belong to the same nest share a common error term. The first assumption states that the total variance associated with each alternative, given as $(\varepsilon_i + \varepsilon_m)$ must be identically distributed and follow a Gumbel with mode zero and scale 1. The second assumption states that the independent error terms, ε_i, also follow a Gumbel with mode zero, but with a different scale. Formally, ε_i are distributed such that they have a cumulative distribution function of

$$exp\left(-\sum_{m=1}^{M}\left(\sum_{i \in A_m} exp^{-\varepsilon_i/\mu_m}\right)^{\mu_m}\right), 0 < \mu_m \leq 1 \tag{4.3}$$

Under the assumptions that $(\varepsilon_i + \varepsilon_m)$ is identically distributed $G(0, 1)$ and that ε_i is distributed $G(0, \frac{1}{\mu_m}), i \in A_m, m = 1, 2, ..., M$, the probability that individual n selects alternative i is given as follows (suppressing the index of n for notational convenience):

$$P_i = \frac{e^{V_i/\mu_m}\left(\sum\limits_{j \in A_m} e^{\frac{V_j}{\mu_m}}\right)^{\mu_m - 1}}{\sum\limits_{l=1}^{M}\left(\sum\limits_{j \in A_l} e^{\frac{V_j}{\mu_l}}\right)^{\mu_l}}, 0 < \mu_m \leq 1 \tag{4.4}$$

A more intuitive expression for the NL choice probability can be derived as the product of a conditional and marginal probability (this derivation is provided in Train (2009), p. 90). This formulation is particularly helpful when extending NL models to include additional levels of nests.

$$P_i = P_{i|m} \times P_m = \frac{e^{\frac{V_i}{\mu_m}}}{\sum\limits_{j \in A_m} e^{\frac{V_j}{\mu_m}}} \times \frac{e^{V_m + \mu_m \Gamma_m}}{\sum\limits_{l=1}^{M} e^{V_l + \mu_l \Gamma_l}}, \tag{4.5}$$

$$\Gamma_m = \ln\left(\sum_{j \in A_m} e^{\frac{V_j}{\mu_m}}\right), 0 < \mu_m \leq 1$$

The first component of the product is the probability of selecting alternative i among all j alternatives in nest m, conditional on the choice of m, and the second product is the probability of selecting nest m among all nests. Γ_m is often called the "log-sum term" because it is the log of a sum (this terminology should not to be confused with μ_m , the "logsum" or "logsum parameter"). The logsum parameter, μ_m, is a measure of the degree of correlation and substitution among alternatives in nest m. More precisely, the correlation is given by $1 - \mu_m^2$. Higher values of μ_m imply less, and lower values imply more, correlation among alternatives in the nest. In turn, higher correlation leads to greater competition effects among alternatives in the nest.

This increased substitution among alternatives that share a common nest can be seen in the direct- and cross-elasticity equations, given in Table 4.3. When $\mu_m = 1$, the NL direct- and cross-elasticities are identical to those of the MNL model. Second, even when $0 < \mu_m < 1$, the MNL and NL cross-elasticities are identical for alternatives that are not in the same nest. Intuitively, this is because alternatives that are not in the same nest do not share a common error term and are independent.

Figure 4.1 shows a two-level NL model that groups alternatives into nests that correspond to level of service, i.e. the first nests contains two nonstop itineraries and the second nest contains two connecting itineraries. In this example, the two-level NL model can be used to incorporate increased substitution for alternatives that share a common nest - but only for one product dimension (e.g., either carrier or level of service, but not both carrier and level of service).

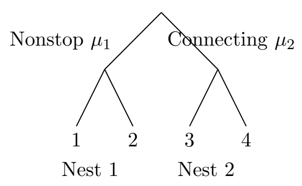


Figure 4.1: Two-level NL model nested by level of service

4.3.3 Three-Level nested logit models

To incorporate increased substitution across multiple dimensions, the analyst has several options. The first option is to use a NL model, such as the one shown in Figure 4.2, that contains three (and potentially more) levels.

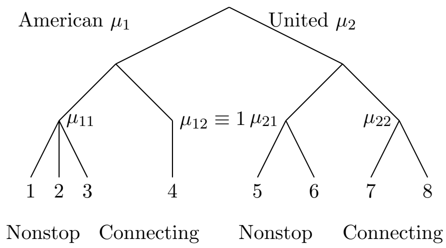


Figure 4.2: Three-level NL model

In this example, the level 3 nests represent carriers nests for American (μ_1) and United (μ_2) and the level 2 nests represent nonstop (μ_{11} and μ_{21}) and connecting (μ_{12} and μ_{22}) level of service nests. Note that in this discussion, n refers to a second level nest, not an individual as used in previous discussions. The probability of choosing alternative i in a three-level nest is given as:

$$P_i = P_{i|nm} \times P_{nm|m} \times P_m = \frac{e^{\frac{V_i}{\mu_{nm}}}}{\sum\limits_{j \in B_n} e^{\frac{V_j}{\mu_{nm}}}} \times \frac{e^{\mu_m \times \mu_{nm} \times \Gamma_{nm}}}{\sum\limits_{k \in A_m} e^{\mu_m \times \mu_{km} \Gamma_{km}}} \times \frac{e^{\mu_m \Gamma_m}}{\sum\limits_{l=1}^{M} e^{\mu_l \Gamma_l}}, \quad (4.6)$$

$$\Gamma_{nm} = \ln\left(\sum_{j \in B_n} e^{\frac{V_j}{\mu_{nm}}}\right), \quad \Gamma_m = \ln\left(\sum_{k \in A_m} e^{\mu_m \times \mu_{km} \Gamma_{km}}\right), \quad 0 \leq \mu_m, \mu_{nm} \leq 1$$

The first component of the product is the probability of selecting alternative i among all j alternatives in nest nm, conditional on the choice of nm. The second product is the probability of selecting nest nm among all two-level nests in nest m, conditional on the choice of m. The third product is the probability of selecting nest m among all three-level nests.

To ensure that the three-level NL model is consistent with utility maximization, the correlation must increase as one moves down the tree. In Figure

4.2, this implies that the correlation (or substitution among alternatives) for alternative 5 is the strongest (or greatest) with alternative 6, and stronger with alternatives 7 and 8 than alternatives 1 to 4. Formally, $\mu_{11} < \mu_1$ and $\mu_{22} < \mu_2$. Note that for nests that contain a single alternative, the corresponding logsum coefficients are set to 1 for identification purposes and do not affect model interpretations. Also, note that the Figure 4.2 shows logsum parameters estimated for each carrier, but in Coldren's (and our) application, the logsums are constrained to be the same for each nest[5].

In practice, it is common to change the ordering of dimensions associated with each level of the nest and empirically determine which nesting structure is theoretically valid and best fits the data. For example, a model that has carrier as the upper nest and level of service as the lower nest with theoretically valid logsum coefficients means that itineraries that share the same carrier and level of service compete the most, and itineraries that share a common carrier compete more than itineraries that do not share the same operating carrier.

4.3.4 Ordered generalized extreme value models with T=1

To incorporate increased substitution across multiple dimensions, the analyst can also use discrete choice modes that allocate alternatives to more than one nest. The ordered generalized extreme value (OGEV) model (Small (1987)) is used in applications in which the ordering of alternatives has a physical meaning. For example, the OGEV model can be used to capture time of day competition effects among airline itineraries. Figure 4.3 shows an OGEV model for six alternatives ($J = 6$) and one adjacent time period ($T = 1$)[6]. In contrast to notation used thus far, note that nest one is not defined as the first nest on the left, but as the first nest that contains more than one alternative. The total number of nests is given as $(J - T + 2T)$ where $(J - T)$ is the number of nests that contain more than one alternative and $2T$ is the number of nests that contain one alternative.

The OGEV probability is given as:

[5] We did this for LOs, TOD and carrier nests (all)

[6] Consistent with Coldren's application, the OGEV models shown in Figures 4.3 and 4.4 constrain the logsum coefficients to be the same across all nests.

