Universitat Autònoma de Barcelona

Three Essays on Incorporating Service and Product Attributes into Economic Models

Charles Baird Howell

Doctoral Dissertation

Three Essays on Incorporating Service and Product Attributes into Economic Models

Charles Baird Howell

Author

Emili Grifell-Tatjé

Thesis Supervisor

Ph.D. program in Economics, Management and Organization (DEMO)

Facultat d’Economia i Empresa

Universitat Autònoma de Barcelona

2020

Doctoral Dissertation

Three Essays on Incorporating Service and Product Attributes Into Economic Models

CHARLES BAIRD HOWELL
Digitally signed by CHARLES BAIRD HOWELL
Date: 2020.08.20 11:48:31 +02'00'

Charles Baird Howell

Author

EMILI GRIFELL TATJE - DNI 39310048N
Firmado digitalmente por EMILI GRIFELL TATJE - DNI 39310048N
Fecha: 2020.08.20 12:00:13 +02'00'

Emili Grifell-Tatjé

Thesis Supervisor

Ph.D. program in Economics, Management and Organization (DEMO)

Facultat d'Economia i Empresa

Universitat Autònoma de Barcelona

2020

Acknowledgements

Acknowledging and thanking everyone who has been part of helping me complete the three-year process of earning a PhD is quite a task in itself! Before acknowledging individuals, I'd like to recognize the support I've received from several "communities". First, the DEMO community and help from all my fellow researchers, whether sitting next to me in the office or in various seminars. Second, the research community in Loughborough who made me feel welcome and helped my work move forward, in spite of the virus lockdown. Finally, on a more personal level, the kendo community in Barcelona who provided both training and comradeship.

Turning to individual acknowledgements, I´d like to begin with my wife Sandi, without her support and encouragement, I never would have started this journey. Next, I'd like to thank my thesis supervisor, Emili Grifell-Tatjé, for all his efforts in getting me this far. I came into the dissertation with a good idea of what I wanted to do, and why, but without many of the tools to make it happen. With Emili's guidance, I found the right ways to pose the questions, define (and refine) the methodology, run the analysis and interpret the results. I also appreciate his willingness to introduce me to the productivity and efficiency community, it has helped me establish my own contacts and network in the research field.

Beyond my advisor, there are many others in the department I would like to thank. The DEMO Coordinator Miguel Ángel García Cestona and the MMOBE Coordinator Pere Ortín were both a huge help to me as I worked through the program. The professors I interacted with as a student, or worked with as a teaching assistant, are too numerous to name, but I appreciate all their help and guidance. A special thanks to Mireia Cirera in the department office, without her help I would never have been able to navigate the bureaucratic channels of the University.

I would like to thank the members of my committee, Hans-Martin Niemeier, Luis Orea and Jonathan Wareham, for taking the time to improve my thesis and allowing me to complete this journey. Beyond the thesis itself, I would also like to thank the three of you for your support, both academic support in my research and personal support in friendship. You were all an important part of the process, whether in Athens, Barcelona, Bremen, or Loughborough.

Finally, although I have already noted my wife Sandi´s support and encouragement, I would like to add that I appreciate her willingness to pack up and leave home and family to pursue this goal with me. Especially in light of the Covid-19 situation, where the typical disruptions to life due to a PhD program were made worse by travel complications and longer than expected absences. To her I say, I´ll be home soon and we´ll be on to the next challenge together!

Abstract

The goal of this dissertation is to explore the effect of changing service characteristics on product cost, price and eventually profitability, or in the case of a product that is a service, changing the characteristics of service delivery. The specific focus is not on physical characteristics that can differentiate products (in the airline industry a 200-mile flight vs 700-mile flight, or in the shelving industry 600 lb. holding capacity vs 400 lb.), but on the characteristics of the transaction and the services provided as part of the product. Throughout this thesis we refer to these generically as “output attributes” and use the term “product” to mean either a product or service. Using a cost-based approach and the tools of productivity and efficiency analysis, we develop a measure of product differentiation by output attribute and then use that measure to explore how the inclusion of the measure affects three different economic models. The empirical setting for all three chapters is the US domestic airline industry.

The first chapter focuses on cost, introduces the measure of differentiation through attributes, and applies it to a model studying price dispersion. Previous studies show that competition affects price dispersion but disagree on the direction of change. To explain this contradiction, this chapter introduces a novel method of measuring product heterogeneity which collects the cost impact of the level of differentiation in a market. Applying this method, we find that the response of price dispersion to changes in competition is conditioned by differentiation. We empirically test our method on 73,981 observations of airfare data from 2002 through 2016. This chapter’s contributions are extending knowledge on the effect of competition on price dispersion and introducing a method of measuring market differentiation.

The main objective of the second chapter is to study the effect of product differentiation on price formation in the airline industry using a hedonic model. For this purpose, we introduce the concept of a core product and examine how differentiation by output attribute beyond the core effects pricing and mark-ups. We measure differentiation from the core product using a Konüs type index of differentiation that is based on cost functions. In this chapter we study how using this index, measures of market power, and controlling for core product cost, can improve the stability and results of a hedonic price model. The model is empirically tested on 103,980 observations of quarterly US domestic airfare data between 2002 and 2016.

The third chapter turns to the question of profitability. The US Airline industry swung from $31 billion in losses over the eight-year period 2002 to 2009 to $90 billion in profits over the seven-year period 2010 to 2016. This reversal of fortune was not driven by specific airlines but can be observed industry wide. To fully explore this change we decompose annual profitability change over the fifteen-year period and analyze the economic drivers, with a focus on the driver of product differentiation. The inclusion of product differentiation, an economic driver not typically included in the analysis of financial performance, is one of the primary contributions of this chapter. This chapter introduces a novel method of decomposing profitability change and uses the combination of a standard cost function paired with a non-standard revenue function. To fully understand this effect, we examine these financial performance measures for twenty individual carriers in the US airline industry between 2002 and 2016.

Contents

Introduction

Introduction

Motivation

The motivation for this thesis is very much rooted in my experience and background as a product management and pricing professional. In my twenty years managing the pricing function for a mid-size US manufacturer I was constantly attempting to balance product cost, service cost, customer demands, customer value and price. Like many firms manufacturing a commodity type product for the business-to-business market, what differentiates the firm is not as much the product itself, but the sales and service package that surrounds the product. This can range from discrete services, such as training, installation, or design services, to the harder to define services like depth of catalog, ease of transaction or nationwide sales coverage. The question that always came to my mind when evaluating a service request was, can we generate more revenue if we provide this service? And maybe more importantly, can we generate more revenue than it will cost to provide the service? This thesis is an attempt to develop quantitative methods and measures to answer those questions.

Differentiation

In this section I would like to briefly discuss the theoretical, empirical, and operational background of differentiation with the intent of defining the terminology and general model used in this thesis. The concept of differentiation is generally understood, but not well defined. As noted by Sharp and Dawes (2001), differentiation is a concept which has very vague meanings, is often referenced without a formal definition, and has a number of non-complementary operationalizations and measures. In this thesis, through new quantitative methods and measures, I define a model and terminology that begins to bring that fuzziness into focus and allow important business questions to be answered.

The theoretical stream of literature that studies differentiation is well-developed. One of the more relevant to our research is the model outlined in Lancaster (1966) who assumes that products are valued not only for their physical attributes, but their non-physical attributes as well. In this work utility is derived not from the product itself, but from the properties or characteristics of the product, and the other products and services consumed in conjunction. Differing mixes of properties or characteristics generates differentiation.

Moving this concept towards empirical application, Caves and Williamson (1985, p. 113) provide a formal condition for differentiation as "… buyers must recognize that goods ("brands") belonging to a product class are close substitutes for one another but face only relatively poor substitutes with goods outside the class." Their work also identifies one of the complexities of empirical study of differentiation, a complexity that ties back to the concepts presented in Lancaster (1966). Using an example of shipping containers, they note that buyers are relatively uninterested in measurements such as density, or container dimensions, but in the relative durability or volume they can hold. Extending the condition for differentiation to a definition we use in this thesis, Sharp and Dawes (2001) state that differentiation exists when a firm's offering is preferred at sometimes, or by some customers all the time, over a rival firm's offering. The empirical stream of literature in this area is less well-developed, partly due to the complexity of measurement. Filling this gap in empirical methods is the primary contribution of this thesis, in addition to the topics tackled in each of the essays.

Due to this complexity in measurement and definition, empirical work in differentiation is often limited to considering a single characteristic. For example, in a study of the video retail industry Seim (2006) measures the return to differentiation based on location choice and spatial differentiation. However, based on data availability, application of the model requires inferences of profitability based on market structure and of costs based on density of businesses. Another approach often used in this empirical stream is to combine all product differences as a single measure of quality (*q),* and the assumption that production cost is a strictly increasing convex function c(*q).* This is the approach taken in Boik and Takahashi (2020) who study price discrimination in the cable industry. In this work they measure differentiation *q* as the number of channels provided.

To operationalize the discussion of characteristics we introduce a modified version of a product model developed by Kotler (1967) as Figure 1. In the center is the *core benefit* that is satisfied by consuming the product or service. For our example of airlines, this would be the need to physically move from one location to another location. The next level, *core product*, is a version of the product containing only those *core characteristics* absolutely necessary for it to function. In the terminology of Caves and Williamson (1985), this core product defines a product class of close substitutes that meet the core need. In our example these core characteristics would include a seat, a take-off, landing, and traveling some distance. All airlines providing service between two city pairs provide these same core characteristics.

The next ring in the model is defined by *output attributes*. These extend the product beyond core characteristic to include the attributes that buyers expect when they purchase the product from a particular firm. These output attributes are available to all buyers of the firm´s product and can create differentiation between firms. In the airline example this includes attributes such as on-time performance, or frequency of flights. The final ring, augmented product, adds characteristics and services that allow a firm to differentiate the product between its own customers, such as service class upgrades or preferential seating. The first and third chapters of this thesis focus on the second ring: output attributes, while core characteristics are the topic of the second chapter.

Methodology

All three chapters of this thesis use the same underlying methodology, but with adjustments to fit the empirical application at hand in the chapter. The thesis proposes a new measure of product differentiation based on a Konüs (1939) approach and measures the ratio of two cost or revenue functions that differ only in their level of core characteristic or output attributes. The Konüs method was first used to measure the cost difference to changes in input price but can be applied to other factor changes as well. We use a cost or revenue function, rather than observed cost or revenue, as differences between firms can arise from a number of sources. Using a cost or revenue function allows us to focus only on the change due to characteristics or attributes.

Beyond the contribution of introducing a new measure of product differentiation, we propose a methodology to use the measure and put into practice some of the concepts from Figure 1. In the first chapter we focus on output attributes and measure the degree of differentiation in the market. The second chapter discusses the core characteristics, using the cost of this core to improve the results of a hedonic estimation. The third chapter returns to output attributes, measuring their difference between firms and how it effects cost, revenue and profitability.

This methodology is fully outlined in the first chapter where we apply it to output attributes with a cost function. In the second chapter we again apply a cost function but focus on core characteristics. In the final chapter we return to output attributes but apply the methodology to both a cost and revenue function.

The Chapters

In the first chapter we introduce and fully develop the methodology that forms the basis for this thesis. This methodological development, the use of a Konüs style ratio of two cost functions, is the primary contribution of this chapter. Using the measure, we determine the cost of differentiation by output attribute for individual firms, and by extension the level of heterogeneity in a market. In an empirical application of this method we study the relationship between competition and price dispersion, controlling for the level of output attribute driven market heterogeneity. The current literature on the relationship between competition and price dispersion is inconclusive, with studies finding contradictory results. We contribute to that literature by showing that once market heterogeneity is accounted for, competition reduces price dispersion under low to moderate heterogeneity. We also find that at high levels of heterogeneity, the relationship can turn positive. In the chapter we discuss how these results contribute to regulatory policy analysis and provide firms with strategic direction. This chapter was presented at the 15th Aviation Student Research Workshop in Bremen Germany, the 33rd Jornadas de Economia Industrial in Barcelona Spain and the OR60, Operations Research Conference in Lancaster UK. It was also awarded the 2018 European Prize in Aviation Economics and Management.

In the second chapter, the focus moves towards price with the objective of studying the effect of product differentiation on price formation in the airline industry. To do this, we introduce a modified hedonic model that measures the effect of characteristics on the mark-up over the cost of the core product, rather than directly on price. Referring to Figure 1, cost of the core product includes the cost of providing the core characteristics in the second ring. We extend the heterogeneity measure introduced in the first chapter to isolate this core cost and show how using this can improve the hedonic model in terms of stability and results. These improvement form one of the primary contributions of this chapter. This modified hedonic model is also a contribution as it always firms to make pricing decisions relative to their specific markets. A further contribution is operationalizing the concepts of the core characteristics in Figure 1, which has implications for further research on differentiation strategy and market entry. This chapter was presented at the EWEPA XVI, European Workshop on Efficiency and Productivity Analysis in London UK and 2019 Aviation Management and Economics Conference (AMEC) in Vienna Austria.

In the first two chapters the focus was on cost and price as separate elements. In the third chapter we bring these together and discuss profitability in the US aviation industry. This industry saw significant change between 2002 and 2016, transforming from a sector that was losing money, to one that was seeing its highest profit levels in decades. To understand the drivers of that change we decompose cost, revenue and profitability change into five sources, output attribute change, efficiency change, technical change, activity level change and input price change. In this chapter we make three primary contributions. First, our inclusion of output attributes and product differentiation as a component of financial measures change. Second, our profitability decomposition pairs a standard cost function and a non-standard revenue function, allowing the decomposition of effects not typically captured. This novel decomposition method can be applied to any industry where the non-standard revenue function is appropriate. Finally, in the empirical application of these methods we examine a period of change in the US airline industry that is relatively unexplored. This chapter was presented virtually at the 17th Aviation Student Research Workshop, Hong Kong Polytechnic University and at the (vNAPW XI) Virtual North American Productivity Workshop, the University of Miami Herbert Business School, USA.

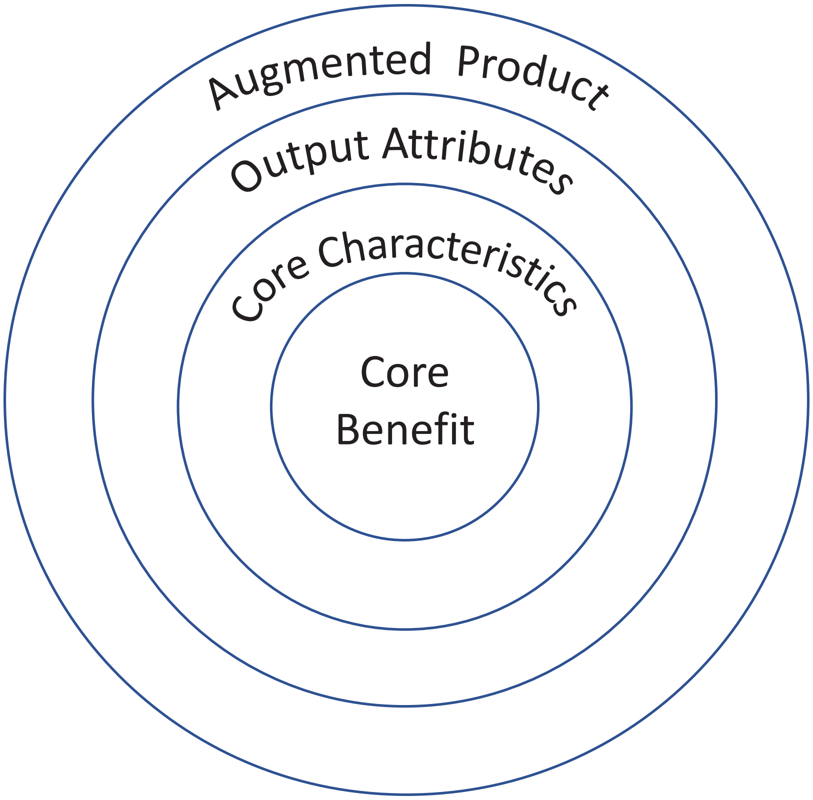


Figure 1: The product and its characteristics adapted from Kotler (1967)

Chapter 1 - Market Heterogeneity and the Relationship Between Competition and Price Dispersion: Evidence from the US Airline Market

Abstract

Previous studies show that competition affects price dispersion but disagree on the direction of change. To explain this contradiction, we introduce a novel method of measuring product heterogeneity which collects the impact on cost of different levels of differentiation in a market. Applying this method, we find that the response of price dispersion to changes in competition is conditioned by differentiation. We empirically test our method on 73,981 observations of airfare data from 2002 through 2016. The results contribute towards extending knowledge on the effect of competition on price dispersion and introduce a method of measuring market differentiation.