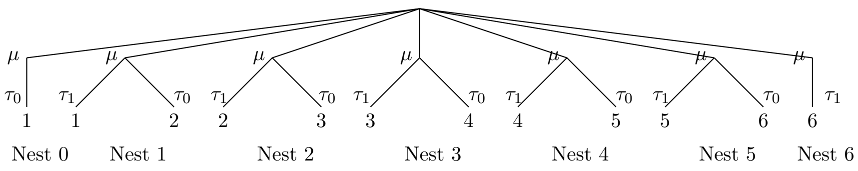


Figure 4.3: Ordered GEV model with one adjacent time period

$$P_i = \sum_{m=i}^{i+T} P_{i|m} \times P_m = \sum_{m=i}^{i+T} \left[\frac{\left(\tau_{m-i} e^{V_i}\right)^{\frac{1}{\mu}}}{\sum\limits_{j \in A_m} \left(\tau_{m-j} e^{V_j}\right)^{\frac{1}{\mu}}} \times \frac{\left(\sum\limits_{j \in A_m} \tau_{m-j} e^{\frac{V_j}{\mu}}\right)^{\mu}}{\sum\limits_{r=1}^{J+T} \left(\sum\limits_{s \in A_r} \tau_{r-s} e^{\frac{V_s}{\mu}}\right)^{\mu}} \right],$$

$$0 < \mu \leq 1, \sum_{m=1}^{J+T} \tau_m = 1 \tag{4.7}$$

where

- T is the number of adjacent time periods in the OGEV model,
- J is the total number of alternatives,
- $j \in A_m$ is the set of all alternatives that belong to nest m,
- r is an index used to sum over all nests,
- τ_{m-1} are unknown allocation parameters that characterize the portion of alternative i assigned to a nest. Allocation parameters are non-negative, i.e., $\tau_{m-1} \geq 0$ and must sum to one for every alternative. Defining nests for a T-step OGEV model as shown in Figure 3, alternative i belongs to nests $i-1, i, , i+1, , i+T$, this last condition is equivalent to $\sum_{m=1}^{J+T} \tau_m = 1$
- μ is the logsum coefficient associated with each nest.

The first component of the product is the probability of selecting alternative i among all alternatives that belong to nest m, conditional on choosing nest m. The second product is the probability of selecting nest m among all nests. The total probability for alternative i is obtained by summing over all nests that contain alternative i. The portion of an alternative that shares a nest with itineraries that depart in the time period immediately before (τ_0) or the time period immediate after (τ_1) is estimated from the data. A constraint is also added to ensure that $\tau_0 + \tau_1 = 1$.

From an interpretation perspective, a value of $0.5 < \tau_0 < 1$ (and $0 < \tau_1 < 0.5$) means that an itinerary departing in time period three would compete more with itineraries in the earlier time period two than with itineraries in the later time period four. Intuitively, this result would be expected for outbound itineraries for travelers that need to arrive at their destinations by a fixed time. The increased substitution among alternatives that depart in the same time period or adjacent time periods is also seen in the direct-elasticity and cross-elasticity equations reported in Table 4.3. In practical applications, it is common to constrain the allocation parameter to 0.5. From a behavioral interpretation, an allocation parameter of 0.5 means that an itinerary departing in time period t competes more (and equally) with itineraries in the earlier time period $t-1$ and later time period $t+1$.

Note that the OGEV cross-elasticity collapses to the MNL cross-elasticity equation for those alternatives that are separated by more than T time periods. The MNL proportional substitution property applies to those alternatives that do not share a nest in common and for which their covariance term is zero. This may be problematic if there is increased substitution among alternatives separated by more than one time step. This can be accommodated using an OGEV model that allocates alternatives to more than one adjacent time period.

4.3.5 Ordered generalized extreme value models with T=2

The OGEV model described above can be extended to more than one adjacent time period as shown in Figure 4.4. The OGEV probability, direct-elasticity, and cross-elasticity formulas discussed above are general and apply to OGEV models with more than one adjacent time period. The OGEV model shown in Figure 4.4 exhibits greater-than-MNL competition for itineraries that depart in the two time periods immediately before or immediately after. Further, itineraries one adjacent period away compete more than itineraries two adjacent periods away.

In practice, the number of time-periods (or T) is determined by comparing model fits between a model that incorporates t adjacent time periods with a model that incorporates $t+1$ adjacent time periods, e.g., the analyst

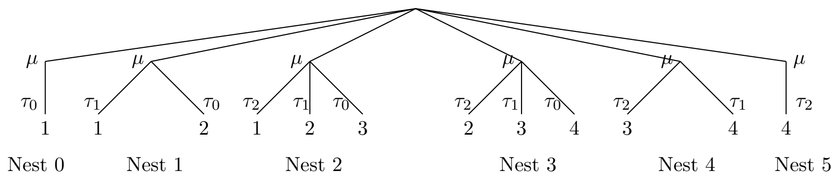


Figure 4.4: Ordered GEV model with two adjacent time periods

compares the model fits between a OGEV model with $t = 1$, $t = 2$, $t = 3$, etc. Sufficient time periods have been incorporated when there is "little" improvement observed in the fit between two models. In practice, it is also common to constrain the allocation parameters so that alternatives are assigned equally to each nest. For the OGEV structure shown in Figure 4.4, this means that $\tau_0 = \tau_1 = \tau_2 = 1/3$.

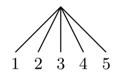
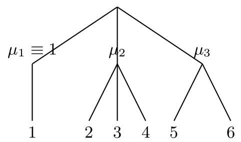
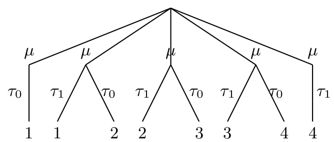
4.3.6 Cross nested logit and network GEV models

The OGEV models described above are a special case of the cross-nested logit (CNL) model, which has also been called a generalized nested logit (GNL) model in the literature (Vovsha (1997), Wen and Koppelman (2001)). Unlike the OGEV model, the CNL is more "general" in the sense that its nesting structures are not restrictive and both allocation and logsum parameters can be estimated (although in practice restrictions on these parameters are often imposed to facilitate estimation and model interpretation). Like the OGEV models, the CNL is based on a two-level nesting structure. Fosgerau et al. (2013) have shown that any choice model can be approximated as precisely as desired by a CNL model. The probability equation for the CNL model is shown in Table 4.2 and the direct- and cross-elasticities in Table 4.3. An example of a CNL model that incorporates increased substitution among products that share the same carrier and/or the same departure time period is shown in Table 4.2[7]. Correlation among alternatives is now a function of both the τ allocation weights and the μ logsum coefficients.

Finally, the network generalized extreme value (NetGEV) model (Daly and Bierlaire (2006)) is more general than the CNL model in that it allows

[7]The CNL model shown in Table 4.2 is equivalent to the two-level "weighted nested logit" models estimated by Coldren; in this case the "weight" of τ_0 applies to the carrier nest and the weight of $\tau_1 = 1 - \tau_0$ applies to the time of day nest.

alternatives to be allocated to more than one nest and incorporates nesting structures with three (or more) levels. The NetGEV model is particularly helpful for itinerary choice modeling applications in which competition exists along multiple dimensions. For example, the NetGEV model can incorporate upper level OGEV nests for departure time of day and lower level carrier nests. Coldren estimated this model, which he called an OGEV-NL model, along with other NetGEV models. The probability equation for the OGEV-NL model and the direct- and cross-elasticities are summarized in Tables 4.2 and 4.3, respectively. See Abbe et al. (2007), Garrow et al. (2010c)) and Newman (2008a,b)) for further details related to NetGEV models.

Model	Probability, P_i (expressed as $P_{i\|m} \times P_m$ where applicable)	Key properties
Multinomial Logit (MNL) McFadden (1974) 	$\dfrac{e^{V_i}}{\sum_J e^{V_j}}$	- Correlation = 0 - IIA - No nests
Nested Logit (NL) McFadden (1974), Williams (1977) 	$\dfrac{e^{\frac{V_i}{\mu_m}}}{\sum_{j \in A_m} e^{\frac{V_j}{\mu_m}}} \times \dfrac{e^{V_m + \mu_m \Gamma_m}}{\sum_{l=1}^{M} e^{V_l + \mu_l \Gamma_l}},$ $\Gamma_m = \ln \left(\sum_{j \in A_m} e^{\frac{V_j}{\mu_m}} \right), 0 < \mu_m \leq 1$	- Correlation $=1 - \mu_m^2$ for i, j in nest m, 0 otherwise - Alternatives allocated to exactly one nest (may be a degenerate nest with one alternative)
Ordered Generalized Extreme Value (OGEV) Small (1987) 	$\sum_{m=i}^{i+T} \left[\dfrac{\left(\tau_{m-i} e^{V_i}\right)^{\frac{1}{\mu}}}{\sum_{j \in A_m} \left(\tau_{m-j} e^{V_j}\right)^{\frac{1}{\mu}}} \times \dfrac{\left(\sum_{j \in A_m} \tau_{m-j} e^{\frac{V_j}{\mu}} \right)^{\mu}}{\sum_{r=1}^{J+T} \left(\sum_{s \in A_r} \tau_{r-s} e^{\frac{V_s}{\mu}} \right)^{\mu}} \right],$ $0 < \mu \leq 1, \sum_{m=1}^{J+T} \tau_m = 1$	- Each alternative shares nest with T adjacent time periods - $\tau_0 < \tau_1$ implies flight in time period 2 competes more with flight in time period 1 than 3

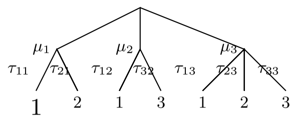
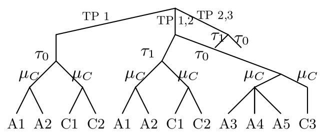
Cross Nested Logit Vovsha (1997), Wen and Koppelman (2001)  	$\sum_m \left[\frac{\left(\tau_{im} e^{V_i}\right)^{\frac{1}{\mu_m}}}{\sum_{j\in A_m} \left(\tau_{jm} e^{V_j}\right)^{\frac{1}{\mu_m}}} \times \frac{\left(\sum_{j\in A_m} \tau_{jm} e^{\frac{V_j}{\mu_m}}\right)^{\mu_m}}{\sum_l \left(\sum_{j\in A_l} \tau_{jl} e^{\frac{V_s}{\mu_l}}\right)^{\mu_l}} \right],$ $0 < \mu_m \leq 1, \sum_m \tau_{jm} = 1$	- Most general two-level GEV structure - All models shown above are special cases of CNL - Allocation and logsum parameters estimated - CNL is a special case of the Network GEV
Ordered Generalized Extreme Value-Nested Logit (OGEV-NL) Coldren and Koppelman (2005b)  	$\sum_{m=1}^{i+T} \left[\frac{e^{\mu_M \Gamma_m}}{\sum_{r=1}^{J+T} e^{\mu_M \Gamma_r}} \times \frac{e^{\frac{\mu_C}{\mu_M}\Gamma_{mc}}}{\sum_{d\in A_m} e^{\frac{\mu_C}{\mu_M}\Gamma_{md}}} \times \frac{\tau_{m-i} e^{\frac{V_i}{\mu_C}}}{\sum_{j\in A_{m,c}} \tau_{m-j} e^{\frac{V_j}{\mu_C}}} \right],$ $0 < \mu_C \leq \mu_M \leq 1, \sum_{m=1}^{T} \tau_m = 1,$ $\Gamma_{mc} = \ln\left(\sum_{j\in A_{m,c}} \tau_{m-j} e^{\frac{V_j}{\mu_C}}\right), \Gamma_m = \ln\left(\sum_{c\in A_m} e^{\frac{\mu_C}{\mu_M}\Gamma_{mc}}\right)$	- Alternatives first grouped by time period in OGEV upper-level, then by one product dimension (e.g. carrier) in lower level. - Alternatives more than T time periods apart exhibit IIA

Table 4.2: Summary of probabilities and key properties for GEV models

Model	Direct Elasticity, $\eta^{P_i}_{X_{ik}}$	Cross Elasticity, $\eta^{P_j}_{X_{ik}}$
Multinomial Logit (MNL)	$(1-P_i)\beta_k X_{ik}$	$-P_i\beta_k X_{ik}$
Nested Logit (NL)	$\left[(1-P_i)+\left(\frac{1-\mu_m}{\mu_m}\right)(1-P_{i\mid m})\right]\beta_k X_{ik}$	$-\left[P_i+\left(\frac{1-\mu_m}{\mu_m}\right)P_{i\mid m}\right]\beta_k X_{ik}$
Ordered Generalized Extreme Value (OGEV)	$\left[(1-P_i)+\sum_{m=i}^{i+T}\left(\frac{1-\mu}{\mu}\right)(1-P_{i\mid m})\right]\beta_k X_{ik}$	$-\left[P_i+\sum_{m=i}^{i+T}\left(\frac{1-\mu}{\mu}\right)\frac{P_{i\mid m}P_{j\mid m}P_m}{P_j}\right]\beta_k X_{ik}$
Cross Nested Logit (CNL)	$\left[(1-P_i)+\sum_{m}\left(\frac{1-\mu_m}{\mu_m}\right)\frac{P_{i\mid m}P_m(1-P_{i\mid m})}{P_i}\right]\beta_k X_{ik}$	$-\left[P_i+\sum_{m}\left(\frac{1-\mu_m}{\mu_m}\right)\frac{P_{i\mid m}P_{j\mid m}P_m}{P_j}\right]\beta_k X_{ik}$
Ordered Generalized Extreme Value Nested Logit (OGEV-NL)	$\left[(\frac{1}{\mu_C}-P_i)+\frac{1}{P_i}\sum_{m=i}^{i+T}\left((1-\frac{1}{\mu_M})P_mP^2_{c\mid m}P^2_{i\mid c,m} + (\frac{1}{\mu_M}-\frac{1}{\mu_C})P_mP_{c\mid m}P^2_{i\mid m,c}\right)\right]\beta_k X_{ik}$	$\left[-P_i+\frac{1}{P_j}\sum_{m=j}^{j+T}\left((1-\frac{1}{\mu_M})P_mP^2_{c\mid m}P_{i\mid m,c}P_{j\mid m,c} + (\frac{1}{\mu_M}+\frac{1}{\mu_C})P_mP_{c\mid m}P_{i\mid m,c}Pj\mid c,m\right)\right]\beta_k X_{ik}$

Table 4.3: Summary of elasticities and cross-Elasticities for GEV models

4.4 Model results

Our modeling approach builds off of Chapter 3 in which we estimated a MNL model for continental U.S. markets based on itineraries that departed in May of 2013. Our utility specification included carrier indicator variables (representing the operating carrier corresponding to the longest flight leg in the itinerary), elapsed time, number of connections, a direct flight indicator, equipment indicators representing whether any leg of the itinerary was operated by a regional jet or propeller aircraft (versus a narrow- or wide-body aircraft), and marketing indicators representing whether the itinerary could be sold only by a single carrier (online), by multiple carriers that had established a code-share relationship (code shares), or by multiple carriers that had not established a code share relationship (interline). Price and departure time of day preferences were also included.

With respect to price, we included two prices: one representing an average high-yield fare for first, business, and unrestricted coach products and the second representing an average low-yield fare for restricted coach and other/unknown products; these averages were calculated for each itinerary origin, destination, carrier, and level of service (nonstop/direct, single connection, and double connection). We corrected for price endogeneity using two instruments. For the MNL model reported in Chapter 3, the first is a Hausman-type instrument that is calculated for itinerary i as the cube of the average price of similar itineraries in other similar markets. Itineraries are considered to be similar if they have the same carrier and level of service (classified as a nonstop/direct versus one or more connections). We assume that markets are similar if they have the same level of competition (i.e., presence of a low-cost carrier or not). For our Stern-type competition instrument, we use a measure of capacity, i.e., the square of monthly seats flown in market by carrier and product type (i.e., first/business seats for high-yield products and coach seats for low-yield products). We use similar instruments for multinomial models developed in this chapter, but vary the markets included in the Hausman-type instrument and/or the functional form used for capacity to ensure instruments remain valid for a given discrete choice model and utility specification.

Departure time of day preferences were included using a continuous specification that combines sine and cosine functions. We estimated models that included six departure time of day (TOD) parameters and then expanded the analysis to include 1260 terms. More precisely, we created ten segments based on the length of haul, direction of travel and number of time zones crossed. For each segment, we estimate separate time of day preferences for departure day of week and itinerary type (defined as outbound, inbound and one-way itineraries[8]).

Figure 4.5 shows an overview of our modeling approach. First, we estimated MNL models that incorporated different functional forms for price and accounted for different price sensitivities by advance purchase periods (which is shown as days from departure (DFD) on the Figure). Given the estimation of the full TOD formulation with 1260 parameters is computationally intensive, we also tested the sensitivity of non-price parameter estimates when six versus 1260 TOD parameters were included in the model. Based on the results of these initial estimations that showed: (1) price was preferred to ln(price); and, (2) parameter estimates were similar for models that did/did not include full TOD formulations, we estimated NL and OGEV models that used a reduced TOD specification (with six parameters) and average high-yield and low-yield prices for different advance purchase periods; for these models we did not instrument price. The results of these NL and OGEV models were compared to those estimated by Coldren and Koppelman (2005b). Based on the fact our NL and OGEV results are strikingly similar to those obtained by Coldren, we review Coldren's CNL and NetGEV model estimations to qualitatively assess the benefit of using these more advanced models, and recommend a preferred model structure for future estimations.