1.1. Introduction

The relationship between market structure and price dispersion, defined as a variation in the price paid for the same product at the same time, has interested economists since merchants first posted prices. Dispersion can come from a variation of price between sellers of the same good or from variation in the price between the customers of a single seller. The second form, price discrimination, is generally believed to only be possible when a firm has market power. Textbook theory would argue that as a market becomes more competitive both forms of dispersion should evaporate. In the competitive market a firm becomes a price taker and loses the ability to price discriminate, with price moving towards marginal cost and dispersion falling.

This textbook theory however is not reflected in real world markets where persistent price dispersion can be observed in markets of all structures. The relationship between market structure and price dispersion has been documented in competitive markets for products such as groceries (Kaplan et al. 2016 and Eden 2018), retail gasoline (Shepard, 1991 and Lach and Moraga-González, 2017), auto insurance (Dahlby and West, 1986) and U.S. domestic airline fares (Borenstein and Rose, 1994; Gerardi and Shapiro, 2009 and Dai, Liu, and Serfes 2014). Despite all these, the relationship between market structure and dispersion remains unclear. This is particularly true in the last industry noted above, U.S. domestic airline fares, where the three articles cited came to widely differing conclusions on the topic.

In one of the first studies on market structure and dispersion of airline fares, Borenstein and Rose (1994) model a market that has two forms of price discrimination, with one form generating a higher level of price dispersion. Testing this model on a pooled data set from the second quarter of 1986 they find that the higher-level form dominates, and that price dispersion increases with competition. These findings are contradicted by Gerardi and Shapiro (2009) who utilize panel data and find that competition strictly reduces price dispersion, arguing the textbook model. Finally, in a more current article on the subject Dai et al. (2014) find an inverse-U-shaped relationship between competition and price dispersion with dispersion initially growing as a concentrated market becomes competitive, but then declining at higher levels of competition.

In a monopoly market, discrimination-based price dispersion can be observed as the result of value-of-service pricing, with consumers who gain the most utility from a product paying a higher price. In a seminal article on the topic of price discrimination in a free-entry market Borenstein (1985) showed that in markets with differentiated brands a firm can also discriminate based on strength of brand preference. The article finds that sorting on brand preference leads to larger market price dispersion than sorting on consumer value. Following Borenstein´s lead, in this chapter we refer to "product" as the good produced by all firms in the market and "brand" as one firm´s output.

In many industries the production of output may have associated qualitative attributes that can differentiate output, Ray and Mukherjee (1996) refer to these as "Output Attributes". In this chapter, using output attributes as the method of differentiation, we extend Bornstein's model, replacing brand preference with output attribute preference. Markets that have a wider range of output attributes between brands are referred to as differentiated or heterogeneous markets, and those that have a narrow range of output attributes between brands as homogenous markets. We find that accounting for the difference in level of output attributes between airlines helps explain the contradictions on price dispersion found by previous research.

The U.S. domestic airline market has many characteristics that make it ideal for empirically testing the relationship between market structure and price dispersion. Defining the product as a non-stop[1] flight between two airport pairs, airlines can sort consumers by product valuation and price discriminate by offering tickets with a range of characteristics: advance purchase discounts, round trip restrictions and refundability to name a few. There is also heterogeneity between airlines in measurable output attributes such as on-time performance, departure frequency and load-factor that allow for differentiation. In addition, the airline market is dynamic and has a wide range of structures, including routes with one carrier, routes with two carrier and more competitive routes with up to seven carriers. We can consider dispersion in air fares in two ways, brand-dispersion being the dispersion in fares charged by a single airline on a route, or product-dispersion as the variance between all prices on the route regardless of airline. In this article we are defining price dispersion as product-dispersion, the dispersion of all prices available on a route.

This study is focused at the micro level; accordingly, competitive changes are driven by a shift in the specific firms serving the market. The response of firms currently in the market may differ depending on how similar their product is to that of a new competitor. Our study shows that the response of price dispersion to a change in market structure is not uniform but depends on the level of similarity between firms in the market. Ignoring this can result in the contradictory results found to date when examining the effect of competition on price dispersion.

Introducing a novel cost-based method of measuring market differentiation we examine the airlines on a route and identify the cost premium driven by providing a higher level of output attributes. Price dispersion arising from cost differences is not considered discrimination, so this premium is removed from the fare price creating a more homogenous product. Remaining price

---

[1] A non-stop flight is defined as a flight between two airport pairs with no intermediate stops between. This is in contrast to a “direct” flight which does not change flight numbers but may have an intermediate stop.

dispersion is measured and analyzed. Our method is empirically tested against a data set of 73,981 observations from the US domestic airline market over 60 quarters between 2002 and 2016. The results are used to answer three specific research questions. First, how much product price dispersion is driven by cost differences resulting from a wide range of output attributes? Second, how does the level of competition affect price dispersion after accounting for output attribute cost-driven price dispersion? Third, is the direction and magnitude of the effect of competition on price dispersion dependent on the level of output attribute differentiation in the market?

This article contributes to the line of research on the effects of market structure on price dispersion and the ability of firms to price discriminate. We also introduce a novel method of measuring the level of product and market heterogeneity. Our findings form the basis for policy recommendations, adding a factor to be considered beyond simply market concentration when regulators examine potential mergers. This article is structured as follows. Section 2 provides a discussion and background of price dispersion and market structure literature. In section 3 we develop the model of Borenstein (1985) to this article and introduce a measure of differentiation based on output attributes and its application to fare prices. Section 4 outlines the data and variable construction whereas section 5 presents the results. Conclusions, policy recommendations and ideas for further research are covered in section 6.

1.2. Background

1.2.1. Price Dispersion

The study of price dispersion has a long history. In part, because its existence runs counter to the textbook model of most market forms. Standard theory allows for price dispersion in a monopoly market, but only when the firm can sort consumers and price discriminate, while under perfect competition all firms accept the single market price. Between the two polar cases, most oligopoly models allow for little to no dispersion in price, while in models of monopolistic competition, price dispersion is primarily driven by product cost differences. Theoretical explanations for price dispersion include costly information search, uncertain demand, price discrimination, and cost differences. Most empirical studies have focused on either product level dispersion that comes from firms selling the same good at different prices, or brand level dispersion that comes from a single firm practicing price discrimination.

In one of the earliest studies, focused on the role of information and search, Stigler (1964) stated it well "Price dispersion is a manifestation - and, indeed, it is the measure - of ignorance in the market." The study focused on dispersion in product price and finds that the cost of search, changing supply and demand conditions and the difficulty of updating prices generate the observed dispersion. Stigler notes that including the "terms of sale" and "services provided" would likely explain some amount of dispersion, but not all. A foreshadowing of the concept of market heterogeneity and dispersion. In the same vein of information and search Stahl (1989) outlines a model of temporal price dispersion, also at the product level. His model separates consumer into two types, informed and uninformed, and shows that sales and intentional fluctuation in prices allow retailers to capitalize on uninformed consumers and generates price dispersion.

The second source of dispersion studied, price discrimination, can create differentials within a firm when a firm´s customers vary in their willingness-to-pay. Both Borenstein (1985) and Holmes (1989) introduce models of third-degree price discrimination in differentiated product markets, showing that market power is not a requirement for discrimination. Holmes (1989) formalized the concept introduced by Borenstein (1985) that firm level elasticity of demand can be decomposed into two components, industry elasticity and cross-price elasticity. Industry elasticity is based on consumer valuation of the basic product whereas cross-price elasticity is based on the price of a substitute. This distinction is important for our analysis because these two forms result in a different amount of price dispersion. Next, we discuss cost-based dispersion before returning to and expanding on Borenstein (1985) later in this section.

Price variances arising from difference in cost are not considered price discrimination, nor dispersion. This distinction makes empirical work on the topic difficult as distinguishing cost-based differentials from discriminatory differentials is complex. Shepard (1991) works around this by taking advantage of a situation where firms differ in their ability to price discriminate, but not in their costs of production. The article compares price differentials between full-service and self-service gasoline at stations that provide both products, to the differentials between full-service and self-service at stations that only offer one or the other. Results show that the differential is greater in multi-product firms and demonstrates price discrimination beyond cost differences in a competitive market.

The perishable nature and uncertainty of demand for the product in aviation has led to studies of a particular form of price dispersion. Dana (1999) models a system where the firm must establish prices and quantities (inventory of seats) prior to knowing what demand will be. In this model, equilibrium is found by establishing multiple price points, but rationing the number of seats available at each point. One of the findings of this model is that as competition increases, the average price level falls and the degree of price dispersion increases.

In another study of the airline industry, Stavins (2001) tests for dispersion driven by price discrimination, using the marginal implicit price of ticket restrictions in interaction with market concentration. The work focuses on restrictions used to separate consumers by valuation of convenience and flexibility, Saturday-night stayover and advanced-purchase discounts. After controlling for cost-based and carrier effects, the model finds that the marginal effect of ticket restrictions is less on routes with higher concentration. In other words, the same ticket restriction yields a greater discount on a more competitive route. This result is attributed to the increase in competition for consumers with higher price elasticity of demand.

As discussed above, there are multiple theoretical explanations for observed price dispersion, costly information search, uncertain demand, price discrimination, and cost differences. Our work focusses on the last two of these. Cost differences due to product differences are not considered dispersion, so by measuring the cost difference driven by differentiation we contribute to that theoretical line. As described in the next section, we also contribute to the literature on price discrimination driven dispersion, and ultimately, what happens to both of these forms of dispersion as competition changes.

1.2.2. Market Structure and Price Dispersion

Borenstein (1985) builds on a generalized price discrimination model introduced by Salop (1979), by developing a characteristic space model of monopolistic competition. In this model brand characteristics and consumer preferences differ on only one dimension and are represented as equidistant points on a unit-circumference circle. The model, through reservation price, allows for consumers to differ in preferred brand and in basic utility for the product. This set-up goes on to show the interaction between the two types of price discrimination. The first is termed monopoly-type discrimination because only one brand can provide positive consumer surplus. Under this form the firm sorts consumers based on industry elasticity and basic valuation of the good. Those with a higher valuation are charged a higher price while the low valuation group receives a discount. Under the second form, competitive-type discrimination, multiple firms can provide consumer surplus. For this type it is more effective for firms to segment by cross-price elasticity and provide a discount to customers on the verge of switching brands. The model shows that price dispersion is greater under competitive-type discrimination than under monopoly-type. Applying this model to price dispersion in airline fares, Borenstein and Rose (1994) hypothesize that competitive-type discrimination will dominate and that price dispersion will increase as airline markets becomes more competitive. They test the model on a pooled cross-sectional data set from the second quarter of 1986 and find that dispersion increases on routes with more competition.

In contrast to this, Gerardi and Shapiro (2009) find that competition strictly reduces price dispersion in airline fares. Citing traditional microeconomic theory, they argue that as competitors enter the market incumbents will find it difficult to price discriminate. Utilizing panel data, they analyze the effect of competition on price dispersion in US airfare between 1993 and 2006 and find that competition has a negative effect on price dispersion. Reconciling their findings with those of Borenstein and Rose (1994) they point to the estimation method, cross-section versus panel, as the main reason for the difference.

In another study of airfares, also using panel data, Dai et al. (2014) find that the effect of competition on price dispersion is non-monotonic, taking on an inverse-U-shaped relationship. Pointing to two opposing effects on price, first a direct price effect which increases dispersion as a highly concentrated market becomes more competitive, and then an indirect quality effect that leads to reduced dispersion as less concentrated market becomes even more competitive.

The three studies noted above have all been on the US domestic market, however this question has been looked at in other markets as well. In a study of the market connecting the UK and Ireland, Gaggero and Piga (2011) find a negative relationship between competition and price dispersion. In their work they also discuss monopoly type and competitive type price discrimination, finding evidence of monopoly type. In Obermeyer et al (2013) the empirical setting for this same question is moved to the European airline market. Their results support the argument for a non-monotonic relationship with the effect of further competition depending on the current degree of concentration. They also link their results to efficiency and productivity analysis, finding that more efficient carriers are better able to differentiate fares. In a work that focusses more on price differentials than dispersion, Chandra and Lederman (2018) study the Canadian market. Categorizing travelers into three groups based on how much the value the trip and brand loyalty, they find that the relationship between competition and price dispersion is ambiguous.

This chapter contributes to this literature stream by considering a variable of market structure that is not typically included in this type of analysis, differentiation and the degree of firm heterogeneity in the market. In the works cited above, and in aviation literature in general, the more typical factors of market concentration and consumer heterogeneity are accounted for, but not the degree of firm differentiation. Some works account for this by using a control for an LCC carrier, but as we discuss in the next section, this is not a very precise control. As one of the factors that affects market structure, our measure of market heterogeneity can be applied to any research questions based on market structure.

1.3. Methodology

Based on the work of Borenstein (1985) we study how the presence of a wide or narrow range of output attributes can alter the response of price dispersion to a change in competition. Building on Borenstein’s model, we replace brand preference with output attributes and consumer preference for those attributes. In homogeneous markets where firms have a similar levels of output attributes monopoly-type discrimination dominates, while in heterogeneous markets that exhibit differentiation competitive-type discrimination is stronger.

For example, visualize the Seattle-Salt Lake City (SEA-SLC) route is served only by Delta and Alaska Airlines, and that they have similar output attributes. With little to differentiate between brands, consumers are sorted on reservation price as under monopoly-type discrimination. Compare this to a second route, Philadelphia to Raleigh-Durham (PHL-RDU) served by American Airlines and Frontier. On this route, if we assume market heterogeneity based on differing output attributes, then the possibility of competitive-type discrimination exists. Using the introduced cost-based measure of differentiation by output attribute to categorize markets as heterogeneous or homogeneous we compare price dispersion on these two types of routes. This lets us explore our first research question; how much product price dispersion is driven by cost differences resulting from a wide range of output attributes.

Moving to a dynamic analysis, and our second research question, we now look at changes in the level of competition and examine the effect on price dispersion after controlling for differences in output attribute level. Returning to the SEA-SLC route, assume there is now an increase in competition coming from the addition of a third airline, JetBlue for example, which provides very similar output attributes to the others on this route. The model would predict a decrease in price dispersion following an increase in competition. However, if instead the change in competition on the SEA-SLC comes from a different airline with lower output attributes on this route, Frontier for example, the opportunity for competitive-type discrimination is created allowing customers to be sorted by output attribute preference. The model would now predict an increase in price dispersion following the increase in competition, as hypothesized by Borenstein and Rose (1994). Comparing the effect of competition on routes that vary in the cost-based measure of differentiation sheds light on the second question, how does the level of competition affect price dispersion after controlling for output attribute cost-driven price dispersion.

Comparing routes that vary in competition and in the level of cost-based differentiation also lets us answer our third research question; is the direction and magnitude of the effect of competition on price dispersion dependent on the level of output attribute differentiation in the market? Taken in combination, we believe that answering these questions will contribute to the understanding of the relationship between price dispersion and market structure and may be able to explain contradictions found by previous work.

An alternative approach to measuring cost-based differentiation by output attribute would be to add a control for the types of carriers in the market, for example Full-Service Carrier (FS), Low-Cost Carrier (LCC). We could then compare the effect of competition on price dispersion in markets with different mixes of these carriers. A market with all of one type would be more homogeneous, while those with a mixture more heterogeneous. There are however some drawbacks to this approach. First, there is no distinction between carriers in each group, this is especially relevant for the LCC group where the product can vary widely between carrier. Second, there would be no distinction between markets. Based on market size and number of rivals a carrier may adjust this product, that strategic change would be lost in a more blanket control indicator. Finally, the approach doesn’t allow for change over time as carriers adjust their product offering. Based on these drawbacks we believe measuring cost-based differentiation is a preferable method.

1.3.1. A Cost Approach.

When the production of a scalar output has associated qualitative output attributes, Ray and Mukherjee (1996) note that these attributes effect the maximum quantity of output producible from a given input bundle, or in the dual problem the minimum cost achievable. Ignoring these output attributes when measuring and comparing the cost of production processes can result in under or overestimating costs. Beginning with standard notation we denote output as $y \in \mathbb{R}_+$, the vector of output attributes as $q \in \mathbb{R}_{+}^{m}$, the vector of inputs as $x \in \mathbb{R}_{+}^{n}$, the vector of input prices as $w \in \mathbb{R}_{+}^{n}$ and the vector of output prices as $p \in \mathbb{R}_{+}^{i}$. Note that one output can be associated with multiple prices. The production possibility set that defines the output and associated output attributes possible for a given set of inputs is noted as

$$T = \{(x, y, q): x \text{ can produce } y \text{ with output attributes } q\}. \tag{1}$$

Given technology set T, we now define the set of inputs x required for every output quantity y with level of output attributes q as $L(y, q) = \{x: (x, y, q) \in T\}$. Technology set T defines all the possible input vectors that can produce a specific output at a given output attribute level. With the addition of input prices w, we move from input quantities to production cost and can define the least expensive bundle of inputs required to generate a specified quantity of output and level of output attributes. For an output quantity y, output attribute vector q and input prices w, the firm allocates the inputs to generate the required quantity of output at a minimum cost, defined as:

$$c(w, y, q) = min\{w^{T}x: x\ L(y, q)\}. \tag{2}$$

Including the transpose of the input price vector "T" we express minimum cost as $w^{T}x(q)$, where $w^{T}x(q) = c(w, y, q)$. As defined in (2), more inputs x would be required to create either a greater quantity of y at the same level q or an equal quantity of y with a greater output attribute level q. Comparing one output quantity y at two output attribute levels q and q^{0}, where $q \geq q^{0}$, we would find that $x(q) \geq x(q^{0})$ and $w^{T}x(q) \geq w^{T}x(q^{0})$.