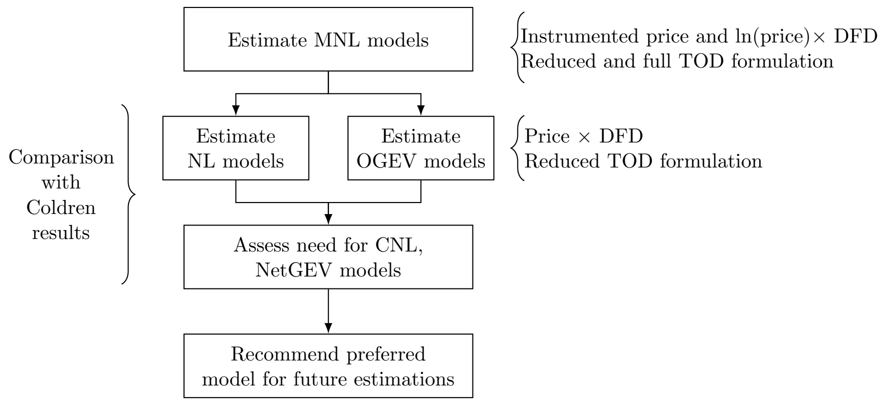


Figure 4.5: Overview of modeling approach

[8]Note that 6 sine and cosine functions × 10 segments × 7 days of week × 3 itinerary types=1260 variables.

4.4.1 Interpretation of MNL models

Table 4.4 summarizes the results of our MNL estimations. Results reported in column 2 and 3 of Table 4.4 indicate that parameter estimates and instruments are similar for models that include a reduced TOD specification (six parameters) versus models that included a full TOD specification (1260 parameters).

In addition, in contrast to estimation databases available to airlines, our database contains information about both booking and departure dates. Consequently, we are able to include different price variables as a function of advance purchase periods. These variables represent the average price paid by consumers for a specific itinerary origin, destination, carrier, level of service (i.e., non-stop/direct, single connection, double connection), product type (i.e., high yield (business) or low-yield (leisure), and purchase period (i.e., 0-6 days in advance, 7-20 days in advance, >21 days in advance). Results, showed in columns 2 and 4 of Table 4.4, indicate that including an average high-yield and low-yield prices for different advance purchase periods produces an increase of the model's fit. The adjusted R-square increases from 0.1800 to 0.1820. Results also demonstrate that price sensitivities vary as a function of advance purchase periods, with those purchasing high-yield products being less price sensitive than those purchasing low-yield products (across any advance purchase periods) and those purchasing closer to departure being less price sensitive.

Finally, results reported in column 5 of Table 4.4 show that using ln(fare) instead of normal fare produces a better fit. The adjusted R-square increases from 0.1820 to 0.1827. However, using ln(fare) provide less intuitive value of time (VOT) estimates. Indeed, the value of time estimate is lower for high-yield leisure products purchased 21 or more days in advance of departure ($28) than for low-yield leisure products purchased 21 or more days in advance of departure ($30). For these reasons, we use fare segmented by days from departure for the remaining model estimations.

4.4.2 Interpretation of NL models

For computational considerations, we estimate all two-level and three-level NL models nested by carrier, level of service (nonstop versus connecting) and departure time of day (grouped into morning flights from midnight - 9:59 AM, afternoon flights from 10 AM - 3:59 PM, and evening flights from

	MNL 6 TOD price	MNL 1260 TOD price	MNL 6 TOD price×DFD	MNL 6 TOD ln(price)×DFD
Avg. HY price ($)				
0-6 DFD			-0.00327	-1.31
7-20 DFD	-0.00331	-0.00332		-1.39
21 + DFD			-0.00564	-2.45
Avg. LY price ($)				
0-6 DFD			-0.00642	-1.82
7-20 DFD	-0.00653	-0.00657	-0.00765	-1.95
21 + DFD			-0.00904	-2.28
Elapsed time (min)	-0.00371	-0.00387	-0.00363	-0.00382
# connections	-2.68	-2.66	-2.82	-2.71
Direct flight	-2.17	-2.14	-2.35	-2.36
Regional jet or prop (ref.)	0	0	0	0
WB or NB	0.378	0.377	0.379	0.389
Online (ref.)	0	0	0	0
Codeshare	0.292	0.284	0.305	0.290
Interline	-0.184	-0.185	-0.187	-0.207
$\widehat{\mu}$ (residuals)	0.00161	0.00161	0.00199	0.422
LL base model	-26,781,806	-26,258,983	-26,718,431	-26,690,436
LL corrected model	-26,776,700	-26,253,933	-26,711,078	-26,687,947
Adj. ρ^2	0.180	0.196	0.182	0.183
Value of time ($/hr)				
0-6 DFD (high/low)			$67 / $34	$52 / $38
7-20 DFD (high/low)	$67 / $34	$70 / $35	$67 / $28	$49 / $35
21+ DFD (high/low)			$39 / $24	$28 / $30

Table 4.4: MNL results

Note: LL= log likelihood, WB=wide-body; NB=narrow-body; LL(0) = -32,652,846.05; TOD = time-of-day; DFD = days from departure; All parameter estimates with the exception of some TOD coefficients are significant at the 0.01 level; Adj. ρ^2 = 1 - (Final LL - #Attributes) / LL(0); Value of times for ln(price) calculated for $300.

4 PM - 11:59 AM)[9], using un-instrumented price and a reduced TOD specification. The results of these models are summarized in 4.5. Only the logsum parameters and measure of model fit are shown as the parameter estimates from the other variables provide similar behavioral interpretations as those obtained from MNL models.

The logsum parameters associated with all two-level NL models are less than 1 and statistically significant at the 0.01 level. The logsums associated with the carrier and level of service nests are both close to 1 (0.9603 for both NL models) and provide a similar model fit ($\rho_0^2 = 0.1818$ for both NL models). The logsum associated with the time of day NL model ($\mu_{TOD} = 0.7890$) provides the best improvement in model fit, particularly when compared to the MNL model ($\rho_0^2 = 0.1826$ compared to 0.1820 for the MNL model). Three of the 3-level NL models are also theoretically valid. The NL model with an upper-level nest by TOD and lower-level nest by carrier ($\mu_{TOD} = 0.8518$ and $\mu_{Carrier} = 0.7572$) provides the best fit with the data among all of the 2-level and 3-level NL models ($\rho_0^2 = 0.1830$).

The logsum parameter estimates and measures of model fit for Coldren's models are shown in the bottom of Table 4.5. In contrast to Coldren's models, we found two additional NL models that were theoretically valid (although the corresponding logsum coefficients were all close to 1). With the exception of the 2-level carrier model, the logsum coefficients obtained from our data and Coldren's data are remarkably similar, and both datasets show the 3-level NL model with an upper nest for TOD and lower nest for carrier best fit the data.

4.4.3 Interpretation of OGEV models

The NL models indicate that there is strong inter-product competition between alternatives that share similar departure times. The OGEV model may be able to better capture this inter-product correlation better than the NL models. In order to estimate OGEV models, we need to create time periods and decide whether the OGEV model should incorporate two adjacent periods (corresponding to T=1 in Equation (4.3.4)), three adjacent periods (corresponding to T=2), or more. Table 4.6 shows the distribution of passengers' departure times in our estimation data and how we created time periods for OGEV models that created hourly time groupings and grouping for two

[9] These are the same time of day grouping used in Coldren and Koppelman (2005b); note that Coldren's data had no flights that departed from midnight to 5 AM.

	MNL	2-level NL models			3-level NL models					
		Carrier	LOS	TOD	TOD LOS	TOD Carrier	LOS Carrier	Carrier LOS	Carrier TOD	LOS TOD
Our Models										
$\mu_{carrier}$		0.9603				0.7572	0.9563			
μ_{LOS}			0.9603		0.7664		0.9672			
μ_{TOD}				0.7890	0.8496	0.8518		NTV	NTV	NTV
LL convergence	-26,718,431	-26,716,760	-26,717,706	-26,690,629	-26,686,156	-26,678,207	-26,717,622			
Adj. ρ^2	0.1817	0.1818	0.1818	0.1826	0.1827	0.1830	0.1818			
Coldren's										
$\mu_{carrier}$		0.8498				0.7370				
μ_{LOS}			NTV		0.8080					
μ_{TOD}				0.8167	0.8248	0.9375	NTV	NTV	NTV	NTV
Adj. ρ^2	0.2830	0.2837		0.2833	0.2833	0.2848				

Table 4.5: NL results and comparison with Coldren's models

Note: LL Zero =-32,652,846 ; LL Conv = Log likelihood at convergence; LOS=level of service; TOD= time of day (nesting by three departure time periods: 12 midnight - 9:59 AM; 10 AM - 3:59 PM; 4 PM - 11:59 PM); NTV = not theoretically valid. Model that fits the data the best in both our study and ColdrenâĂŹs study is highlighted. All logsum parameters are significant at the 0.01 level.

hour periods. Note that all of the time periods fit the "hourly" or "two-hour" definition except for the last time period which combines departures into a longer period.

Departure times	# Passengers	% Passengers	Hourly	Two hours
5:00 - 5:59 AM	188,885	1.9%	TP1	TP1
6:00 - 6:59 AM	866,02	8.6%	TP2	
7:00 - 7:59 AM	826,315	8.2%	TP3	TP2
8:00 - 8:59 AM	732,289	7.3%	TP4	
9:00 - 9:59 AM	576,637	5.7%	TP5	TP3
10:00 - 10:59 AM	642,073	6.4%	TP6	
11:00 - 11:59 AM	674,584	6.7%	TP7	TP4
12:00-12:59 PM	637,975	6.4%	TP8	
1:00 - 1:59 PM	618,462	6.2%	TP9	TP5
2:00 - 2:59 PM	587,411	5.9%	TP10	
3:00 - 3:59 PM	638,908	6.4%	TP11	TP6
4:00 - 4:59 PM	595,851	5.9%	TP12	
5:00 - 5:59 PM	798,367	8.0%	TP13	TP7
6:00 - 6:59 PM	607,23	6.1%	TP14	
7:00 - 7:59 PM	522,531	5.2%	TP15	TP8
8:00 - 8:59 PM	252,176	2.5%	TP16	
9:00 - 9:59 PM	116,179	1.2%	TP17	TP9
10:00 - 10:59 PM	89,491	1.5%	TP18	
11:00 - 11:59 PM	37,562			
12:00-12:59 AM	19,258			
1:00 - 1:59 AM	6,731			
TOTAL	10,034,935	100%		

Table 4.6: Time periods included in OGEV models

Results for the OGEV models (specifically the allocation and logsum parameters and model fit statistics) are shown in Table 4.7. Columns 2 to 4 contain results for OGEV models that incorporated two adjacent periods (or T =1). The model that incorporated hourly time groupings fit the data better than the model that incorporated two-hour time groupings. However, the behavioral interpretation of the model changes when different time groupings are used. In the hourly model, the allocation parameters are approximately equal (0.585 vs. 0.415) with a slightly higher correlation associated with alternatives that depart in earlier versus later periods. In the two-hour model, the applications are further apart (0.355 vs. 0.645) and show a higher correlation with alternatives that depart in later versus earlier periods. Both models fit the data better than the three-level NL TOD | carrier model ($\rho_0^2 = 0.1830$ for NL vs. 0.1833 and 0.1832 for OGEV models). This result contradicts the result found by Coldren, who estimated an OGEV model with T=1 using larger time periods of three hours. In his case, his three-level TOD | carrier model slightly fit the data better than his OGEV model with T=1. It

	Our models		Coldren's Model		Our models		Coldren's Model
	Two Alloc. (T=1) Hourly	Two Alloc. (T=1) Two Hour	Two Alloc. (T=1) Three Hour	Three Alloc. (T=2) Hourly	Three Alloc. (T=2) Two Hour Model 1	Three Alloc. (T=2) Two Hour Model 2	Three Alloc. (T=2) Two Hour
μ	0.7747	0.7461	0.7932	0.7217	0.7469	0.7223	0.7586
τ_0	0.5851	0.3548	0.7785	0.3221	0.6312	0.6105	0.6752
τ_1	0.4149	0.6452	0.2215	0.2403	0.3688	0.2397	0.0728
τ_2				0.4376	0.00*	0.1497	0.4376
LL convergence	-26,667,089.6	-26,672,315.0	-1,557,214	-26,660,741.2	-26,672,582.4	-26,672,457.8	-1,556,869
Adj. ρ^2	0.1833	0.1832	0.2834	0.1835	0.1831	0.1831	0.2836

Table 4.7: Summary of OGEV results and comparison with Coldren's models

Note: *Allocation parameter is not statistically different than 0, LL Zero =-32,652,846; Coldren's Two Allocation model includes six nests (5-6:59 AM; 7-9:59 AM; 10 AM -12:59 PM; 1-3:59 PM; 4-6:59 PM; 7-10:59 PM). Coldren's Three Allocation model includes eight nests (5-6:59 AM; 7-8:59 AM; 9-10:59 AM; 11 AM-12:59 PM; 1-2:59PM; 3-4:59PM; 5-6:59 PM; 7-10:59PM).

is unknown whether an OGEV model estimated using smaller time period grouping on his data would have outperformed his NL model.

Similar results are observed for the shown in columns 5 - 8 for the OGEV models that incorporated three adjacent periods (or T =2). Using our data, the OGEV model with hourly time groupings fits the data the best (over both the OGEV model with hourly grouping and T=1 and the three-level NL TOD | carrier model). The OGEV models using two-hour groupings were ill behaved, and converged to local (vs. global) optima. The first OGEV model (shown in column 7) had one allocation that converged to zero, which is equivalent to an OGEV model with two hour groupings and two adjacent periods (T=1); however, the log likelihood associated with this model (-26,672,582.42) is more negative (and does not fit the data as well as) the OGEV model with two adjacent time periods and two hour groupings (-26,672,314.97); this indicates that the OGEV model shown in column 7 has converged to a local optimum. We restarted this OGEV model using different starting values; the results are shown in column 8. This model allocates alternatives to all nests, but also has a log likelihood value that is more negative (-26,672,457.77) than the OGEV model with two adjacent time periods and two groupings; this indicates that this model has also converged to a local (vs. global) optimum. Indeed, the likelihood function appears to be quite flat near the optimum, with (very different allocations) providing similar log likelihood values. For completeness, Coldren's OGEV model with T=2 (that used three-hour time groupings) is shown in the last column. This model better fits the data than his OGEV model with T=1; however, none of his OGEV models better fit his data than his three-level NL TOD | carrier model.

To summarize, our OGEV models with hourly time periods slightly fit the data better than our three-level NL TOD | carrier nest; however, there is evidence that OGEV models with different allocations parameters (corresponding to different behavioral interpretations) provide similar model fits and that the log likelihood function near convergence is flat, which increases computational times. For example, our NL models converged in approximately four hours whereas our OGEV models converged in approximately 30 hours. For implementation purposes, we recommend that OGEV models with hourly time groupings be estimated for T=1 and T=2 and that the allocation parameters be constrained to be equal across nests. If the model fits between these OGEV models are similar with the three-level NL TOD | Carrier model, we recommend the simpler NL structure be used for implementation.

4.4.4 Assessment of CNL and NetGEV models

Our results from the NL and OGEV models are similar to those obtained from Colden's analysis. Given the similarities with Coldren's data, it is useful to examine which of Coldren's CNL and NetGEV models better fit his data over his 3-level time | carrier NL model. Table 4.8 reports the model fits (ρ_0^2) for seven CNL and NetGEV models Coldren estimated and his best-fitting NL and OGEV models. Only three of the seven CNL and NetGEV models (highlighted in the table) fit the data better than the 3-level time | carrier NL model. Interestingly, all of these advanced NetGEV models maintained the 3-level nesting structure with time as the upper-level nest and a lower level carrier nest. None of the CNL models he estimated fit the data as well as his 3-level time | carrier NL model. Stated another way, the multi-level nesting structure of the 3-level NL or NetGEV model better captures inter-product competition than the single-level CNL nesting structure that includes carrier and time nests.

Based on an assessment of Coldren's models, we conclude that the marginal improvement in model fit seen with his NL-OGEV (T=2) time | carrier model compared to his 3-level time | carrier NL model ($\rho_0^2 = 0.2855$ versus $\rho_0^2 = 0.2848$) does not justify the estimation of CNL and NetGEV models on our data. Indeed, the largest improvement in model fit for Coldren's data as well as ours is observed in transitioning from a MNL model to the 3-level time | carrier NL model ($\rho_0^2 = 0.2830$ vs. 0.2848 for Coldren and $\rho_0^2 = 0.1817$ vs. 0.1830 for our data). We thus select the 3-level time | carrier NL and the OGEV models that include hourly time periods and constrained allocation parameters as our preferred models.

Model	Adj. ρ^2
Best NL model: 3-level NL time | carrier	0.2848
Best OGEV model: OGEV (T=2)	0.2836
CNL and NetGEV models	
Weighted NL time | carrier	0.2842
Weighted NL time | carrier, time | service	0.2848
Nested weighted nested logit (NWNL)	0.2848
NL-OGEV (T=1) time | service	0.2835
NL-OGEV (T=1) time | carrier	0.2852
NL-OGEV (T=2) time | carrier	0.2855
NL-OGEV (T=2) time | carrier, time | service	0.2852

Table 4.8: Results from Coldren's CNL and NetGEV models

A comparison of the MNL model and the 3-level time | carrier NL model is given in Table 4.9. Both models have been corrected for price endogeneity using the control function method. We use similar instruments for both models but vary the functional form used to ensure instruments remain valid. Results show that the value of time estimates are lower for the 3-level time | carrier NL model than for the MNL model. The most important difference is for the high-yield leisure products purchased 0 to 6 days in advance of departure: the value of time changes from \$67 with the MNL model to \$67 with the the 3-level time | carrier NL model.