From (2) we derive minimum cost of providing output quantity y, with output attributes q and input prices w. However, the cost between periods for a single firm, or between individual firms in a single period, can vary greatly based on scale of operations. For this reason, we move to a unit or average cost function to better understand the effect of changes in output attributes. The move to a unit or average cost function is also useful when applying cost changes to unit prices. With output defined as y we can find a unit or average cost function $ac(w, y, q)$ as,

$$ac(w, y, q) = \frac{c(w, y, q)}{y}. \tag{3}$$

Since our interest lies in capturing the cost premium of providing a greater level of output attributes, we measure cost differences between different attribute levels for the same firm and period. Extending (3) we define $ac(w_h, y_h, q_h) \;\; h = 1, \dots, k$ as the average minimum cost given w, y and output attribute level q, which defines the attribute level currently provided by a firm h in a specific market. Similarly, $ac(w_h, y_h, q^0) \;\; h = 1, \dots, k$ is the minimum cost for the same output quantity and price level, but with the lowest level of output attributes in the market q^0, where $q^0 = \{q_1^0, \dots, q_m^0\}$ and $q_j^0 = Min\left\{q_{jh}^0, h = 1, .., k\right\}, j = 1, .., m$.

1.3.2. Output Attribute Index.

In the literature, changes in average cost are thought to come from a limited number of sources. Capturing the change from the price of inputs, Konüs (1939) defines a theoretical input price index as a scalar function that compares input price vectors w^0 and w^1 for a given output y in the context of a cost function. Building on the Konüs approach, Grifell-Tatjé and Lovell (2015) discuss the drivers of unit cost change. They show that a difference in average cost can only come from two sources, a change in the price of inputs or a change in productivity. In this article we extend these concepts introducing a Konüs type output attribute index which collects the impact on the unit cost of changes between the level of output attributes q and q^0 as,

$$
\begin{aligned}
Q(w_h, y_h, q_h, q^0) &= \frac{ac(w_h, y_h, q^0)}{ac(w_h, y_h, q_h\,)};\;\; h\; = 1, .., k \\
&= \frac{c(w_h, y_h, q^0)}{c(w_h, y_h, q_h)}\;;\;\; h\; = 1, .., k.
\end{aligned}
\qquad (4)
$$

The index of output attribute differentiation defined in the first row holds output quantities and input prices equal, allowing measurement of the impact on unit cost of a change in the output attributes from q to q^0. The second row, after simplifying by dropping y, shows that the impact on total minimum cost of a change in the output attributes from q to q^0 is equal to the unit cost result from the first row. These results show that the output attribute index can be calculated on either minimum unit cost or minimum cost, meaning the model can be easily applied to the case of multiple outputs as well[2].

Recalling that q_h represents the observed level of output attributes for a firm in a market and that q^0 represents the lowest level of observed attributes in the market, then for all cases where $q_h \geq q^0$ and consistent with (4), $ac(w_h, y_h, q_h) \geq ac(w_h, y_h, q^0)$ and $c(w_h, y_h, q_h) \geq c(w_h, y_h, q^0)$, thus $Q(w_h, y_h, q_h, q^0)\; \leq\; 1$. When the index is 1 it signifies perfect market homogeneity or no differentiation in terms of output attributes. A value of 0.5 would indicate that firm h has a unit cost that is double that of providing the lowest level of attributes[3]. In section IV we discuss moving this measure to the market level to measure market heterogeneity. Note that

---

[2] In addition to multiple outputs, this method could be applied to a revenue function to capture differentiation by input attribute.

[3] The output attribute index can be formulated in the form of difference instead of ratio. The formulation is based on the first row of (4) and is given by $ac(w_h, y_h, q_h) - ac(w_h, y_h, q^0)\; \geq 0$

the output attribute index $Q(w, y, q, q^0)$ is bounded between 0 and 1 with the same rank of results as the Herfindahl-Hirschman Index (HHI) and the Gini coefficient.

There are several methods of estimating and constructing the empirical production frontier needed to estimate minimum cost $c(w_h, y_h, q_h)$. These methods generally fall into one of two classes, parametric and non-parametric, the preceding methodology could be applied to a production frontier developed by either method. In this article we apply the non-parametric data envelopment analysis (DEA) method, a method introduced by Charnes, Cooper, and Rhodes (1978)[4]. In this choice we follow the lead of Färe, Grosskopf, and Lovell (1985) who moved the non-parametric data envelopment analysis (DEA) approach to the economic context, and Ray and Mukherjee (1996) who introduced the inclusion of output attributes to DEA. In appendix 1.1. we outline the use of a parametric method to estimate the minimum cost $c(w_h, y_h, q_h)$ and the output attribute index defined in (4).

With our focus on finding the minimum cost for a given level of output and output attributes we utilize a cost minimization model with variable returns to scale. The minimum cost is found solving the DEA linear programming problem for one observation h with output quantity y_h and output attribute levels q.

---

[4] See Liu et al. (2013) for a general survey of DEA or Schefczyk (1993) for a detailed discussion of this technique to airlines.

$$c(w_h, y_h, q) = \min w_h^T x \quad s.t.$$

$$\sum_{l=1}^{k} \lambda_l x_{li} \leq x \quad i = 1, \dots, n$$

$$\sum_{l=1}^{k} \lambda_l y_l \geq y_h, \qquad \sum_{l=1}^{k} \lambda_l q_{lj} \geq q \quad j = 1, \dots, m \tag{5}$$

$$\sum_{l=1}^{k} \lambda_l = 1, \qquad \lambda_l \geq 0.$$

Minimum cost is solved for twice, first with q taking the value of observed q_h to find $c(w_h, y_h, q_h)$ and then as q^0 to find $c(w_h, y_h, q^0)$. Using the output attribute index $Q(w_h, y_h, q_h, q^0)$ defined in (4) we find the cost premium of providing a higher level of attributes.

We then remove this cost premium from price, allowing us to measure the amount of price dispersion resulting from differentiation by output attribute. Beginning with observed prices by market and quarter, $p_h(q_h)$ represents a vector of observed prices for firm h in a specific market and quarter. Noting the definition of the output attribute index from (4) we utilize the formula below to generate the vector $p_h(q^0)$ that equalizes prices for differences in output attributes between firms and removes any cost premium for differentiation by output attribute

$$p_h(q^0) = p_h(q_h) \cdot Q(w_h, y_h, q_h, q^0) \; ; \; i = 1, .., l; \; h = 1, .., k. \tag{6}$$

An alternative way of explaining formula (6) is that the unit margin, defined here as $\alpha_h = \frac{p_h(q_h)}{ac(w_h, y_h, q_h)}$, can be applied to the minimum unit cost with the lowest level of attributes to find an adjusted price as $p_h(q^0) = \alpha_h \cdot ac(w_h, y_h, q^0)$. With some rearrangement and recalling the definition of $Q(w_h, y_h, q_h, q^0)$, we have formula (6).

1.3.3. Estimating Market Price Dispersion and the Effect of Competition.

We now move the analysis from the firm to the market level, and the focus to changes in price dispersion over time. With $p(q)$ as the vector of all observed prices in a market and $p(q^0)$ as the vector of all adjusted prices, we define $S_{jt}(p(q))$ as observed price dispersion in the market *j* period *t* and $S_{jt}(p(q^0))$ as dispersion on adjusted prices. The method of measuring price dispersion will be defined in the next section. The calculation of price dispersion $S_{jt}(p(q^0))$ on prices that have been adjusted to remove the effect of differentiation is one of the primary contributions of our research. Hereafter we note price dispersion in market *j* at time period *t* as $S_{jt}(\cdot)$, to signify either $S_{jt}(p(q))$ or $S_{jt}\big(p(q^0)\big)$.

Following the structure of previous literature, we define $S_{jt}(\cdot)$ as dependent variables and *competition* (*Com*) as the independent variable. Like Borenstein and Rose (1994) and Gerardi and Shapiro (2009), we include dummy variable χ_{jt} as a control for the presence of a bankrupt carrier in the market. Market fixed effects are represented as υ_j and we include a full set of quarterly time dummies as γ_t to control for exogenous demand and cost effects. Residuals are captured by ε_{jt}.

As we are measuring product level price dispersion rather than brand level dispersion, we add a control variable for the carriers present in the market. Firms differ in their pricing strategies and vary in amount of price discrimination. When measuring brand level dispersion, firm strategy would be captured in the fixed effects, but in the case of product level dispersion it must be accounted for. To control for this, we add a dummy variable for each of the carriers represented in our study as δ_{ht}. The following panel regression is then estimated as a baseline,

$$S_{jt}(\cdot) = \beta_1\, Com_{jt} + \theta\, \chi_{jt} + \delta_{ht} + \gamma_t + \upsilon_j + \varepsilon_{jt}, \tag{7}$$

which measures the effect of competition on price dispersion over time as coefficient β_1, but does not account for the level of market heterogeneity.

To account for the level of market heterogeneity we add Q_{jt} to the estimation as the value of the output attribute index for market *j* in period *t*. In the previous section we defined $Q(w_h, y_h, q_h, q^0)$ as a firm-level output attribute index, moving this measure to the market level is discussed in subsection 4.5. The addition of this variable is unique to this stream of literature and comprises our second primary contribution. The following panel regression addresses our first two research questions and is estimated with the same control variables as (7),

$$S_{jt}(\cdot) = \beta_1\, Com_{jt} + \beta_2\, Q_{jt} + \theta\chi_{jt} + \delta_{ht} + \gamma_t + \upsilon_j + \varepsilon_{jt}. \tag{8}$$

Our third research question focuses on the effect of the interaction between competition and product differentiation on price dispersion. In specific, when the market is homogenous (low differentiation) the effect of an increase in competition is thought to reduce price dispersion, but when the market is heterogeneous (high differentiation) the effect of increasing competition is an increase in price dispersion. To capture this effect, we add the interaction of Com and Q_{jt} to equation (8) and estimate the following panel regression with the same control variables as (7),

$$S_{jt}(\cdot) = \beta_1\, Com_{jt} + \beta_2\, Q_{jt} + \beta_3\, (Com \cdot Q)_{jt} + \theta\chi_{jt} + \delta_{ht} + \gamma_t + \upsilon_j + \varepsilon_{jt}. \tag{9}$$

To address potential problems of endogeneity with the competition variables in formulas (7 - 9) we adopt the instrumental variable approach. Higher price dispersion can make a market more attractive to prospective entrants and this reverse causation can create positive bias in the least squares estimates of β. We tested for the presence of endogeneity with a Hausman Endogeneity test, including the predicted residuals from a first stage regression on Com. Although this test did not indicate endogeneity in competition, we have chosen to follow the previous literature and use the same instrumental variables to keep our results consistent. It also possible that endogeneity may arise from omitted variables, a problem also solved by the instrumental approach. We use market level variables of distance between end points, arithmetic and geometric mean of endpoint populations, total enplaned passenger and two variables introduced by Borenstein and Rose (1994) as instruments for competition.[5] We also used the Hausman test to check for endogeneity in our Q variable and found no evidence of endogeneity.

1.4. Data

In this section we provide a discussion of the sources and construction of the data used in the applied portion of this study. All data is provided by the Bureau of Transportation Statistics (BTS)[6], an independent statistical agency within the US Department of Transportation (DOT). The BTS collects data on traffic, passenger flow, employment, financial conditions and performance of US commercial aviation. It is a federally designated "principal statistical agency" and adheres to a set

---

[5] See Borenstein and Rose (1994) or Gerardi and Shapiro (2009) for full details on these instruments. In the empirical section instrumented variables are denoted by a hat.

[6] https://www.transtats.bts.gov/homepage.asp (accessed May 15, 2019)

of professional standards and operational practices designed to ensure the quality, integrity and credibility of their statistical activities.[7]

Our unit of analysis, the product, is defined as a non-stop, coach class, one-way or round-trip flight between two airport pairs. Non-stop refers to a flight with no intermediate stops between airport pairs. The market is defined by a set of airport pairs which we term a *route*. For example, three airlines service the route San Francisco (SFO) to Salt Lake City (SLC) with non-stop flights: Delta, Alaska Airlines and United Airlines. We measure the price dispersion of all products in the market. In the previous example, this would be all prices charged by Delta, Alaska and United on the SFO-SLC route. Following both Gerardi and Shapiro (2009) and Dai et al. (2014), in cities with multiple close-by airports that are easily substitutable, we put together the observations of the airports. For example, O´Hare (ORD) and Midway (MDW) are both located in the Chicago metropolitan area so observations for these airports are treated as a single location[8], therefore flights from San Francisco to O´Hare (SFO-ORD) or Midway (SFO-MDW) constitute a single route. This combination of close-by airports best fits our definition of a unique market.

1.4.1. Flight Operations, Output Quantity and Output Attributes

Flight operations data is used to generate our best practice production frontier and derive minimum costs. It is sourced from the BTS Form 41, which is collected for all large certified air carriers subject to the Federal Aviation Act of 1958. Traffic data is reported monthly and financial data is

---

[7] https://www.bts.dot.gov/learn-about-bts-and-our-work/statistical-methods-and-policies/statistical-policy-directives (Accessed June 1, 2019)

[8] The following close-by airports are combined in the results:DFW(Dallas–FortWorth) and DAL (Love Field); LGA (La Guardia), EWR (Newark) and JFK (J. F. Kennedy); AZA (Phoenix-Mesa Gateway) and PHX (Phoenix Sky Harbor); TPA (Tampa) and PIE (St. Petersburg Clearwater); DCA (Reagan) and IAD (Washington Dulles); ORD (O'Hare) and (MDW) Midway

collected quarterly. Data reported by carriers includes statistics on traffic, capacity, cost, profit and loss accounts and balance sheet accounts. From this data we specify the input and output sets used to generate the best practice production frontier.

When measuring airline output a number of studies have focused on the generation of output as capacity, measured as either ton miles available or available seat miles (see Assaf, 2011; Arjomandi and Seufert, 2014 and Lee and Worthington 2014), while others have focused on revenue generating measures (see Färe, Grosskopf, and Sickles, 2007 and Wang et al., 2014) such as revenue passenger miles (RPM) or revenue ton miles. As our focus is on passenger ticket prices and the cost of transporting those passengers, we have chosen to use RPM, calculated as the number of paying passengers multiplied by the miles travelled, as our measure of output (y).[9]

Three variables were chosen for output attribute vector q: i) on-time arrival performance, ii) flight frequency and, iii) load-factor. Numerous previous studies (see Borenstein, 1989; Douglas and Miller, 1974; Ippolito, 1981; Suzuki, 2000 and Gayle and Yimga 2018) have identified these attributes as important in differentiating air carriers. Beginning with on-time arrival, Gayle and Yimga (2018) showed that travelers would be willing to pay $1.56 per minute to avoid late arrival. We source on-time arrival data from the BTS On-Time Performance database which includes actual and scheduled arrival times for all non-stop domestic flights by major air carriers. From this data set we build a panel of quarterly on-time performance measures for each carrier for each of the sixty quarters analyzed. A flight is considered on-time if it arrives within fifteen minutes of

---

[9] By using only passenger miles, we are ignoring freight and mail as an output, however, because passenger revenue is typically 98% of total revenue for this group of carriers the impact of freight and mail is negligible.

scheduled arrived time. Average on-time arrival performance over the period was 80.6%, with a minimum of 11.5%, maximum of 100% and standard deviation of 8.3%.

Flight frequency is sourced from the T-100 Domestic Segment table, collected as part of Form 41. The T-100 is a 100% census of all non-stop segment data reported by US carriers. Flight frequency has been noted by Borenstein (1989) and Douglas and Miller (1974) as a product differentiator that increases the value of the brand to consumers, in particular to more inelastic consumers who place a higher value on time. We calculate it as the average number of daily flights provided by the carrier[10]. On the previously mentioned SFO-SLC market-route, in Q4 2016 United Airlines averaged 0.86 flights per day, Alaska Airlines 1.02 and Delta Airlines 2.37.

Load-factor, the proportion of aircraft seats with revenue paying passengers, is also sourced from the T-100 Domestic Segment table. High load factors can affect passengers in several ways. First, as noted by the study on demand factors by Ippolito (1981), high load factors make it less probable that a passenger will get their preferred flight time. In addition, a fully loaded flight will take longer to embark and disembark and can increase the likelihood of denied boarding. As passengers prefer a lower load factor, the output attribute is measured as (1 – load factor) to be consistent with other output attributes. Average load factor during the period under observation was 73% with a standard deviation of 12% points.