4.5 Limitations and contributions

There are several limitations of our research. First, our sample is non-representative in the sense that low cost carriers are under-represented. We are therefore implicitly assuming that those customers who purchase tickets on low-cost carriers have similar itinerary preferences as those who purchase on major carriers. Our ticketing database provides no information about the customers who purchased the ticket, preventing us from examining differences in preference based on trip purpose and socio-economic factors.

Nonetheless, our analysis provides several important contributions. First, we find strong evidence that inter-alternative correlation patterns have remained stable over the past 10-15 years for itinerary choice models estimated for domestic U.S. itineraries; further, these inter-alternative correlation patterns do not appear to be heavily influenced by the mix of distribution channels in the estimation database. Second, models that incorporate correlation across itineraries that share similar departure times result in significant improvements in model fit. Among the GEV model structures estimated, those structures that included an upper-level departure time of day nests consistently fit the data better. There are several GEV models that provide similar model fit statistics: these include the 3-level time | carrier NL models and OGEV models. Third, incorporating increased substitution across itineraries that share similar departure times is important even when more refined departure time of day preferences are incorporated into the utility function.

Our paper also contributes to the literature by examining the robustness of value of time estimates when different discrete choice models are used to model itinerary choices. We find that using price versus ln(price) provides more intuitive behavioral interpretations and that price sensitivities vary as a function of advance purchase periods.

	MNL 6 TOD price×DFD	3-level NL 6 TOD price×DFD
Avg. HY price ($)		
0-6 DFD	-0.00327	-0.00324
7-20 DFD	-0.00327	-0.00324
21 + DFD	-0.00564	-0.00526
Avg. LY price ($)		
0-6 DFD	-0.00642	-0.00581
7-20 DFD	-0.00765	-0.00691
21 + DFD	-0.00904	-0.00808
Elapsed time (min)	-0.00363	-0.00297
# connections	-2.82	-2.38
Direct flight	-2.36	-1.99
Regional jet or prop (ref.)	0	0
WB or NB	0.379	0.318
Online (ref.)	0	0
Codeshare	0.305	0.249
Interline	-0.187	-0.116
μ_{TOD}		0.854
$\mu_{carrier}$		0.756
$\widehat{\mu}$ (residuals)	0.00199	0.00218
LL base model	-26,781,805.82	-26,678,207
LL corrected model	-26,711,079	-26,671,051
Adj. ρ^2	0.182	0.183
Value of time ($/hr)		
0-6 DFD (high/low)	$67 / $34	$55 / $31
7-20 DFD (high/low)	$67 / $28	$55 / $26
21+ DFD (high/low)	$39 / $24	$34 / $22

Table 4.9: Comparison of MNL and NL results

Note: LL= log likelihood, WB=wide-body; NB=narrow-body; LL(0) = -32,652,846.05; TOD = time-of-day; DFD = days from departure; All parameter estimates with the exception of some TOD coefficients are significant at the 0.01 level; Adj. ρ^2 = 1 - (Final LL - #Attributes) / LL(0).

There are several research extensions we are planning to pursue as part of future work. These extensions are explained in details in Chapter 5.

Bibliography

Abbe, E., M. Bierlaire, and T. Toledo (2007). Normalization and correlation of cross-nested logit models. *Transportation Research Part B: Methodological 41*(7), 795–808.

Airport Cooperative Research Program (ACRP) (2013). *ACRP Report 98: Understanding Airline and Passenger Choice in Multi-Airport Regions* Washington, D.C.: National Academy of Sciences.

Atmosphere Research Group (2012). *The Future of Airline Distribution: A Look Ahead to 2017.* Special report prepared for the International Air Transport Association (IATA). Cambridge, MA: Atmosphere Research Group.

Bureau of Transportation Statistics (BTS) (2016). *Airline Origin and Destination Survey (DB1B).* `http://www.transtats.bts.gov/Tables.asp?DB_ID=125`.

Coldren, G. M. (2005). *Modeling the Competitive Dynamic among Air-travel Itineraries with Generalized Extreme Value Models.* Ph.D. Dissertation, Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL.

Coldren, G. and F. Koppelman (2005a). Modeling the proximate covariance property of air travel itineraries along the time-of-day dimension. *Transportation Research Record: Journal of the Transportation Research Board* (1915), 112–123.

Coldren, G. M. and F. S. Koppelman (2005b). Modeling the competition among air-travel itinerary shares: GEV model development. *Transportation Research Part A: Policy and Practice 39*(4), 345–365.

Coldren, G. M., F. S. Koppelman, K. Kasturirangan, and A. Mukherjee (2003). Modeling aggregate air-travel itinerary shares: Logit model development at a major US airline. *Journal of Air Transport Management 9*(6), 361–369.

Fosgerau, M., McFadden D, and and M. Bierlaire (2013). Choice probability generating functions. *Journal of Choice Modelling 8*, 1–18.

Daly, A. and M. Bierlaire (2006). A general and operational representation of generalised extreme value models. *Transportation Research Part B: Methodological 40*(4), 285–305.

Garrow, L. A., G. M. Coldren, and F. S. Koppelman (2010a). MNL, NL, and OGEV models of itinerary choice. In *Discrete Choice Modelling and Air Travel Demand: Theory and Applications*, Chapter 7, pp. 203–252. Ashgate Publishing: Aldershot, United Kingdom.

Garrow, L. A. (2010b). *Discrete Choice Modelling and Air Travel Demand: Theory and Applications.* Ashgate Publishing: Aldershot, United Kingdom.

Garrow, L. A. and J. P. Newman (2010c). Network GEV models. In *Discrete Choice Modelling and Air Travel Demand: Theory and Applications*, Chapter 5, pp. 137–174. Ashgate Publishing: Aldershot, United Kingdom.

Government Accountability Office (GAO) (2003). *Airline Ticketing: Impact of Changes in the Airline Ticket Distribution Industry.* GAO Report 03-749.

Guevara, C. A. (2010). *Endogeneity and Sampling of Alternatives in Spatial Choice Models.* Ph.D. Dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.

Guevara, C. A. (2015). Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research Part A: Policy and Practice 82*, 240–254.

Guevara-Cue, C. A. and P. Navarro (2013). Control-function correction with weak instruments in logit models. Working Paper, Universidad de los Andes, Chile.

Howard, C. (2016). Email communication between Chris Howard at the Airlines Reporting Corporation and Laurie Garrow on February 22, 2016.

IdeaWorksCompany (2015). Airline ancillary revenue projected to be $59.2 billion worldwide in 2015. press release `http:`

//www.ideaworkscompany.com/wp-content/uploads/2015/11/Press-Release-103-Global-Estimate.pdf. [Online; accessed 2-February-2016].

Jacobs, T. L., L. A. Garrow, M. Lohatepanont, F. S. Koppelman, G. M. Coldren, and H. Purnomo (2012). Airline planning and schedule development. In *Quantitative Problem Solving Methods in the Airline Industry: A Modeling Methodology Handbook.* Part of the Fred Hillier International Series on Operations Research and Management Sciences. Vol. 169. Eds. C. Barnhart and B. Smith. New York: Springer. pp. 35-100.

Koppelman, F. S., G. M. Coldren, and R. A. Parker (2008). Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B: Methodological 42*(3), 263–273.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In: Zarembka,P. (Ed.), *Frontiers in Econometrics.* Academic Press, New York, pp. 105–142.

McFadden, D. et al. (1978). *Modelling the choice of residential location*, in *Spatial Interaction Theory and Residential Location*, edited by A. Karlqvist, *et al.* Amsterdam: North-Holland, 75–96. California.

Nesset, E. and Ø. Helgesen (2014). Effects of switching costs on customer attitude loyalty to an airport in a multi-airport region. *Transportation Research Part A: Policy and Practice 67*, 240–253.

Newman, J. P. (2008a). *Normalization and Disaggregation of Networked Generalized Extreme Value Models.* Ph.D. Dissertation, Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL.

Newman, J. P. (2008b). Normalization of network generalized extreme value models. *Transportation Research Part B: Methodological 42*(10), 958–969.

Newman, J. P. (2016). Larch documentation. Discrete choice modeling software available at http://larch.readthedocs.org/en/latest/model.html.

Seelhorst, M. and Y. Liu (2015). Latent air travel preferences: Understanding the role of frequent flyer programs on itinerary choice. *Transportation Research Part A: Policy and Practice 80*, 49–61.