We measure the output attributes of on-time performance and load factor at the carrier level and flight frequency at the carrier-route level[11]. An extended description of output, on-time

---

[10] As carriers can enter or exit a market during a quarter, we find the average number of daily flights for each month the carrier was active in the market, then find the average for the quarter based only on the months they were active. For example, if Delta entered a market in March, its flight frequency for that quarter would be based only on the number of daily flights performed in March.

[11] Output attributes are measured as the difference between the observed value and the minimum value. For example, if United Airlines had the lowest on-time arrival performance at 0.765 and

performance, frequency, load factors and routes by carrier can be found in table I. Recall that load factor is not the traditional value but is measured as (1 – load factor). Output is presented as average passenger miles per quarter and route and we can see significant variation in values.

(Table I near here)

1.4.2. Inputs

Regarding inputs, it is generally agreed that airlines use fuel, labor, flight capital and purchased materials and services (see Oum, Fu, and Yu, 2005; Färe et al., 2007; Wu, He, and Cao, 2013 and Wang et al., 2014). These four inputs form our input quantity vector x and input price vector w. Fuel data is sourced from Form 41 schedule P-12(a) and is measured as total gallons of fuel consumed. The price is calculated as a ratio of total cost to gallons of usage. Regarding labor, we follow Wang et al. (2014) and Oum et al. (2005) and use the number of full-time equivalents (FTE) as the measure of labor. Total salary and benefits divided by the number of FTE provides the labor price. Form 41 Schedule P6 provides payroll data and Schedule P10 employee data.

The input flight capital has been measured in a few ways in the literature. Wang et al. (2014) compounded different types of aircraft into a fleet quantity index, while Wu et al. (2013) used the number of planes as an input. We are following a process similar to Färe et al. (2007) and develop the flight capital input, defined as the total number of seats available, based on the number of planes in service and the seat configuration utilized by the carrier. This data comes from Form 41, Schedule T2 (US Air Carrier Traffic and Capacity Statistics by Aircraft Type). The cost of capital comes from two sources, leasing rates and capital depreciation, both available in Form 41,

---

Alaska Airlines the best at 0.863, their output attribute for on-time arrival would be 0.00 and 0.098 respectively.

Schedule P52 (Quarterly Aircraft Operating Expenses). The price is calculated as the ratio of total cost of capital to quantity.

The final input is ground equipment and purchased materials. Following previous work, Färe et al. (2007), this input is calculated as total operating expenses less the cost of all other identified inputs. The result is deflated by the Bureau of Labor Statistics (BLS) producer price index (PPI) of air transportation support activities[12] to obtain quantities and the price is set equal to the index value. Table II provides a quarterly average of inputs by carrier and the average route. As with the outputs, we can observe a significant amount of variation.

(Table II near here)

1.4.3. Competition and Market Structure

Following previous work, we measure competition, defined in formulas (7 – 9) as (*Com),* using two variables, the Herfindahl-Hirschman index (HHI) and the number of carriers on a route. The HHI is a standard measure of concentration and is defined as the sum of the squares of market shares of all firms in the market[13]. With market share calculated as a fraction, the value varies from 0.0 to 1.0 with higher values indicating greater levels of concentration. The US Justice department guidelines for horizontal mergers consider an HHI below 0.15 as indicating a non-concentrated market, between 0.15 and 0.25 moderate concentration and above 0.25 high concentration.[14] Our

---

[12] https://www.bls.gov/data/ Series PPI industry group data for Air transportation support activities, not seasonally adjusted

[13] $HHI = \sum_{h=1}^{k} s_h^2$ where s is the share of firm h in the market and k is the number of firms.

[14] https://www.justice.gov/atr/horizontal-merger-guidelines-08192010 (Accessed February 27, 2019)

definition of a market as non-stop flights between two airport pairs is relatively narrow, and this is reflected in the average HHI score of 0.76 over the period of analysis. This value is in line with the 0.79 HHI reported by Dai et al. (2014) and the 0.72 to 0.78 noted by Gerardi and Shapiro (2009). To facilitate comparison of the HHI value and the number of carriers we follow Gerardi and Shapiro (2009) and use the negative of the log of the HHI, noted in our estimation as the instrumented variable $-ln\ \widehat{HERF}$.

1.4.4. Ticket Data

We analyze domestic US, coach-class tickets over the fifteen-year period of the first quarter 2002 to the fourth quarter 2016. Ticket price data comes from the Airline Origin and Destination Survey (DB1B), which is a 10% sample from all reporting carriers collected quarterly by the BTS. The data provided by carriers contains itinerary specific information, such as the number of coupons, the ticketing carrier, operating carrier, origin, destination and connecting airports, fare, miles flown and service class. This data is reported by all U.S. carriers that have at least 1% of the total scheduled-service domestic passenger revenue. This group has seen changes over the period of observation, ranging from ten to eighteen firms as new carriers have entered the market and existing carriers have merged or gone under.

Following previous airline literature (see Borenstein and Rose, 1994; Gerardi and Shapiro, 2009 and Dai et al., 2014) we include only non-stop, coach class itineraries for flights within the US. We include both one-way and round-trip tickets, but define ticket price as a one-way fare, thus round-trip tickets are included as half of the full fare. Itineraries less than \$10 one-way or \$20 round trip are excluded to eliminate frequent flyer, promotional tickets or non-revenue passengers. We also exclude fares that the BTS has flagged as questionable.

Following previous work, we are utilizing the Gini coefficient as our primary measure of dispersion $S(\cdot)$ for estimations (7 – 9). The Gini coefficient is a value between 0 and 1 and is equal to twice the expected absolute difference as a proportion of mean price between any two randomly drawn ticket prices. For example, a Gini coefficient of 0.20 would indicate an expected absolute price difference of 40% of the mean fare for any two randomly drawn tickets. Our calculation for the Gini coefficient of fares follows the formula established by Borenstein and Rose (1994) and replicated by Gerardi and Shapiro (2009) [15]. As the Gini coefficient is bounded between zero and one we follow Gerardi and Shapiro (2009) and use the log-odds ratio given by $S(\cdot) = \ln[\mathrm{Gini}/(1-\mathrm{Gini})]$, which provides an unbounded statistic. To further understand how competition affects price and dispersion we also look at the 10th percentile price (P10), the 90th percentile price (P90) and the spread between these two prices (P90-P10).

The fifteen-year period between 2002 and 2016 provides 73,981 separate quarterly route observations with 20 individual carriers and 2,079 routes represented. At 0.23, the overall Gini measure is in line with previous studies, Borenstein and Rose (1994) using Q2 1986 data calculated a Gini of 0.18, Gerardi and Shapiro (2009) calculated a Gini value of 0.22 for the 1993 – 2006 period and Dai et al. (2014) found 0.23 for the 1993 – 2008 period.

1.4.5. Differentiation

As defined in (4), the output attribute index $Q(w_h, y_h, q_h, q^0)$ measures the impact on the unit cost of a change between output attribute level q_h and q^0. However, for the purpose of estimations (7 - 9) we need a measure at the level of the market to indicate the degree of differentiation by output

---

[15] $GINI = 1 - 2\ x\ \left(\sum_{i=1,N} fare_i\ x\ \frac{PAX_i}{total\ Revenues}\right)\ x\ \left[\frac{PAX_i}{total\ PAX}\ x\ \left(1 - \sum_{j=1,i} \frac{PAX_j}{total\ PAX}\right)\right]$ where N is the number of different fare level tickets reported by a carrier on specific route, $fare_i$ is the fare for the ith ticket and PAX_i is the number of passengers traveling at that fare.

attribute in the market. Forsund and Hjalmarsson (1979) demonstrated that it was possible to create a firm representative of an industry based on the arithmetic average of all firm inputs, outputs and output attributes q_h. The value for q^0 is already at the market level and does not need to be averaged. An approach similar to this was used by Yu, Chen, and Hsiao (2018) in their study of ferry transportation. Following these examples, we create a representative carrier for each route and quarter and calculate a market output attribute index defined as $Q(w, y, q, q^0)$ where w, y and q are the average of all carriers in the market. The interpretation of $Q(w, y, q, q^0)$ at the market level is the same as for the case of firm h provided after expression (4). Over the period observed, the average value for $Q(w, y, q, q^0)$ was 0.95 with a standard deviation of 0.09. For the purpose of our estimation, we want a value that increases with differentiation to match the value of $-ln\ \widehat{HERF}$ which increases with competition so we use the inverse value noted as $Q(w, y, q, q^0)^{-1}$.

1.5. Results

1.5.1. Descriptive Results

Results for observed and adjusted fare and price dispersion values are presented in table III. To provide detail and context we have separated the results by degree of market concentration. Following Borenstein and Rose (1994) we are grouping together markets where one carrier has a 90% or greater share as *Monopoly,* markets where the top two carriers combined have a 90% or greater share as *Duopoly* and then all other markets as *Competitive*. These groupings are not intended to infer specific market characteristics, but simply to group together markets of similar concentration. Results are further categorized into five levels of market differentiation using the value of $Q(w, y, q, q^0)^{-1}$. The first grouping contains markets we would consider homogenous. To allow for some slack in measurement, and in consumer's ability to discern differences, we define a market as homogeneous when minimum unit cost at attribute level q and q^0 are equal or no more than 1% different[16]. In other words, the value of $Q(w, y, q, q^0)^{-1}$ is between 1.00 and 1.01. The next four groupings would be considered heterogeneous, those with values greater than 1.01 but less than 1.05, then values between 1.05 and 1.10, values between 1.10 and 1.15 and finally those observations with a value greater than 1.15. The higher the value of $Q(w, y, q, q^0)^{-1}$ the greater the level of market heterogeneity.

(*Table III near here)*

We observe a sharp difference in per mile fares between the three market structures. At an average of $0.34 per mile, fares in the *Monopoly* group are 35% greater than the average of $0.25 in the *Duopoly* group, which themselves are 32% greater than the $0.19 in the *Competitive* group. Fares that have been adjusted to remove the costs of a higher level of attributes are lower but retain a similar difference between concentration levels. We can also observe that generally per mile fares increase as market heterogeneity increases.

The level of price dispersion, given by the average Gini coefficient in the third column, also varies by level of market concentration and differentiation, but the interpretation is more complex. The table shows the complex relationship between market concentration and level of differentiation, which is one of the objects of this study that the econometrical analysis of the next section will help to disentangle. The values for the adjusted Gini coefficient show an overall increase which highlights a shortcoming of the adjusted Gini as a measure of dispersion for our use. As a measure of dispersion, the Gini measures dispersion across the whole range of fares and

---

[16] For context, a 1% difference in unit costs adds roughly $1.00 to the total cost per passenger on a flight of average length.

can be sensitive to changes in the middle. There are often common fare points in a market that are shared by carriers, for example multiple carriers may offer a $145 advance-purchase fare. Our method of adjustment may move one or both carriers off that point, creating more dispersion. To better understand the data, we have added a second measure of dispersion that is more focused on the tails of the dispersion, the 10th (P10) and 90th (P90) percentile price per mile and the spread between them. We define the spread (P90-P10) as simply the 90th percentile price per mile less the 10th percentile price per mile. An increase in the spread value indicates a wider dispersion.

Our first research question laid out in section III compares the levels of dispersion in markets with a wide and narrow range of attribute. To explore this, we perform a comparison of means, the results for this are outlined in Table IV. Using the definition of *homogeneous* as a value of $Q(w, y, q, q^0)^{-1}$ between 1 and 1.01 and *heterogeneous* as values greater than 1.01 we calculate weighted means of dispersion on the Gini coefficient and the (P90-P10) spread. Our test establishes the null hypothesis that mean dispersion in a homogeneous market is greater than or equal to the mean in a heterogeneous market. We reject the null hypothesis and accept the alternative that dispersion is greater in heterogeneous markets in every case except the Gini measure in the *Monopoly* structure[17].

(Table IV near here)

---

[17] One possible explanation for this is that in the more highly concentrated Monopoly markets carriers providing a higher level of output attributes do not offer price reductions and engage in limited price discrimination. This explanation is supported by table IV where that same group has a per mile fare significantly higher than any other group. This also may explain the 0.410 measure for P90-P10 spread in this group, many higher price tickets, and only a few lower priced tickets.

We also looked at the dispersion on prices that have been adjusted in concordance with (6) by removing the cost of providing a higher level of attribute and find mixed results. In terms of the Gini measure we see an increase in dispersion, but the (P90-P10) spread shows a reduction in dispersion. These results may be explained by the issue with “common” fare points noted above[18].

1.5.2. Fixed Effects Panel Estimation Results

Table V contains estimation results using the HHI as the measure of competition and the Gini log odds ratio as the measure of dispersion, while VI presents the same measure of competition but with the (P90-P10) as the measure of dispersion. In tests of instrument validity, we found all instruments were significant, and with an F test value of 6962 we can reject that they are weak instruments. Both tables report results for estimation formulas (7 - 9). In the table V baseline estimation (7) we see that effect of competition on price dispersion, measured as $-ln\,\widehat{HERF}$, is negative and significant. This echoes the findings of Gerardi and Shapiro (2009). These findings are reinforced in table VI where we see that an increase in competition reduces the P90 price by more than the P10 price, resulting in a narrower spread and a reduction in dispersion. We also see in the second column of table V that the effect of competition on price dispersion, after removing the cost of product differentiation, is stronger than the coefficient on observed price. This would be expected as adjusted prices reflect a market that is more homogenous in terms of output attributes. For brevities sake we do not report adjusted values in table VI but results mirror those of table V.[19] Results for estimation (8), which bridges (7) and (9) are interesting, but not very

---

[18] On prices that have been adjusted the Gini value for all market structures together is 0.248 for Homogeneous markets and 0.246 for Heterogeneous. For the P90-P10 spreads these values are 0.259 and 0.304 showing contradictory results.

[19] Like table VI the reduction in dispersion is greater on the adjusted prices. The P90-P10 spread is -0.511 on observed prices and -0.546 on adjusted prices.

informative. We see that the significantly negative effect of competition on price dispersion is strengthened over (7) once the level of differentiation is accounted for. Although the negative coefficient on $Q(w,y,q,q^0)^{-1}$ was unexpected, in fact our model did not make predictions on what happens within a market when the level of differentiation changes.

(Tables V and VI near here)

Turning to formula (9) we now address the question of the interaction effect of competition on price dispersion in homogenous and heterogeneous markets. As can be seen in both tables V and VI, the interaction effect is significant and positive. To verify the validity of the interactive term we performed a Wald test on the results of estimations (8) and (9) and find that in addition to being significant, it adds to the model. With an F value of 52.21 we reject the null hypothesis that model (8) is as good as (9). A similar test was run based on both measures of competition and both measures of dispersion, all showed the same validity of the interactive measure. Understanding the question of interaction requires a more detailed level of information that is provided by the rewriting of expression (9), and the specification of different levels of market differentiation in concordance with Table III. To simplify interpretation, we provide table VII as the resulting coefficient on competition at various levels of $Q(w,y,q,q^0)^{-1}$. This can be understood as reducing formula (9) to the formula below, and then finding $(\beta_1+\beta_3 Q)$ at various levels of Q,

$$S_{jt}(\cdot) = (\beta_1+\beta_3 Q_{jt})\, Com_{jt} + \beta_2\, Q_{jt} + \theta\chi_{jt} + \delta_{ht} + \gamma_t + \upsilon_j + \varepsilon_{jt}. \tag{10}$$

Results in Table VII clearly show that considering interactive effects, the effect of competition on dispersion is negative in homogenous markets, but then turns positive as the market becomes more

---

differentiated in terms of output attributes. This can be seen in the Gini measure or in the relative changes of the P10 and P90 coefficient and the resulting (P90-P10) spread. In both cases, the effect of competition on dispersion changes from negative to positive at a $Q(w, y, q, q^0)^{-1}$ around a value of 1.10. Referring to our adaption of Borenstein (1985) model developed in section 1.3, dispersion reduction in homogeneous markets is driven from monopoly-type discrimination as firms compete based on market elasticity. While in the more heterogeneous markets the increase in dispersion arises from competitive-type discrimination and more cross-price elasticity competition.

(Table VII near here)

Estimations from tables V, VI and VII are reproduced in tables VIII, IX and X with the log of the number of carriers in the market ($ln\widehat{N}$) as the measure of competition. Under this measure we see in table VIII with the Gini as the measure of dispersion that the effect of competition has almost no effect on price dispersion and the coefficient is not significant. In fact, it is not until we add the interactive effects of formula (9) and allow for different effects based on the level of differentiation, that the coefficient becomes significant. Results from table IX, using the P10 and P90 to measure dispersion, closely mirror those of table VI in terms of sign, coefficient value and significance. In table IX the effect of competition is significant and negative even before controlling for differentiation. Finally, table X reports the effective coefficient based on interactive effects at various levels of $Q(w, y, q, q^0)^{-1}$. Under this measure of competition, we see the effect of competition on the Gini coefficient and the (P90-P10) changes sign when $Q(w, y, q, q^0)^{-1}$ values are greater than 1.15.

(Tables VIII, IX and X near here)

1.5.3. Robustness Checks

Because one of our primary contribution is the generation of output attribute index $Q(w,y,q,q^0)$ we first test for robustness in this value. As discussed in section IV, the value for output attributes of on-time arrival and load factor are measured at the level of the carrier, whereas the value for flight frequency is measured at the route level. An alternative would be to measure all output attributes at the route level. The logic being that although the information is less accessible at the route level, it is the route level output attribute that directly affects consumers. To test for robustness, we replicate the method with this alternative measure and find that results are essentially unchanged, though not as statistically strong. This loss in statistical significance is largely because a significant number of routes are served by a single carrier. When differentiation by output attribute is measured at the route level, these single-carrier routes show no differentiation, masking the true level of heterogeneity in service levels.

In addition to testing robustness of the output attribute index, we check for robustness in our regression results from formula (9). Our presented results are based on a static fixed effects panel within estimator. As there is a fair amount of persistence in the measure for dispersion in any given market, we first test our results with the addition of a lag of the dependent variable. As might be expected, the lagged variable is highly significant both statistically and economically. However, our main results for competition, differentiation and the interaction are robust to the addition of this lag.

We test further and utilize dynamic panel data methods to account for the possibility of correlation between the transformed lagged variable and transformed error term in the within model. We test several specifications, including the Arellano and Bond (1991) generalized-method-of-moments (GMM) estimator, the Blundell and Bond (1998) system GMM estimator and the Arellano and Bover (1995) system GMM with forward orthogonal deviations from Roodman (2003). This last specification, which differences observations by subtracting the mean of future observations, fits our data well. Because airlines can enter or exit a market easily, or switch between direct and connecting service, our panel has many gaps. The forward orthogonal deviation method retains observations that would have been lost through first differences. Though the absolute value of coefficients vary somewhat between the methods, again our main results for competition, differentiation and the interaction hold up.

As a further check for robustness we test for the existence of indirect effects of competition on dispersion through our measure of heterogeneity Q. A concern would be that heterogeneity has no effect on its own, but only through mediation of competition. As a first step in this check we run a simple regression with competition predicting heterogeneity Q. We find that the relationship is statistically significant at the 5% level but has a very small effect with a coefficient of -0.005 on competition. If we replicate this with instrumented competition the effect is stronger at -0.015. These results might indicate that there are indirect effects of competition on dispersion through Q. A simple way to test for these results is suggested by Judd and Kenny (1981). They suggest comparing the result coefficient between a model that includes the mediator and one that does not, subtracting one from the other. We have this already as equations (7) and (8). Using the results from Table V we find that the indirect effect of competition on dispersion, through Q, is 0.031 compared to the direct effect of -0.222.

In the introductory part of section 1.3, we noted the possibility of using a control for LCC as a method of measuring market heterogeneity. As a robustness check, we ran estimations (8) and (9) replacing Q with a dummy for the presence of a mix of carrier types in the market. In neither estimation was the dummy significant, nor was the model fit was not as good as the base model.

As a further test we included a control for the presence of an LCC in the market. The coefficient on the control was negative and significant, but the there was no change to the other variables.

1.5.4. Reconciliation with Previous Studies

In this section we reconcile our results with those of previous studies and explore how our findings can help explain some of the contradictions between those studies. Like previous studies, we perform a panel analysis of the effect of competition on price dispersion using fixed-effects estimation to control for time-invariant market specific factors. However, we add a measure of market differentiation and its interaction with competition, showing that the effect of competition on price dispersion differs depending on whether the market is homogeneous or heterogeneous based on output attribute.

Our results from estimation formula (7) are in line with those of those of Gerardi and Shapiro (2009), the difference is our extension of the model to include the level of market heterogeneity in formula (8) and the interaction effect in formula (9). This extension refines their work, finding that the effect of competition can differ based on the degree of differentiation. However, in a sample where markets are predominantly more homogeneous the result would indicate a negative relationship, just as Gerardi and Shapiro (2009) found.

Dai et al. (2014) find an inverse-U-shaped relationship, with price dispersion increasing as a highly concentrated market initially becomes more competitive but decreasing when a less concentrated market becomes even more competitive. They do not control for the level of differentiation. What we find does not contradict this but can provide further insight to what is happening in the market. When a monopoly market first receives a second carrier there is a move from highly concentrated to more competitive, concurrent with that change there is a probability of an increase in differentiation and therefore an increase in dispersion. For markets that are already less concentrated, differentiation is more likely to have peaked and competition increases are likely to come in the form of battles between carriers serving the same niche. We see this effect in the descriptive data of table III in the column Avg Gini-Observed. In the competitive structure, as markets become more differentiated dispersion increases from 0.250 to 0.277, but then begins decreasing, dropping to 0.255 at the highest levels of heterogeneity. A similar effect is seen in the duopoly structure.

Reconciling our findings with Borenstein and Rose (1994) we look to two particular points, the difference in econometric methods and the lack of a variable controlling for differentiation. Gerardi and Shapiro (2009) provide a full discussion on the differences between the cross section method used by Borenstein and Rose (1994) and the fixed-effects panel method, but in short, they show how the cross section method would bias the coefficient on competition. At the end of their discussion of a few possible causes, they note that it is possible that other time-invariant factors are biasing the cross-sectional estimates. We would argue that the level of differentiation in the market is one of the most important of these. As a test of this we run a cross section estimation on our data set, and like Borenstein and Rose (1994) and Gerardi and Shapiro (2009) find a positive relationship between competition and dispersion. However, once we add $Q(w, y, q, q^0)^{-1}$ to the estimation, either on its own or as an interaction variable, the sign on competition turns negative.

1.6. Conclusions

In this study we have introduced a Konüs type output attribute index which collects the impact on cost change between various levels of output attributes. This new method allows us to measure the amount of cost-driven differentiation by output attributes in a given market. Markets with wide difference in output attribute and associated cost difference are categorized as heterogeneous markets, whereas those with a narrow range of output attributes and cost difference are classified as homogeneous markets.

Returning to our research questions we can now draw conclusions. First, price dispersion is greater in heterogeneous markets. This is true both for observed price dispersion and for price dispersion after removing the effect of cost differences due to a higher level of output attributes. Our second question explores how competition effects dispersion after accounting for output attribute cost-driven dispersion. We find that after removing the premium for a higher level of attributes the negative effect of competition on price dispersion is greater. Further, if we control for the level of output attribute differentiation, we find that competition has a strictly negative effect on price dispersion.

Finally, using a fixed-effects panel estimation with interactions we find that the direction and magnitude of the effect of competition on price dispersion is dependent on the level of output attribute driven market heterogeneity. This can be seen in the positive coefficient on the interaction effect between competition and market heterogeneity on price dispersion. We also find that at higher levels of differentiation an increase in competition can increase price dispersion. Whereas homogeneity is an absolute, heterogeneity is gradable and has levels. We see that the switch from price dispersion decreasing to increasing as competition grows, only occurs at a certain level of differentiation. This might imply that although these results can be generalized to other industries, the point at which this affect occurs may be different. We also reconcile our findings with previous research and show how our research fits with them and helps to explain contradictory findings.

Our findings provide regulators additional information to consider when reviewing potential mergers. Typically, regulators focus on the HHI and how the merger will affect market concentration. However, as we have shown, the market HHI alone is not enough when trying to understand changes related to pricing. Regulators should also consider the level of differentiation provided by the firms and how a merger might affect differentiation in the markets they are part of. A merger or horizontal alliance between two carriers can result in a greater increase in fare price, and price dispersion than a merger between two carriers with different attributes. A further consideration for policy makers is that concentration alone does not determine the level of competition. A market with two carriers of similar attributes will be more competitive than one where the carriers are quite different since Market power is not determined by concentration alone. This is clearly seen in Table IV where the highest per mile fares are found in the more heterogeneous markets, regardless of the market structure.

From a strategic perspective, these results can provide carriers some direction when choosing routes to enter or exit. Entering a route where the current competitors are more similar would be less attractive, unless the carrier can adjust their product offering. On the other side, entering a route where the incumbents are dissimilar will allow the carrier greater latitude to price discriminate and generate higher fares. In Table III we also see that higher fare prices can be earned even after adjusting for the cost of differentiation.

Our study has focused on the effects on price dispersion. Further research utilizing these techniques could analyze the effect on average prices and price deciles to further understand how changes in competition affects markets, conditioned on the level of differentiation. This method could also be applied to other economic models to understand how explicitly controlling for the level of differentiation by output attribute affects model outcomes. Some possible examples may be productivity analysis, the hedonic price equation that is explored in the next chapter or other models where market power plays a significant role.

Appendix 1.1. – Alternative Cost Estimation Method

The cost function and Q in (4) can be estimated through one of two methods, parametric or non-parametric. In this article we have used the non-parametric DEA method, however in this appendix we explore a parametric method of estimating (4). Unlike non-parametric models, which estimate the technology using the minimum extrapolation method, parametric models are defined a priori and require assumptions on the functional form of production. From these functions a finite set of parameters are estimated from the data. In relation to airline cost functions, previous literature has seen the application of both deterministic parametric frontiers (Gillen, Oum, and Tretheway, 1990 and Johnston and Ozment, 2013) and stochastic parametric frontiers (Ahn and Sickles, 2000 and Assaf, 2009).

The functional forms most commonly used in airline cost functions is the translog, which was proposed by Christensen, Jorgenson, and Lau (1973) and has been widely used in airline applications. With *C* as the total cost per firm at time *t*, *w* as the price of inputs, *y* as output and *q* as output attributes and, additionally, including a dummy variable to account for the panel nature of our data by adding a time trend variable (*t*), the translog cost function to be estimated is

$$\begin{aligned}\ln C_{ht} = {} & \sum_{i=1}^{n} \alpha_i \ln w_{iht} + {}^{1}\!/_{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} \ln w_{iht} \ln w_{jht} \\ & + \sum_{i=1}^{l} \alpha_{iy} \ln y_{iht} + {}^{1}\!/_{2} \sum_{i=1}^{l} \sum_{j=1}^{l} \alpha_{ijy} \ln y_{iht} \ln y_{jht} \\ & + \sum_{i=1}^{n} \sum_{j=1}^{l} \alpha_{ijwy} \ln w_{iht} \ln y_{jht} + \sum_{k=1}^{m} \alpha_{kq} \ln q_{kht} \\ & + \alpha_t t + \varepsilon_{ht}.\end{aligned}$$

Given the log nature of the translog cost function, the value $Q(w_h, y_h, q_h, q^0)$ in (4) can be directly calculated after parameters $\widehat{\alpha} = \{\hat{\alpha}_i, \hat{\alpha}_{ij}, \hat{\alpha}_{iy}, \hat{\alpha}_{ijy}, \hat{\alpha}_{ijwy}, \hat{\alpha}_{kq}, \hat{\alpha}_t\}$ are estimated. Recalling (4) as $Q(w_h, y_h, q_h, q^0) = c(w_h, y_h, q^0)/c(w_h, y_h, q_h)$, we have,

$$Q(w_h, y_h, q_h, q^0) = \exp[\ln c(w_h, y_h, q^0) - \ln c(w_h, y_h, q_h)]$$

$$= \exp\left[\sum_{k=1}^{m} \hat{\alpha}_{qk} \ln q_{kt}^0 - \sum_{k=1}^{m} \hat{\alpha}_{qk} \ln q_{kht}\right].$$

Appendix 1.2. – Table of Variables and Computations

Variable	**Definition**
Output Quantity *y*	Output quantities: Revenue Passenger Miles
Output price *p*	Output prices: All fare prices from the 10% sample DB1B
Input Quantity *x*	Input Quantities: Labor (Full-Time Equivalents), Gallons of Fuel, Capital (measured as quantity of aircraft seats), All Other Materials (quantity measured as cost divided by a price index)
Input Price *w*	Input Prices: Total expenditures for the input divided by quantity *x*
Output Attributes *q*	Output Attributes: Load Factor (passenger seat miles divided by available seat miles), Flight Frequency (average daily flights on the route), On-Time Arrival (percentage of flights that arrive within 15 minutes of scheduled arrival time)
Heterogeneity Measure: $Q(w_h, y_h, q_h, q^0)$ for firm *h*	$\frac{ac(w_h, y_h, q^0)}{ac(w_h, y_h, q_h\ )};\ \ h\ = 1,..,k$
Minimum attributes q in market *j*	$q^0 = \{q_1^0, \dots, q_m^0\}$ and $q_j^0\ = Min\ \{q_{jh}^0, h = 1,..,k\}, j = 1,..,m.$

Variable	**Definition**
Herfindahl-Hirschman index (HHI)	$HHI = \sum_{h=1}^{k} s_h^2$ where s is the share of carrier h in the market and k is the number of carriers
Gini Coefficient	$1 - 2\ x\ \left(\sum_{i=1,N} fare_i\ x\ \frac{PAX_i}{total\ Revenues}\right)\ x\ \left[\frac{PAX_i}{total\ PAX}\ x\ \left(1 - \sum_{j=1,i} \frac{PAX_j}{total\ PAX}\right)\right]$ where N is the number of different fare level tickets reported by a carrier on specific route, $fare_i$ is the fare for the ith ticket and PAX_i is the the number of passengers traveling at that fare
lgini (variable to measure price dispersion)	$\ln[\mathrm{Gini}/(1 - \mathrm{Gini})]$
P90-P10 (variable to measure price dispersion)	90th percentile fare price – 10th percentile fare price
lnherf (variable to measure concentration)	- ln (HHI)

Figures, Tables and Graphs

Table I

Statistics of Output, Output Attributes and Routes – Average Route

Carrier	Passenger Miles y (millions)	On-Time Arrival $q1$ (%)	Flight Frequency $q2$ (#)	Load Factor $q3$ (%)	Routes (#)
AirTran Airways Co.	31	79	2.7	24	119
Alaska Airlines Inc.	56	83	2.5	21	87
America West Airlines Inc.	65	81	3.2	23	83
American Airlines Inc.	105	79	4.3	19	208
ATA Airlines	64	78	2.1	26	35
Continental Air Lines Inc.	86	79	3.3	19	120
Delta Air Lines Inc.	79	82	3.6	18	260
Envoy Air	10	78	3.8	29	169
ExpressJet Airlines Inc.	10	78	2.6	25	209
Frontier Airlines Inc.	35	78	2.2	15	68
Hawaiian Airlines Inc.	90	93	4.0	13	23
JetBlue Airways	65	77	2.9	17	97
Mesa Airlines Inc.	7	81	1.8	21	154
Northwest Airlines Inc.	52	79	3.0	20	197
PSA Airlines Inc.	4	70	1.8	30	92
SkyWest Airlines Inc.	9	82	2.4	21	336
Southwest Airlines Co.	43	82	3.5	24	455
Spirit Air Lines	31	73	1.2	15	147
United Air Lines Inc.	95	80	3.4	17	197
US Airways Inc.	61	81	3.5	20	160
Virgin America	95	81	3.1	18	27

Table II

Statistics of Input Quantity and Input Prices (Route Average)

Carrier	Fuel ($x1$)		Labor ($x2$)		Flight Capital ($x3$)		Other Material ($x4$)	
	Gal (000's)	Price ($)	FTE (#)	Price ($)	Seats (#)	Price ($)	Quantity	Price ($)
AirTran Airways Co.	661	2.00	62	15.69	29	17.51	9	141.51
Alaska Airlines Inc.	1,037	2.02	113	22.96	47	14.51	22	143.73
America West Airlines Inc.	1,222	1.05	133	14.32	56	18.15	24	123.41
American Airlines Inc.	2,162	1.88	254	21.49	96	11.64	49	143.73
ATA Airlines	1,160	1.43	138	14.29	55	22.64	28	131.10
Continental Air Lines Inc.	1,566	1.68	203	20.10	72	16.65	60	135.55
Delta Air Lines Inc.	1,496	2.00	175	24.02	69	11.94	52	143.73
Envoy Air	352	1.85	55	12.31	17	13.80	6	137.50
ExpressJet Airlines Inc.	222	1.46	28	16.45	14	17.95	4	141.12
Frontier Airlines Inc.	770	2.38	76	16.80	30	21.69	13	153.56
Hawaiian Airlines Inc.	1,424	2.13	134	23.22	73	19.04	29	151.11
JetBlue Airways	1,236	2.19	112	20.89	50	13.03	17	149.31
Mesa Airlines Inc.	155	2.64	16	10.99	10	26.82	3	150.74
Northwest Airlines Inc.	1,242	1.52	131	23.21	58	9.89	31	131.46
SkyWest Airlines Inc.	126	2.50	28	14.89	11	23.93	2	149.31
Southwest Airlines Co.	834	1.79	85	24.96	40	10.34	11	143.73
Spirit Air Lines	431	1.66	30	22.22	21	21.15	6	167.44
Trans World Airways LLC	1,461	0.90	198	15.98	65	23.17	26	114.18
United Air Lines Inc.	1,720	1.90	237	20.65	77	11.89	63	143.73
US Airways Inc.	1,196	1.92	156	18.83	60	15.59	43	140.57
Virgin America	1,553	2.23	94	27.05	66	28.23	28	164.01