Small, K. A. (1987). A discrete choice model for ordered alternatives. *Econometrica: Journal of the Econometric Society*, 409–424.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica: Journal of the Econometric Society 65*, 557–586.

Train, K. E. (2009). *Discrete Choice Methods With Simulation.* Cambridge, UK: University Press.

Vovsha, P. (1997). The cross-nested logit model: Application to mode choice in the Tel-Aviv metropolitan area. *Transportation Research Record, 1607*, 6–15.

Wen, C.-H. and F. S. Koppelman (2001). The generalized nested logit model. *Transportation Research Part B: Methodological 35*(7), 627–641.

Williams, H. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A 9*(3), 285–344.

Chapter 5

Conclusions and Future Research

This dissertation contains two main bodies of work. In the first part, Chapter 2, we examine how the determination of optimal loading plans for cargo aircraft can help airlines to reduce their fuel and handling operation costs. In the second part, Chapter 3 and Chapter 4, we present the results of a series of itinerary choice models that account for price endogeneity.

These two bodies of work share one major common point. They combine mathematics and statistics with economics to assist airlines in solving complex and real industry problems and to help them to make better decisions. In this sense, both projects are applications of management sciences in the field of air transportation. Discrete choice modeling and operations research are two important disciplines for airlines. Researchers from these two communities have generally different perspectives and research priorities. Yet, the air travel community would gain in a deeper collaboration between these two communities. In that sense, this dissertation can help practitioners and researchers to see the usefulness of both disciplines for the airline industry.

However, each of these studies has its own conclusions and recommendations for future research, which are outlined in the following sections.

5.1 Airline container loading problem

5.1.1 Major conclusions

In the second chapter, a new mixed integer programming model for the Airline Container Loading Problem with pickup and delivery (ACLPPD) is presented. By recognizing that cargo aircraft can visit multiple destinations, our formulation is the first to model cargo transport as a series of trips that consist of one or more legs with pickups and deliveries. The consideration of a series of trips compelled us to model the handling operations on the ground, and to adapt the basic constraints of the Weight & Balance Problem to the multi-leg trip. Our model is based on international standards and is valid for most commercial operators. We defined our objective function as the sum of fuel and handling operation costs. This function is expressed in money terms, which makes its interpretation easier for practitioners. We provided evidence that the Airline Container Loading Problem is NP-complete. To the best of our knowledge, this had never been done before.

Our results demonstrated that it was possible to quickly find near optimal or excellent loading plans. Based on real data, we compared our solution to the solution obtained manually by the loadmasters and found that our approach led to substantial savings with respect to current typical practices.

5.1.2 Directions for future research

There are some directions for future research.

First, the ACLPPD has been showed to be NP hard, which opens the way to heuristics. Although our results demonstrated that it was possible to quickly find excellent solutions, it has been showed that adding a second leg significantly increased the complexity of the problem. Depending on the number of containers to be loaded, the number of airports to be visited or the nature of the cargo (dangerous goods), we cannot expect to be able to solve quickly all practical problem instances to optimality. The development of heuristics could enable practitioners to still find good solutions for difficult problems in reasonable time. Our model can be used to test the quality of different heuristics.

Second, the ACLPPD exhibits symmetry and it is well known that the presence of symmetry can have a significant negative effect on the performance of the branch-and-bound algorithm. Different approaches exist to reduce the symmetry of a problem. One can reformulate the problem or we can use symmetry breaking methods. CPLEX Optimizer offers techniques for symmetry breaking and for this research, we decided to let CPLEX decide the level of symmetry breaking. However, it would be interesting to further explore the symmetry issue as well as the different methods of reducing symmetry.

Another research extension might help to overcome an important limitation of our model. Indeed, one limitation of our model is to assume that we know all the containers and pallets (ULDs) to be delivered. However, at the time of the elaboration of the loading plans, the loadmaster has often only a rough idea of what will be loaded at intermediate airports. A future extension could be to transform our model into a stochastic model in which the ULDs are associated to random variables.

Another direction for future research that would be interesting is to investigate the applicability of our model in the case of aircrafts that carry both cargo and passengers. Indeed, in their willingness to maximize revenues, many airlines will additionally transport cargo on scheduled passenger flights. Optimizing the loading plans of these aircrafts poses additional challenges and merits further attention. As mentioned in Chapter 2, over the past years, more attention has been paid to the problem that precedes ACLPPD by considering how to optimize freight loading within ULDs independently of the routing of the aircrafts. An interesting problem would be to analyze the two problems simultaneously.

Finally, although our model considers a large set of realistic constraints, our discussions with industry practitioners revealed that some extensions to add specific constraints would be useful. As explained in Limbourg et al. (2012), when the aircraft is not loaded at full capacity, extra caution must be taken. In particular, it is common to pack the shipments around the center of gravity. A "packing" constraint could be added in our model to ensure that no ULD is too far away from the center of gravity. As additional examples, new constraints could be also added to take into account what loadmasters call "indents" and "overhangs". When long pieces have to be loaded on a pallet, it can happen that the pieces exceed the length of the pallet. This is an overhang. In order to optimize the loading, loadmasters create a corresponding "indent" pallet that contains an empty space that will allow to position both pallets next to each other. These are two, of many, examples of how the problem can be modified to incorporate additional business rules.

5.2 Itinerary choice models

5.2.1 Major conclusions

In Chapter 3, we presented an itinerary choice model that is consistent with those used by industry and corrects for price endogeneity using a control function that uses several types of instrumental variables. We estimated our model using a database of more than 3 million tickets provided by the Airlines Reporting Corporation. Results based on Continental U.S. markets for May 2013 departures showed that models that fail to account for price endogeneity overestimate customers' value of time and result in biased price estimates and incorrect pricing recommendations. To the best of our knowledge, this is the first itinerary choice model representative of those used in practice that accounts for price endogeneity.

As part of the itinerary choice model developed in Chapter 3, we modeled consumer departure time preferences using a continuous time of day representation using the initial part of a Fourier sine and cosine series, a formulation first proposed by Abou-Zeid et al. (2006) and applied to airline itinerary choice by Koppelman et al. (2008). Due to the size of our analysis database, we have been able to estimate highly re- fined departure time of day preferences that are a function of distance, the direction of travel, number of time zones crossed, departure day of week, and itinerary type (inbound, outbound, oneway). To the best of our knowledge, these are the most refined time of day preferences published in the literature. The results of this model (which includes 1260 parameter estimates) can be used to assess the impacts of congestion pricing, slot restrictions, and other demand-management strategies.

Within the aviation literature, there has been different formulations used to model continuous time-of-day preferences. In particular, Carrier (2008) proposed the use of a truncated function that accounts for periods of the day in which little no flights depart. In Appendix A3 of Chapter 3, we compare these two continuous time-of-day formulations to a discrete one and show that in our application the 24-hour cycle better fits the data. However, due to differences observed in utilities for itineraries that depart very early or very late in the day, we recommend that both continuous time-of-day specifications be tested for a particular application.

In Chapter 4, we extend the multinomial logit (MNL) model we developed in Chapter 3 to more advanced discrete choice models that incorporate inter-product substitutions; these models include the nested logit and ordered generalized extreme value (OGEV) models. We compared our results to those reported in the literature for United Airlines using 2000 data. Despite the many structural changes that occurred in the airline industry over the past 15 years (including industry consolidation, the increased use of online distribution channels, and product de-bundling trends), we found strikingly similar results to those observed based on 2000 data. Based on our estimations, we recommend using either a three-level NL model that includes upper-level nests for time of day and lower-level nests for carriers or (potentially) an OGEV model with hourly time periods in which each itinerary is allocated equally between three nests.

Finally, in contrast to estimation databases available to airlines, our database contains information about both booking and departure dates. Consequently, our price formulation is the first within the itinerary choice literature to account for different price sensitivities as a function of advance purchase periods. This is important, as the airline industry can use our results to account for different customer segments as revealed through high-yield and low-yield booking curves to evaluate the profitability of airline schedules.

5.2.2 Directions for future research

There are various natural directions for future research build upon the itinerary choice model estimated in this dissertation. These directions are the subject of a postdoctoral research proposal. The central idea is to further enhance these itinerary choice models for consumer welfare applications.

The consumer welfare literature uses BLP methods (for Berry, Levinson, and Pakes, see Berry et al. (1995, 1999, 2004)), which is similar in spirit to the mixed logit methods from the transportation community. Both BLP and mixed logits are simulation-based methods to allow the researcher to incorporate the distribution of customer preferences (if customer attributes are available). As shown by Armantier and Richard (2008) in the context of consumer welfare models for air travel, this simulation-based framework is also convenient for addressing issues related to missing or incomplete data. This is important, because I can use their methods to account for the fact that some low cost carriers are underrepresented/missing in my data.