Table III

Statistics by Market Structure and Output Attribute Index (Average Weighted by Number of Passengers)

Market Structure	Output Attribute Index Range	Output Attribute Index Avg	Avg Fare per Mile ($) Observed	Avg Fare per Mile ($) Adjusted	Avg Gini Observed	Avg Gini Adjusted	Count of Markets
Monopoly	1.00 - 1.01	1.00	0.29	0.28	0.249	0.249	18,794
	1.01 - 1.05	1.02	0.37	0.36	0.246	0.246	15,747
	1.05 - 1.10	1.07	0.42	0.39	0.234	0.234	7,314
	1.10 - 1.15	1.12	0.42	0.37	0.225	0.225	3,407
	1.15 +	1.38	0.40	0.30	0.189	0.189	8,526
	Average	1.07	0.34	0.32	0.238	0.238	-
	Total	-	-	-	-	-	53,788
Duopoly	1.00 - 1.01	1.00	0.24	0.24	0.245	0.246	5,659
	1.01 - 1.05	1.03	0.23	0.22	0.254	0.257	4,672
	1.05 - 1.10	1.07	0.26	0.24	0.253	0.257	2,138
	1.10 - 1.15	1.12	0.30	0.26	0.245	0.253	916
	1.15 +	1.27	0.35	0.26	0.224	0.239	1,423
	Average	1.05	0.25	0.24	0.247	0.250	-
	Total	-	-	-	-	-	14,808
Competitive	1.00 - 1.01	1.00	0.17	0.17	0.250	0.250	2,698
	1.01 - 1.05	1.03	0.17	0.16	0.262	0.265	1,576
	1.05 - 1.10	1.07	0.22	0.20	0.277	0.281	610
	1.10 - 1.15	1.12	0.28	0.24	0.274	0.281	217
	1.15 +	1.30	0.43	0.32	0.255	0.269	284
	Average	1.04	0.19	0.18	0.257	0.260	-
	Total	-	-	-	-	-	5,385
All Market Structures	1.00 - 1.01	1.00	0.24	0.24	0.248	0.248	27,151
	1.01 - 1.05	1.03	0.27	0.26	0.253	0.254	21,995
	1.05 - 1.10	1.07	0.31	0.29	0.251	0.254	10,062
	1.10 - 1.15	1.12	0.35	0.30	0.242	0.247	4,540
	1.15 +	1.34	0.39	0.29	0.209	0.215	10,233
Overall	Average	1.060	0.28	0.26	0.245	0.247	-
	Total	-	-	-	-	-	73,981

Table IV

Independent Samples Weighted Unequal Variances T-test

H_0: $\mu_{het} \leq \mu_{hom}$

H_1: $\mu_{het} > \mu_{hom}$

	Mean Observed Gini Coefficient			Mean Observed P90-P10 Spread		
	Monopoly	Duopoly	Competitive	Monopoly	Duopoly	Competitive
Homogenous	0.249	0.245	0.250	0.316	0.260	0.182
Heterogenous	0.229	0.248	0.266	0.410	0.286	0.250
P-Value	1.00	0.00	0.00	0.00	0.00	0.00
Reject H_0	No	Yes	Yes	Yes	Yes	Yes

Table V

Panel Estimates

Dep Var: Gini Log Odds Ratio (Observed or Adjusted Prices)

Estimation Formula:	Observed (7)	Adjusted (7)	Observed (8)	Observed (9)
$-\ln \widehat{HERF}$	-0.164**	-0.253***	-0.222***	-2.380***
	(0.068)	(0.068)	(0.069)	(0.338)
$Q(w,\ y,\ q,\ q^0)^{-1}$			-0.208***	-0.286***
			(0.009)	(0.015)
Interaction Variables				
$-\ln \widehat{HERF} \cdot Q(w,\ y,\ q,\ q^0)^{-1}$				2.234***
				(0.067)
Observations	73981	73981	73981	73981

Notes: All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table VI

Panel Estimates

Dep Var: Log of 10th or 90th Percentile Observed Price

	Log(P10)	Log(P90)	Log(P10)	Log(P90)	Log(P10)	Log(P90)
Estimation Formula:	(7)	(7)	(8)	(8)	(9)	(9)
$-\ln \widehat{HERF}$	-1.044***	-1.555***	-1.076***	-1.603***	-5.772***	-11.155***
	(0.060)	(0.059)	(0.061)	(0.061)	(0.064)	(0.103)
$Q(w, y, q, q^0)^{-1}$			-0.096***	-0.224***	-0.257***	-0.562***
			(0.009)	(0.009)	(0.018)	(0.022)
Interaction Variables						
$-\ln \widehat{HERF} \cdot Q(w, y, q, q^0)^{-1}$					5.172***	10.102***
					(0.366)	(0.443)
Observations	73,981	73,981	73,981	73,981	73,981	73,981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table VII

Coefficient ($\beta_1 + \beta_3 Q$) on *Com* ($-\ln \widehat{HERF}$) at Different Output Attribute Index Levels

Dependent Variable: Gini log odds ratio, 10th Percentile Price or 90th Percentile Price

Output Attribute Index		$-\ln \widehat{HERF}$	$-\ln \widehat{HERF}$		
Range	Avg	Gini	P10	P90	(P90-P10)
1.00 - 1.01	1.00	-0.146	-0.600	-1.053	-0.453
1.01 - 1.05	1.03	-0.079	-0.445	-0.750	-0.305
1.05 - 1.10	1.07	0.010	-0.238	-0.346	-0.108
1.10 - 1.15	1.12	0.122	0.021	0.159	0.139
1.15 +	1.34	0.614	1.158	2.382	1.223

Table VIII

Panel Estimates

Dep Var: Gini Log Odds Ratio (Observed or Adjusted Prices)

Estimation Formula:	Observed (7)	Adjusted (7)	Observed (8)	Observed (9)
$\ln \widehat{N}$	0.007	-0.072	-0.020	-1.041***
	(0.074)	(0.074)	(0.073)	(0.186)
$Q(w, y, q, q^0)^{-1}$			-0.194***	-0.260***
			(0.008)	(0.013)
Interaction Variables				
$\ln \widehat{N} \cdot Q(w, y, q, q^0)^{-1}$				0.893***
				(0.144)
Observations	73,981	73,981	73,981	73,981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table IX

Panel Estimates

Dep Var: Log of 10th or 90th Percentile Observed Price

Estimation Formula:	Log(P10) (7)	Log(P90) (7)	Log(P10) (8)	Log(P90) (8)	Log(P10) (9)	Log(P90) (9)
$\ln \widehat{N}$	-0.985***	-1.269***	-0.991***	-1.293***	-2.959***	-5.241***
	(0.056)	(0.051)	(0.056)	(0.051)	(0.185)	(0.187)
$Q(w, y, q, q^0)^{-1}$			-0.029***	-0.124***	-0.187***	-0.419***
			(0.008)	(0.007)	(0.013)	(0.013)
Interaction Variables						
$\ln \widehat{N} \cdot Q(w, y, q, q^0)^{-1}$					2.137***	3.983***
					(0.144)	(0.146)
Observations	73,981	73,981	73,981	73,981	73,981	73,981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table X

Coefficient ($\beta_1 + \beta_3$ Q) on Com (ln $\hat{N}$) at Different Output Attribute Index Levels

Dependent Variable: Gini log odds ratio, 10th Percentile Price or 90th Percentile Price

Output Attribute Index		ln $\hat{N}$	ln $\hat{N}$		
Range	Avg	Gini	P10	P90	(P90-P10)
1.00 - 1.01	1.00	-0.148	-0.822	-1.258	-0.436
1.01 - 1.05	1.03	-0.121	-0.758	-1.139	-0.381
1.05 - 1.10	1.07	-0.085	-0.672	-0.979	-0.307
1.10 - 1.15	1.12	-0.041	-0.566	0.780	-0.214
1.15 +	1.34	0.156	-0.095	0.096	0.192

Chapter 1 Annex – Robustness Tests of Chapter 1

1.A1 Robustness Results

1.A1.1.Output Attribute Index

This annex has been prepared to provide detailed results of the robustness tests for the article of the same title. Its structure follows the layout of section 1.5.3. *Robustness Checks* and first discusses robustness in the measure $Q(w, y, q, q^0)$, then the introduction of a lag to the primary regression (9) and finally a dynamic panel version of the same regression. For ease of reference (9) is replicated below:

$$S_{jt}(\cdot) = \beta_1\, Com_{jt} + \beta_2\, Q_{jt} + \beta_3\, (Com \cdot Q)_{jt} + \theta\chi_{jt} + \delta_{ht} + \gamma_t + \upsilon_j + \varepsilon_{jt}. \qquad (9)$$

Output attribute index $Q(w, y, q, q^0)$ is a measure of the average cost premium of providing a higher level of output attributes on a route. It measures the degree of differentiation on a route and is constructed from the output attributes of the airlines present on the route. To recall, the three output attributes used in this article on-time arrival performance, flight frequency and load-factor. It is possible to measure these output attributes at the level of the route or the national level, for example Delta's on-time arrival performance on flights between Seattle-Salt Lake City (SEA-SLC) or Delta's national performance. There are arguments for both methods. The level of the route most directly effects consumers and may more accurately reflect resource allocation and cost on the route; however, route level information is not easily accessible by consumers. Performance at the national level is reported in media and is accessible by consumers, it also may better reflect the resources an airline has dedicated toward producing the output attribute system-wide.

A further argument for measuring output attribute at the national level is that on a route where only a single-carrier reports performance, measuring at the route level does not generate differentiation. A carrier may be generating a higher level of output attribute compared to potential entrants, recent exiting carriers or non-reporting carriers, but this differentiation would be lost at the route level if there were no carrier to compare against. For these reasons, our primary results are presented based on on-time arrival performance and load-factor measured at the national level and flight frequency at the route level. As a test of robustness for the measure $Q(w, y, q, q^0)$ we now replicate the primary results with all output attributes measured at the route level.

Tables XI, XII, XIII and XIV replicate the results from the main text tables V, VI, VIII, and IX respectively. The only difference is that now all the output attributes are measured at the route level. One apparent result is that the measure $Q(w, y, q, q^0)$ and its interaction is less significant, both economically and statistically then before. This is not surprising as all the single airline routes would now show no differentiation as there are no other carriers to compare to. Though the results are less significant, the main findings, negative coefficient on competition and positive coefficient on the interaction, still hold.

(Tables XI, XII, XIII and XIV near here)

Tables XV and XVI replicate the results from main text tables VII and X and present the results arising from the interaction effect at various levels of output attribute index. Again, the main results are similar, as differentiation based on output attribute grows, the negative effect of competition on price dispersion moves from negative to positive. However, the results again are not as strong as when differentiation is measured at the national level.

(Tables XV and XVI near here)

1.A1.2. Lagged Dependent Variable

Price dispersion in a market can be very persistent and it would not be unreasonable to argue that the level of dispersion in the prior period influences the current period. For this reason, we add a lag of the dependent variable to equation (9) and estimate the following,

$$S_{jt}(\cdot) = \beta_0 S_{jt-1} + \beta_1\, Com_{jt} + \beta_2\, Q_{jt} + \beta_3\, (Com \cdot Q)_{jt} + \theta\chi_{jt} + \delta_{ht} + \gamma_t + \upsilon_j + \varepsilon_{jt}. \qquad (11)$$

Results of estimation (11) are shown in table XVII with competition measured as $\ln \widehat{HERF}$ and in table XVIII with competition measured as $\ln \widehat{N}$. In all cases the lag of the dependent variable is highly significant both statistically and economically. The addition of the lag weakens the effect of our independent variables of interest, but in all cases except the Gini measure with $\ln \widehat{N}$ our main findings hold with significance. These results show that even after accounting for the previous periods level of dispersion our hypothesis holds.

(Tables XVII and XVIII near here)

1.A1.3. Dynamic Panel Data Estimator

Our final test for robustness is based on the lagged dependent variable in the previous section. Equation (11) is estimated as a fixed effects panel applying the within demeaning transformation. Because this transformation can create a correlation between the lagged dependent variable and the transformed error term our results may be biased. The dynamic panel estimator essentially creates instruments for the lagged dependent variable to avoid this bias. In table XIX we present estimation results for formula (11) utilizing three of the more common dynamic panel estimators, Difference GMM, System GMM and System GMM with Forward Orthogonal Deviations (FOD)

Figures, Tables and Graphs

Table XI
Panel Estimates - Output Attributes at Route Level
Dep Var: Gini Log Odds Ratio (Observed or Adjusted Prices)

Estimation Formula:	Observed (7)	Adjusted (7)	Observed (8)	Observed (9)
$-\ln \widehat{HERF}$	-0.164**	-0.169**	-0.142**	-0.382***
	(0.068)	(0.072)	(0.064)	(0.103)
$Q(w,\ y,\ q,\ q^0)^{-1}$			-0.007**	-0.045***
			(0.003)	(0.010)
Interaction Variables				
$-\ln \widehat{HERF} \cdot Q(w,\ y,\ q,\ q^0)^{-1}$				0.296***
				(0.067)
Observations	73981	73981	73981	73981

Notes: All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table XII

Panel Estimates - Output Attributes at Route Level

Dep Var: Log of 10th or 90th Percentile Observed Price

	Log(P10)	Log(P90)	Log(P10)	Log(P90)	Log(P10)	Log(P90)
Estimation Formula:	(7)	(7)	(8)	(8)	(9)	(9)
$-\ln \widehat{HERF}$	-1.044***	-1.555***	-1.006***	-1.463***	-1.687***	-2.671***
	(0.060)	(0.059)	(0.057)	(0.056)	(0.099)	(0.099)
$Q(w, y, q, q^0)^{-1}$			-0.016***	-0.036***	-0.143***	-0.246***
			(0.003)	(0.003)	(0.022)	(0.018)
Interaction Variables						
$-\ln \widehat{HERF} \cdot Q(w, y, q, q^0)^{-1}$					1.028***	1.689***
					(0.065)	(0.065)
Observations	73981	73981	73981	73981	73981	73981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table XIII
Panel Estimates - Output Attributes at Route Level
Dep Var: Gini Log Odds Ratio (Observed or Adjusted Prices)

Estimation Formula:	Observed (7)	Adjusted (7)	Observed (8)	Observed (9)
$\ln \widehat{N}$	0.007	0.145*	0.024	-0.063
	(0.074)	(0.078)	(0.074)	(0.059)
$Q(w, y, q, q^0)^{-1}$			-0.005*	-0.033***
			(0.003)	(0.011)
Interaction Variables				
$\ln \widehat{N} \cdot Q(w, y, q, q^0)^{-1}$				0.048**
				(0.019)
Observations	73981	73981	73981	73981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table XIV

Panel Estimates

Dep Var: Log of 10th or 90th Percentile Observed Price

Estimation Formula:	Log(P10) (7)	Log(P90) (7)	Log(P10) (8)	Log(P90) (8)	Log(P10) (9)	Log(P90) (9)
$\ln \widehat{N}$	-0.985***	-1.269***	-0.999***	-1.256***	-0.963***	-1.249***
	(0.056)	(0.051)	(0.057)	(0.052)	(0.049)	(0.044)
$Q(w, y, q, q^0)^{-1}$			0.005*	-0.005*	-0.139***	-0.231***
			(0.003)	(0.003)	(0.010)	(0.009)
Interaction Variables						
$\ln \widehat{N} \cdot Q(w, y, q, q^0)^{-1}$					0.246***	0.387***
					(0.017)	(0.015)
Observations	73,981	73,981	73,981	73,981	73,981	73,981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table XV

Coefficient ($\beta_1 + \beta_3 Q$) on $Com\ (-\ln \widehat{HERF})$ at Different Output-Attribute Index Levels
Dependent Variable: Gini log odds ratio, 10th Percentile Price or 90th Percentile Price

Output-Attribute Index		$-\ln \widehat{HERF}$	$-\ln \widehat{HERF}$		
Range	Avg	Gini	P10	P90	(P10-P90)
1.00 - 1.01	1.00	-0.086	-0.659	-0.982	-0.323
1.01 - 1.05	1.03	-0.077	-0.628	-0.931	-0.303
1.05 - 1.10	1.07	-0.065	-0.587	-0.864	-0.277
1.10 - 1.15	1.12	-0.050	-0.536	-0.779	-0.244
1.15 +	1.34	0.015	-0.309	-0.408	-0.098

Table XVI

Coefficient ($\beta_1 + \beta_3 Q$) on *Com* ($\ln \hat{N}$) at Different Output-Attribute Index Levels

Dependent Variable: Gini log odds ratio, 10th Percentile Price or 90th Percentile Price

Output-Attribute Index		$\ln \hat{N}$	$\ln \hat{N}$		
Range	Avg	Gini	P10	P90	(P10-P90)
1.00 - 1.01	1.00	-0.015	-0.690	-0.862	-0.172
1.01 - 1.05	1.03	-0.014	-0.683	-0.850	-0.168
1.05 - 1.10	1.07	-0.012	-0.673	-0.835	-0.162
1.10 - 1.15	1.12	-0.009	-0.660	-0.816	-0.155
1.15 +	1.34	0.001	-0.606	-0.730	-0.124

Table XVII
Panel Estimates
Dep Var: Gini Log Odds Ratio or Log of 10th or 90th Percentile Observed Price

Estimation Formula:	Gini Log Odds (11)	Log(P10) (11)	Log(P90) (11)
$-\ln \widehat{HERF}$	-0.620**	-2.077***	-2.932***
	(0.249)	(0.278)	(0.208)
$Q(w,\, y,\, q,\, q^0)^{-1}$	-0.115***	-0.105***	-0.162***
	(0.013)	(0.015)	(0.011)
Lag Dependent Variable	0.631***	0.511***	0.715***
	(0.003)	(0.003)	(0.003)
Interaction Variables			
$-\ln \widehat{HERF} \cdot Q(w,\, y,\, q,\, q^0)^{-1}$	0.577**	1.858***	2.652***
	(0.228)	(0.255)	(0.191)
Observations	73981	73981	73981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table XVIII

Panel Estimates

Dep Var: Gini Log Odds Ratio or Log of 10th or 90th Percentile Observed Price

	Gini Log Odds	Log(P10)	Log(P90)
Estimation Formula:	(11)	(11)	(11)
$\ln \widehat{N}$	-0.151	-1.283***	-1.593***
	(0.137)	(0.153)	(0.110)
$Q(w,\, y,\, q,\, q^0)^{-1}$	-0.101***	-0.086***	-0.131***
	(0.011)	(0.012)	(0.009)
Lag Dependent Variable	0.632***	0.509***	0.712***
	(0.003)	(0.003)	(0.003)
Interaction Variables			
$\ln \widehat{N} \cdot Q(w,\, y,\, q,\, q^0)^{-1}$	0.165	0.898***	1.211***
	(0.107)	(0.119)	(0.086)
Observations	73981	73981	73981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Table XIX
Panel Estimates
Dep Var: Gini Log Odds Ratio

Estimation Formula:	Difference GMM (9)	System GMM (9)	FOD Orthogonal (9)
$-\ln \widehat{HERF}$	-0.475***	-1.380***	-3.096***
	(0.113)	(0.290)	(0.190)
$Q(w, y, q, q^0)^{-1}$	-0.032***	-0.368***	-0.493***
	(0.012)	(0.036)	(0.014)
Lag Dependent Variable	0.413***	0.586***	0.566***
	(0.011)	(0.072)	(0.010)
Interaction Variables			
$-\ln \widehat{HERF} \cdot Q(w, y, q, q^0)^{-1}$	0.490***	1.264***	2.783***
	(0.103)	(0.263)	(0.205)
Observations	73981	73981	73981

Note - All regressions include, quarter and carrier dummies, and a dummy variable indicating whether any carrier in the market was currently in bankruptcy. Standard errors are in parentheses.

* Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level

Chapter 2 - Product Differentiation and Hedonic Prices: An Analysis of Airfares

Abstract

The main objective of the chapter is to study the effect of product differentiation on price formation in the airline industry. For this purpose, we introduce the concept of a core product and examine how differentiation beyond the core product effects pricing and mark-ups. We measure differentiation from the core using a Konüs type index of differentiation that is based on cost functions. We explore how using this index, measures of market power, and controlling for core product cost can improve the stability and results of a hedonic price model. The model is empirically tested on 103,980 observations of quarterly US domestic airfare data between 2002 and 2016.

2.1. Introduction

Firms selling differentiated products or services must make two important decisions, which product characteristics to include and at what price. This decision is often not only a question of which characteristics, but the level or degree of the characteristic. For example, an airline may choice to provide a customer service counter, should it be staffed 8, 12 or 24 hours a day? The airline must also decide how much, if any, first-class seating to provide, or decide how frequently to offer flights on a given route, bi-weekly, daily or more often? Expected service level goals have to be set, for example how is on-time arrival prioritized. Each choice drives a different level of required resources and presumably a different value for the customer.

Three key questions must be answered when a firm is deciding the level of product characteristics to provide. First, how much will it cost to provide the characteristic? Second, how much do consumers value the characteristic? Third, can the firm earn more profit by changing the level they provide of the characteristic? While these seem like relatively simple questions, answering them is not so easy.

The firm may have available accounting mechanisms to estimate the costs but answering the question on how much consumers will pay is more challenging. Differentiated products of the same general type are not perfect substitutes, as consumers place different valuations on characteristics. The challenge for the firm is to uncover consumer valuation and the distribution of tastes and demands. In one standard method of modelling characteristics and product differentiation, Mussa and Rosen (1978) define a single underlying characteristic that measures product quality, where larger levels of the characteristic indicate higher quality. The product is available on the market at multiple quality levels, and for any single quality level all purchasers pay the same price. In this model, market equilibrium is described by a price and quantity for each level of quality, and the range of qualities available.

When products are thought of as a bundle of characteristics, one method frequently used for measuring consumer valuation of the characteristics is the hedonic pricing model. Rosen (1974) defines hedonic prices as the set of implicit prices of characteristics that are revealed from the observed level of characteristics and prices of the product. Unfortunately for pricing practitioners, empirical applications of hedonic pricing models often do not provide clear answers on how much consumers are willing to pay. There are three specific problems with the standard hedonic model that our model attempts to alleviate.

The first of these is the “wrong” sign or unreasonable coefficient value. This problem is discussed by Triplett (2004) who points to a number of possible reasons. The first relates to correlation errors that can occur duet to engineering relations along the full scale of the characteristics. For example, in auto industry hedonics the fuel consumption variable often has the wrong sign due to correlation with the size (cost) of the auto. A second common problem of the hedonic model is coefficient instability. This problem is discussed by Pakes (2003) who shows that hedonic coefficients are the combination of a marginal cost function and a complex function that summarizes the relationship between mark-up and the characteristics. Changes in either function can cause coefficient changes. The final problem concerns competing products. By construction, a hedonic coefficient is the measure of consumer value at the mean of the characteristic for all products on the market. However, for products that compete in more localized markets, for example retail stores or airline routes, the more relevant comparison is to available substitutes, not the whole industry.

The model we present in this chapter solves these three problems and provides a more useful tool for practitioners. First, to disentangle customer valuation for characteristics correlated with the cost of the product we introduce the concept of a "core product" and show how accounting for core cost can improve the results of a hedonic function. Separating this core cost, and controlling for market power, helps us solve the second problem raised. By removing most of the cost component from the hedonic coefficient we increase stability by minimizing the sources of coefficient change. Finally, removing the core product allows our analysis to focus on a narrower variable range and only focus on the difference between competitors in the markets they serve. In a study of brand equity Baltas and Saridakis (2010) discuss this concept as the premium that costumers are willing to pay for a characteristic in comparison to a competitor offering a similar product. Being able to understand consumer valuation of a characteristic above the market minimum can be more useful to a firm in product positioning than understanding customer valuation at the industry average.

This chapter aims to answer three specific research questions. First, can we improve the stability and interpretability of a hedonic estimation by accounting for market power, and the cost of the core product? Second, does the effect on mark-up vary by characteristic? Third, in an empirical application, can these methods answer the firm questions we started this chapter with. Is the firm able to make a higher profit by providing these characteristics?

To answer these questions, we combine two well-known models. First, utilizing cost functions we estimate the cost of the core product, and the minimum cost of providing different levels of characteristics. Also using cost functions, we generate an index of firm heterogeneity that measures firm differentiation by characteristic. We then present a modified version of the Pakes (2003) hedonic price model that accounts for both core product cost, market power, and differentiation. The model is empirically tested on the US domestic airfare market on product characteristics such as class of service, on-time arrival, and frequency of departure.

This chapter contributes to the ongoing research into how to handle the effects of market power in hedonic regressions, (Taylor and Smith, 2000; Chwelos, Berndt, and Cockburn, 2008; Cotteleer, Gardebroek, and Luijt, 2008; Abrate and Viglia, 2016) and the line of research on improving the usability of hedonic models, (Heckman, Matzkin, and Nesheim, 2010; Bourassa, Cantoni, and Hoesli 2016). Our inclusion of the level of firm differentiation is a unique contribution to this stream of literature. While the use of cost functions to analyze the airline industry is relatively common, the application of hedonic pricing models to airfares is not. In one of the few, Morrison and Winston (1995) employ a series of hedonic equations to calculate marginal valuation of choice characteristics such as frequent flyer miles, fare restrictions and multimarket contact. A study by Good, Sickles, and Weiher (2008) also applied the hedonic method to airfares, but their focus was on adapting it's use for inclusion as part of the consumer price index rather than as tool to diagnose product decisions. Finally, Borenstein (1989) used a hedonic model to investigate market power that arises when a carrier dominates a hub. To the best of our knowledge, our use of product characteristics to explain mark-up and our definition of mark-up outlined in section 5 is a unique contribution, as is our inclusion of estimated costs directly within the hedonic pricing model.

This chapter is organized as follows. In section 2 we review the relevant literature related to both hedonic pricing and cost functions. In section 3 we provide background and develop our proposed hedonic function. Section 4 outlines the methodology used for the cost function and the hedonic functions. Section 5 presents the data used in the study, Section 6 the results and Section 7 concludes.

2.2. Literature Review

2.2.1. Hedonic Price Literature

We begin this section with a review of hedonic price literature to provide a background to the method and to highlight the issues we hope to confront. For more complete surveys of the method we refer the reader to Armknecht and Ginsburg (1992) and Good et al. (2008) for applications to airfares, and Triplett (2004) and Hill (2013) for more general surveys.

Much of the theoretical work related to hedonic prices has been motivated by attempts to improve price indexes. In one of the first works in the field, Court (1939) starts with the first line "No valid price comparisons can be made without adequate commodity standards, i.e., definition of the articles priced in terms of their usefulness and desirable physical characteristics." The focus of this work was on showing how the U.S. Bureau of Labor Statistics auto price index overestimated increases in prices by ignoring changes in size, equipment and quality. The article introduces a hedonic method to establish and measure the usefulness and desirability of different characteristics of an automobile (e.g. horsepower, seat width, tire size). Usefulness is measured through multiple regression, with price as the dependent variable and characteristics as independent variables. Including a time trend allows for a measure of change in price over time holding characteristics constant. In an extension of this, Griliches (1961) looks at the relationship between prices and dimensions of autos in the years 1937, 1950 and 1954 through 1960. The article uses the results of a hedonic equation to build a quality index that can deflate an existing price index. It also raises the issue of coefficient instability, and points to variable correlation, shifting supply conditions and changes in consumer tastes as causes of instability.

Supply and demand conditions play a large role in the seminal work by Rosen (1974). This paper provides a theoretical foundation for hedonic analysis of differentiated products by combining a family of utility maximizing bid functions and profit maximizing offer function. The collection of tangent points between bid and offer curves for each characteristic form a locus of equilibrium prices in a competitive market. Feenstra (1995) expanded this theoretical foundation beyond the competitive case, showing that under imperfect competition we must consider any mark-up as an omitted variable. This means that the coefficients on product characteristics cannot be interpreted directly as marginal consumer values. Since market power allows for mark-up, this value is embedded in the coefficient along with marginal consumer value of the product characteristic. If market power is not accounted for, the resulting coefficient is biased upward by the price-cost ratio. In an empirical study on the cable TV Industry, Anstine (2001) addresses market power by including two control variables, the presence of a competitor and the availability of over the air programming.

Building on this, Pakes (2003) shows that when mark-ups exist, the characteristic hedonic function is a sum of the marginal cost function and a function that summarizes the relationship between mark-ups and characteristics. When either of these “primitives” change we would expect the hedonic coefficient to change. Following Feenstra (1995) and Pakes (2003) a number of articles attempt to control for the presence of market power and address the concern that mark-ups can bias hedonic coefficients. In an application to farmland pricing, Cotteleer et al. (2008) proxy for market power based on the potential number of buyers and sellers. More recently, in research on hotel revenue management, Abrate and Viglia (2016) include the number of competitors with available hotel rooms and real-time room prices.

2.2.2. Cost Function Literature

In this section we provide a short review of the literature we base our cost function on. The focus of this chapter is not on cost function estimations, but we do rely heavily on their use. We do contribute to this stream in terms of the use of cost functions in airline research.

In one of the earlier papers on the subject, Christensen and Greene (1976) extend the work of Nerlove (1961) noting that given certain regularity conditions there exist both a production and cost function which would be dual to each other. Specifying a certain production function implies a certain cost function, and vice-versa. They also note that using a production function is more appropriate when output levels are endogenous, while cost functions are better suited when output levels are exogenous. This supports our use of a cost function as we assume exogenous demand. Concerning functional form, they demonstrate the attractiveness of the translog form as it places no restrictions on input substitution and allows for scale economies to vary by output level.

The translog form is also used by Caves, Christensen, and Tretheway (1984) in one of the first applications to airline cost. Their focus was on comparing cost structures, returns to scale and returns to density between trunk carriers and regional carriers. Using standard inputs and outputs[20], and controlling for stage-length and load factor, they find evidence of constant returns to scale and increasing returns to density for both trunk and regional carriers. In terms of unit cost, they attribute most of the difference to density and stage length.

In a more recent paper, Bitzan and Peoples (2016) focus on comparing cost change over time at low-cost and full-service carriers in the US airline industry. They take a total factor productivity approach and estimate a translog cost function defined by factor prices, output, technological characteristics and a time trend. They find that even though the sources of cost

---

[20] These will be discussed in section 6

change differ between the two, there is evidence of cost convergence over the 1993 to 2014 period studied. Full-service carriers have reduced cost, while low cost carriers have seen increases, though they still retain a cost advantage. A final example, especially relevant given our use of hedonic pricing, can be found in Arrondo, Garcia, and Gonzalez (2018). In this work, they use stochastic frontiers to examine pricing efficiency, combining hedonic pricing with frontier analysis.

2.2.3. Cost Pass-Through Literature

A standard hedonic function generally only accounts for product price and product characteristics. By including cost and mark-up in our model some parallels can be drawn between our work and the literature that studies cost pass-through. While the source of the cost shock differs, some of the concepts discussed and tools used are similar.

In a study of pass-through of emissions cost to electricity prices Fabra and Reguant (2014) find that 80% of emissions cost increase is passed through to price. In their article they discuss three channels that account for incomplete pass through, adjustment of mark-up, the presence of costs not affected by the cost shock, and price rigidities that affect ability to adjust prices. In this chapter we focus on the first channel of adjusting mark-up as a mechanism. Studying cost pass-through in the airline industry, Agarwal et al. (2015) found complete pass-through of fuel price increases. In another study, the effect of product differentiation on cost pass-through was examined by Loy and Weiss (2019) who found that in the yogurt industry, higher levels of differentiation were associated with lower pass through rates of shocks to milk prices. They attribute this to an initially higher margin point and rigidity in retail prices.

There are however three primary difference between the pass-through literature and the model we present in this chapter. First, the cost shocks in pass-through models only affect product cost, and not consumer benefit. In our model a cost increase associated with a higher level of characteristic also increases product value to the consumer. Second, cost shocks in pass-through are thought of as more transitory and random, whereas a change in product characteristics can be considered as a change to a new level that persists until another change occurs. Finally, cost shocks in pass-through literature are generally considered to affect the whole industry, where our cost changes are firm specific.

2.2.4. Aviation Business Models

In this chapter we look at product price formation as a function of the separate characteristics that make up the product. An alternative method of analysis may be examining the effect that different business models have on price formation, where a business model is loosely defined as a common set of characteristics. This was the route taken by Gillen (2006) who studied airline business models and network. This work describes two existing models, the full-service airline (FSA) model and the Low-Cost Carrier (LCC) model, with the primary distinction between the two being network type, hub-and-spoke for FSA and point-to-point for LCC carriers. Based on compatibility with business model and network, the characteristics we examine would tend to be higher for FSA carriers. In another work that discusses these two models Franke (2007) points out the danger of being positioned in the middle between premium carriers and LCC and notes that many of the legacy FSA carriers have not adjusted their business models.

A third business model is discussed by Bachwich and Wittman (2017) in their study of the effects of Ultra-Low-Cost Carrier (ULCC). These carriers are characterized by significantly lower cost structures, led by low labor cost, and aggressive unbundled fares and sales of ancillary services. This study, coming ten years after the previous two, notes the FSA and LCC models are beginning to blend. Our work differs from this literature stream in that we focus on the characteristics as separate elements, rather than bundling them into a business model allowing us to analyze the effect of each characteristic separately.