The paper by Armantier and Richard (2008) is arguable the most advanced consumer welfare application to air travel demand in the literature. However, like the itinerary choice models used by industry, their formulation suffers from two major limitations: (1) it does not include customer characteristics, and (2) makes strong assumptions about the underlying decision process consumer use when choosing an itinerary.

Therefore, a first research extension that would be interesting to explore is to expand our itinerary choice models to include consumer characteristics by linking a random sample of the ticketing database to a targeted marketing (TM) database that contains detailed and extensive socioeconomic information about individuals. Several large firms assemble detailed socioeconomic and financial information for adults in developed countries and sell these data to firms who use the data to customize their marketing campaigns. The level of detail attained by TM datasets has increased dramatically with the explosion of electronic financial transactions and includes household characteristics (e.g., number of adults and number of children in the household, household income) and individual characteristics (e.g., age, gender, marital status, education). These data have been previously used in several transportation applications and have been successfully linked to ticketing data by ARC and American Airlines. To date, there have been very few applications of BLP that have been based on data that links purchases to individual consumer-level characteristics. To my knowledge, there are no itinerary choice models used by industry or reported in the consumer welfare literature that incorporate consumer characteristics (such as age, gender, and household income). The approach for linking TM data to purchase transactions provides a roadmap for overcoming an important limitation that has faced demand modelers.

A second research extension arises from the recognition that despite the wave of mergers and acquisitions - and the potential harm these may have on consumers - there has been little research into whether state of the art consumer welfare models accurately capture consumers' decision-making process. Both the itinerary choice and welfare estimation literatures use discrete choice models, which are based on utility-maximizing theory and make strong assumptions about individuals' underlying decision-making processes. For example, both generate a "complete" set of itineraries and assume individuals look at and evaluate all itineraries in the choice set. This may not be a realistic assumption, especially for long-haul trips where the number of itineraries returned by an online search engine can easily exceed 300 non-stop and single connections in one day. For example, recent work by Collins et al. (2012, 2013) used a “mock-up” of an airline website for a long-haul international market and found strong evidence that individuals use sorting tools to eliminate alternatives with particular attributes (such as itineraries with connections or high fares). Their findings are consistent with recent research that has suggested consumers often resort to a two-stage “consider-then-choose” decision process when the choice task is too complicated. In a first (screening-) stage individuals may identify a consideration-set of alternatives that needs further evaluation (Hauser (2014), Leong and Hensher (2012)) or eliminate attributes they find irrelevant (Hensher et al. (2005)). In a second stage they make a choice using a standard compensatory model on the outcome of the screening-stage. Different screening rules have been proposed in the literature, which I will explore as part of my postdoctoral research. Consideration-set screening (CSS) rules assume that individuals first evaluate whether the attribute levels of an alternative are acceptable and next use combination rules to combine the attribute level evaluations across attributes (e.g., see Jedidi and Kohli (2005), Gilbride and Allenby (2004), Hauser (2014)). Attribute non-attendance (ANA) rules assume that individuals ignore part of the attributes and is particularly popular in the discrete choice modeling community (e.g., see Hensher and Rose (2009), Hensher et al. (2012), Hess et al. (2013) for examples within transportation).

The objective of the postdoctoral research project is to benchmark the state of the art welfare estimation models based on classic discrete choice modeling with one that parameterizes choice sets as a function of consumer characteristics and search behaviors. A series of online experiments will be conducted in order to examine how sorting and other online behaviors can be used as an indication of consumer preferences. This information can be used to generate choice sets as a function of consumer search characteristics such as sorting behavior, trip characteristics such as the length of the trip or how far in advance of the departure date the individual paid for the trip, and socioeconomic information such as age, gender, and household income. It will then possible to develop a model to generate choice sets as a parameterized function of these characteristics, potentially using methodologies similar to those of Başar and Bhat (2004); note that these choice sets can be probabilistically assigned to consumers to capture different behaviors across a similar group of consumers. This project provides an important bridge between classic and behavioral economics by examining how sensitive state of the art models that predict consumer welfare effects are to underlying assumptions about consumers’ decision-making behavior. I hypothesize that in today’s markets, where hundreds of options can be shown to consumers, it is critical to incorporate different decision rules related to choice set generation and that failure to do so will result in very different demand predictions and consumer welfare estimates.

Finally, these models will be used to assess the effects of industry consolidation on firms and consumers. In the airline context, consolidation can result in additional product offerings (through coordination of flight schedules and capacity), improved operational performance, and lower costs. However, it is unclear whether an increase in prices after a merger is due to firms charging more for this expansion of additional products and services or firms acting in a more coordinated (and monopolistic) fashion, or other external factors (such as increases in production costs). My modeling approach allows me to disentangle these effects. For example, if firms are acting in a more monopolistic fashion we should see capacity reductions associated with rising prices. I hypothesize that theoretical models of competitive behavior for oligopolies and government policies based on Herfindahl-Hirschman Index (HHI) and other concentration metrics may not reflect levels of coordination possible in today's information-driven society. Consequently, increased industry consolidation may have resulted in firms being able to mimic monopoly behavior thereby harming consumers. Welfare impacts will be examined as the industry has recently consolidated and created new joint ventures (e.g., Delta - Virgin America). Moreover, these effects can be related to traditional measures of concentration used by regulators to understand whether, in the face of consolidation, we lack adequate measures. Even though the focus is the airlines, this theoretical approach can easily be applied to any oligopoly industry that may be trying to further consolidate. Ultimately, our work may provide regulators with yet an additional tool to help analyze potential mergers within oligopoly industries.

Bibliography

Abou-Zeid, M., T. Rossi, and B. Gardner (2006). Modeling time-of-day choice in context of tour-and activity-based models. *Transportation Research Record: Journal of the Transportation Research Board* (1981), 42–49.

Armantier, O. and O. Richard (2008). Domestic airline alliances and consumer welfare. *The RAND Journal of Economics 39*(3), 875–904.

Başar, G. and C. Bhat (2004). A parameterized consideration set model for airport choice: An application to the san francisco bay area. *Transportation Research Part B: Methodological 38*(10), 889–904.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society 63*(4), 841–890.

Berry, S., J. Levinsohn, and A. Pakes (1999). Voluntary export restraints on automobiles: Evaluating a trade policy. *American Economic Review 89*(3), 400–430.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Econometrica: Journal of Political Economy 112*(1), 68–105.

Carrier, E. (2008). *Modeling the Choice of an Airline Itinerary and Fare Product Using Booking and Seat Availability Data.* Ph. D. thesis, Massachusetts Institute of Technology, Department of Civil and Environmental Engineering.

Collins, A. T., S. Hess, and J. M. Rose (2013). Choice modelling with search and sort data from an interactive choice experiment. *Transportation Research Part E: Logistics and Transportation Review 56*, 36–45.

Collins, A. T., J. M. Rose, and S. Hess (2012). Interactive stated choice surveys: A study of air travel behaviour. *Transportation 39*(1), 55–79.

Gilbride, T. J. and G. M. Allenby (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science 23*(3), 391–406.

Hauser, J. R. (2014). Consideration-set heuristics. *Journal of Business Research 67*(8), 1688–1699.

Hensher, D. A., J. Rose, and W. H. Greene (2005). The implications on willingness to pay of respondents ignoring specific attributes. *Transportation 32*(3), 203–222.

Hensher, D. A. and J. M. Rose (2009). Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. *Transportation Research Part E: Logistics and Transportation Review 45*(4), 583–590.

Hensher, D. A., J. M. Rose, and W. H. Greene (2012). Inferring attribute non-attendance from stated choice data: Implications for willingness to pay estimates and a warning for stated choice experiment design. *Transportation 39*(2), 235–245.

Hess, S., A. Stathopoulos, D. Campbell, V. O'Neill, and S. Caussade (2013). It's not that I don't care, I just don't care very much: Confounding between attribute non-attendance and taste heterogeneity. *Transportation 40*(3), 583–607.

Jedidi, K. and R. Kohli (2005). Probabilistic subset-conjunctive models for heterogeneous consumers. *Journal of Marketing Research 42*(4), 483–494.

Koppelman, F. S., G. M. Coldren, and R. A. Parker (2008). Schedule delay impacts on air-travel itinerary demand. *Transportation Research Part B: Methodological 42*(3), 263–273.

Leong, W. and D. A. Hensher (2012). Embedding multiple heuristics into choice models: An exploratory analysis. *Journal of choice modelling 5*(3), 131–144.

Limbourg, S., M. Schyns, and G. Laporte (2012). Automatic aircraft cargo load planning. *Journal of the Operational Research Society 63*, 1271–1283.