2.3. Background

2.3.1. Hedonic Price Function

As described in Rosen (1974), a hedonic equation maps the locus of consumer bid and producer offer functions for the characteristics that make up a product. Expanding on the producer offer function, this function indicates the unit price a firm will accept for different designs of the product at a constant profit, when production quantities are optimally chosen[21]. At that point, the parameters on characteristics indicates the marginal reservation supply price for the characteristic at constant profit. In this paper, as with others in the hedonic literature, we then refer to the parameter on the characteristic as the marginal cost of the characteristic from the production standpoint and the marginal value from the consumer point of view.

Through relation to price, the model captures consumer valuation and producer cost. This dual valuation adds complexity to variable selection, since not all characteristics are relevant to both groups. In our airline example product characteristics that are relevant to consumers and producers may include; miles travelled, duration of flight, necessity of a connection, probability of arriving on-time and choice of departure time. For the firm, additional characteristics such as route density, input prices and load factors are also relevant.

Before continuing our discussion of the hedonic approach we would like to define the variables used in the following, sections. We denote output quantity as $y \in \mathbb{R}_{++}$, the vector of

---

[21] See Rosen (1974) for a full development of this model and additional detail

inputs as $x \in \mathbb{R}_{+}^{m}$, the vector of input prices as $w \in \mathbb{R}_{++}^{m}$, the vector of output prices as $p \in \mathbb{R}_{++}^{l}$, and finally $z \in \mathbb{R}_{+}^{n}$ as a vector of *n* characteristics. Following Rosen (1974), the hedonic model defines *a class of products* that can be described by the vector of characteristics *z*, where $z = (z_1, z_2, \dots, z_n)$. The components of z are objective measures that are perceived by all consumers and producers identically but may be valued differently. We use the term "version" to designate a product of a given specification with characteristics *z*. An excellent illustration of this idea can be found in Carew (2000). In this chapter the class of products is "apples" and the vector of defining characteristics includes cultivar, grade, package size and fruit size. A single version *j* of the product can be specified by the vector of characteristics *z*, for example version z_j = (Fuji, Canada Extra Fancy, tray pack, medium) and version z_h = (Spartan, Canada Extra Fancy, consumer package, medium). In our empirical analysis, the class of products is a flight between two city pairs.

Each product also has an observed market price and associated characteristic vector z such that the market for the product implicitly reveals a function $p(z) = p(z_1, z_2, \dots, z_n)$. The function $p(z)$ represents the minimum price for any bundle of characteristics, because if the same bundle was offered at other prices consumers would only choose the less expensive one. Rosen (1974) adopts the convention of measuring characteristics $z_i,\ i = 1, \dots n,$ in a way that all can be considered "goods". This would imply that all consumers give them a positive marginal valuation in the neighborhood of their minimal technically feasible amount.

To estimate the hedonic function, we define p_j as the observed price of version j of the product with observed vector of characteristics z_j. We then relate p_j with each of the characteristics $z_i,\ i = 1, \dots n$ and add an error term. A simple example of a linear form would be,

$$p_j = \beta_0 + \sum_{i=1}^{n} \beta_i z_i + \varepsilon. \tag{1}$$

In this example, the hedonic weights β_i are the portion of the product price that can be attributed to characteristic z_i. Hulten (2003) outlines two basic approaches in the literature to interpreting the characteristic prices. The first relates the coefficient β_i to consumer willingness-to-pay, and is where the term "hedonic" was coined by Court (1939) and other early writers. This utility based interpretation was the basis for the proposal by Lancaster (1966) that consumer utility theory be based on characteristics of goods, rather than on goods themselves.

The second approach, introduced by Rosen (1974), likens the hedonic function to utility driven bid and offer curves for individual characteristics. Bid curves formed by consumers with heterogeneous tastes for combinations of characteristics and offer curves formed by the firms cost of supplying characteristics. The function then defines the intersections of bid and offer for characteristics and can be viewed as an envelope linking together the equilibriums. Under this approach coefficients are considered both customer value driven, and producer cost or resource driven. This is most clear in the competitive market case where price equals marginal cost. Under the assumption of a competitive market the market price is equal to marginal cost, so we can interpret (1) as either price or cost and that characteristics z_i are both valued by consumers and drive production cost. This interpretation of β_i has inspired much debate, outlined in Triplett (1991). The debate centered on whether to interpret hedonic measures as a resource-cost or user-value concept. His conclusion was that the concepts are not necessarily competing, but that they represent different uses of the data.

The model of Rosen (1974) is now the more generally accepted view and has been built on by Feenstra (1995) and Pakes (2003) who extend it beyond the competitive market. Relaxing the assumptions of the perfectly competitive case we now have the possibility that price exceeds cost. Pakes (2003) shows that if we let (z_j, p_j) denote the vector of characteristics and price of version j and (z_{-j}, p_{-j}) the vector of characteristics and price for all other versions of product j, then demand for version j is $D_j(\cdot) = D(z_j, p_j, z_{-j}, p_{-j}, A)$ where A is the distribution of consumer preferences for characteristics. Assuming a single product firm and marginal cost of $mc(\cdot)$, then $p_j = mc(z_j) + D_j(\cdot)/|\partial D_j/\partial p|$, where the second term is typically referred to as the mark-up. Simply put, the price of version j is the marginal cost of producing a version with characteristics z_j plus a mark-up based on characteristics z_j and other available versions. From this we can define the hedonic function $h(z_j)$ for version j, as the expectation on price conditional on z as,

$$\begin{aligned} h(z_j) &\equiv E[p_j|z_j] \\ &= mc(z_j) + E\left(\frac{D_j(\cdot)}{|\partial D_j/\partial p|}\,|z_j\right). \end{aligned} \tag{2}$$

The expected price of version j is the sum of the marginal cost of producing characteristics z_j plus the expected mark-up conditional on z_j. The mark-up can be described as a complex function which varies inversely with elasticity of demand and is dependent on characteristics of competing goods and consumer preferences. The resulting coefficient of the hedonic equation is comprised of the marginal cost function and the complex mark-up function, changes in either of these two will result in changes in the coefficients and can create instability.

2.3.2. Hedonic Mark-Up Function

In this chapter we attempt to sidestep this instability and the debate over coefficient interpretation by presenting a different way of thinking about the hedonic price model. We begin by defining $\bar{z}$

as a vector of characteristics where $\bar{z} \subseteq z$, in other words $\bar{z}$ cannot have more elements than z. Vector $\bar{z}$, which we term "core characteristics", only includes *quantifiable* characteristics that can be measured for all versions and is used to define what we term the "core product". As a simple example, all customers on a specific route travel the same distance, and all have one take off and one landing, regardless of which airline they fly. In addition, this core product may include characteristics measurable for all airlines such as frequency of flight or on-time performance.

In the case of the airline industry, defining the product as a flight between two city pairs, characteristic vector z can include the distance traveled, operating carrier, class of travel, seat selection, day of travel, quality of service, flight frequency or on-time performance among others. The subset of core characteristics $\bar{z}$ would only include distance traveled and those characteristics that are quantifiable. This echoes Rosen (1974) who defined characteristics as those that can be considered goods and have a positive marginal valuation.

Using subset $\bar{z}$ we define the "core product" as the version with the most basic characteristics of the product available and formally described it by the vector $\bar{z} \in \mathbb{R}^n_+$. The operational approach of the core product is to consider the minimum values of characteristics available in the market for all versions j of the product, or $\bar{z} = \{\overline{z_1}, \dots, \overline{z_n}\}$ and $\overline{z_i} = Min\{z_{ij}, j = 1, .., k\}, i = 1, .., n$. We also define $d \in \mathbb{R}^n_+$ as the vector of differentiating characteristics or the difference between z_i and $\overline{z_i}$. This would also imply $z_{ij} = \overline{z_i} + d_{ij}, i = 1, .., n; \; j = 1, .., k$. In other words, the characteristics for any version *j* of the product are equal to the value of the core characteristic, plus the value of any differentiating characteristic. The value of a core characteristic can be zero, meaning that the characteristic is not provided.[22]

---

[22] As a numerical example, picture a market with two airlines, Delta and Southwest. Now consider z to include four characteristics (*Carrier, Flight Frequency, On-Time Performance, Seat Selection)* and that *Seat Selection* is measure as 1 if available and 0 if not. If characteristics z are

Using the previous operational definition of the core product to extend Pakes (2003), we now let $(\bar{z}, d_j, p_j)$ denote the core characteristics, differentiating characteristics and price of version *j*, and $(\bar{z}, d_{-j}, p_{-j})$ for all other version. Recall from the definition of the core product that $\bar{z}$ is the same for all versions. Demand for version *j* now becomes $D_j(\cdot) = D(\bar{z}, d_j, p_j, d_{-j}, p_{-j}, A)$ where *A* again is the distribution of consumer preference over characteristics. Based on this, we extend (2) to include both the core product and output attributes as,

$$
\begin{aligned}
h(z_j) &\equiv E\left[p_j \middle| z_j(\bar{z}, d_j)\right] \\
&= mc(\bar{z}) + mc(d_j) + E\left(\frac{D_j(\cdot)}{|\partial D_j / \partial p|} \middle| z_j(\bar{z}, d_j)\right).
\end{aligned}
\tag{3}
$$

We can understand (3) as the expectation on price of version *j* conditional on $z_j(\bar{z}, d_j)$ is the sum of the cost of producing core characteristics $\bar{z}$, the cost of producing differentiating characteristics d_j, and a mark-up conditional on $z_j(\bar{z}, d_j)$. This extension of Pakes (2003) separates cost into two components, the cost of producing the core product and the cost of producing characteristics beyond the core that differentiate this version. In the next two sections we discuss estimations of the cost of producing the core product and the differentiating characteristics, and an empirical estimation of (3).

Our extension in formula (3) is conceptually similar to Savioli and Zirulia (2020) who analyze how the presence of add-ons effects baseline prices. In their duopoly model they find that

---

respectively (*Delta, 4, 75%, 1)* and (*Southwest, 2, 80%, 0)*, then $\bar{z}$ would include (*Flight Frequency, On-Time Performance, Seat Selection)* and values for this market would be (*2, 75%, 0)*, d_{Delta} would be *(Delta, 2, 0, 1)* and $d_{Southwest}$ would be *(Southwest, 0, 5%, 0)*.

when both firms offer the add-on, the price for the baseline product is lower than in cases where no add-on is offered. However, in the asymmetric equilibria, where only one firm offers the add-on, the effect is dependent on the degree of market power. There are some major differences in our approach though. Their definition of baseline product allows for horizontal differentiation between the firms, while our core product defines a product that is equivalent for all firms. A further difference, and one of our primary contributions, is the inclusion of a cost function as part of the hedonic model.

2.4. Methodology

2.4.1. Cost Approach

In a seminal paper on hedonic cost functions Spady and Friedlaender (1978) suggest treating output as a function of physical output and its characteristics, instead of treating each combination of output and characteristics as a unique output. This method has been applied to a wide range of industries, for example freight transport (Spady and Friedlaender, 1978), airlines (Gillen, Oum, and Tretheway, 1990) and the military (Hanson, 2016).

To begin, we define the production possibility set T as the output quantity and the associated product characteristics that can be produced for a given set of inputs as $T = \{(x, y, z)\text{: } x \text{ can produce } y \text{ with characteristics } z\}$. Technology set T defines the input set x required to produce an output quantity y with characteristics z as $L(y, z) = \{x\text{: } (x, y, z) \in T\}$. In other words, output quantity y with characteristics z can be produced with the input vector x. Our treatment of this model is somewhat different from Spady and Friedlaender (1978) as we allow output quantity and product characteristics to separately affect cost.

Our interest is on the cost of producing output quantity y with characteristics z. To estimate this, we move to a cost function, the dual of the above production function. Adding input prices w, we define the least expensive bundle of inputs to generate a given output quantity and level of characteristics. The minimum cost for a given output quantity y, characteristic level z and input prices w can be defined as

$$c(w,y,z) = \min\{w^T x : x \ \in L(y,z)\}. \tag{4}$$

From (4) we see that minimum cost would be greater with an increase in output quantity y or an increase in characteristic level z. Recalling the definition of a core product, we would find that $z \geq \bar{z}$ and that $x(z) \geq x(\bar{z})$ and therefore $c(w,y,z) \geq c(w,y,\bar{z})$.

In (4) we derive minimum total cost of providing output quantity y with core characteristics $\bar{z}$ and input prices w. However, our intention is to define the cost of the core product at the level of a single passenger fare. Based on this we move from total cost to an average or unit cost as,

$$ac(w,y,\bar{z}) = \frac{c(w,y,\bar{z})}{y}. \tag{5}$$

The average to produce a given output quantity y with core characteristics $\bar{z}$ at input prices w available at the time. In the following section we use this core product unit cost in conjunction with the hedonic price function. Previous literature on estimating airline cost functions has used a number of different methods[23] and functional forms. In this chapter, we specify a stochastic

---

[23] As examples, stochastic parametric frontiers (Gillen et al., 1990; Ahn and Sickles, 2000; Assaf, 2009 and Johnston and Ozment, 2013), deterministic non-parametric frontiers (Barros et al., 2013 and Arjomandi and Seufert, 2014), and Tsionas (2003) who combined a non-parametric method with a stochastic frontier.

parametric cost function and a translog functional form[24]. Characteristics enter the function as shifters and are assumed to have a direct linear influence on the production structure and cost. Put another way each firm potentially faces a different frontier at each period given the effects of characteristics on the technology. From this estimation we can extract the cost elasticity estimates on product characteristics, allowing us to compare them to those from the hedonic regression.

In addition to output quantity, input prices and characteristics, we include a time variable (*t)* to account for the panel nature of our data set. Variable *t* takes the value of 1 in the first quarter, 2 in second quarter, and so on, and captures cost changes due to technical progress or regress over time[25]. All values are in logarithmic form and all dollar figures have been converted to 2002 dollars using the producer price index (PPI) for the airline industry. Product characteristics *z* are treated as cost shifters and are not included in interaction effects. With C_{ht} as the total cost for firm *h* at time *t* the translog cost function to be estimated is,

$$\begin{aligned}\ln C_{ht} = {} & \sum_{k=1}^{m} \alpha_k \ln w_{kht} + {}^1\!/_2 \sum_{k=1}^{m} \sum_{j=1}^{m} \alpha_{kj} \ln w_{kht} \ \ln w_{jht} \\ & + \alpha_y \ln y_{ht} + {}^1\!/_2 \ \alpha_{yy} (\ln y_{ht})^2 + \sum_{k=1}^{m} \alpha_{ky} \ln w_{kht} \ln y_{ht} \\ & + \sum_{i=1}^{n} \alpha_i \ln z_{iht} + \alpha_t t + \varepsilon_{ht}; \ \ h = 1, \ldots, l\,, t = 1, .., T.\end{aligned} \tag{6}$$

To ensure the estimated cost function is homogeneous of degree one in input prices, and to ensure symmetric cross effects, the following typical restrictions are imposed.

---

[24] The translog form, proposed by Christensen, Jorgenson, and Lau (1973), has been widely used in airline cost functions since Caves et al. (1984) study on the economies of density and scale..

[25] We considered including a variable to indicate the airline, however in most cases product characteristics and the airline are very closely correlated.

$$\sum_{k=1}^{m} \alpha_k = 1, \ \sum_{k=1}^{m} \alpha_{kj} = \sum_{k=1}^{m} \alpha_{jk} = \sum_{k=1}^{m} \alpha_{ky} = 0$$

As an additional restriction we apply Shephard's lemma that input shares be equal to the derivative of the cost function with respect to input prices and impose the following to obtain the share of input k in total cost (S_k),

$$S_k = \frac{\partial \ln C}{\partial \ln w_k} = \alpha_k + \sum_{k=1}^{m} \alpha_{kj} \ln w_{jht} + \alpha_{ky} \ln y_{ht} \,; k = 1, \dots m.$$

Where $S_k = w_k x_c / C$ it follows that $\sum_{k=1}^{m} S_k = 1$. To satisfy homogeneity of price we follow previous work and normalize input prices and cost by one of the input prices.

The coefficients found in (6) are cost elasticities. In the applied section we use these elasticities to predict cost at the observed level of the characteristic z_{iht} and at the level of the core product in the market, given the firms output level.

2.4.2. Hedonic Price Function

Our first research question outlined in the introduction asks whether it is possible to improve the results of a hedonic estimation by accounting for market power and the cost of the core product. To that end we now introduce a series of hedonic estimations to test this. A hedonic regression can be expressed in a few functional forms, with linear, semi-log or double-log being some of the most common. We adopt the double-log form for all the hedonic regressions as it is compatible with the functional form of the cost function introduced in the previous section.

Individual fares are determined by many factors that are not observable in our data. Factors like purchase lead time (number of days before departure), purchase methods (agent, direct, on-line), seats available at time of purchase or other fares available at time of purchase. To avoid bias due to unobserved factors on individual fares we define price as the average fare paid on an airline and route in a quarter. We define a route as a non-stop flight[26] between any two city pairs. Like input prices in the cost function, we adjust all fare prices to 2002 prices using the bureau of labor statistics consumer price index for airline fares[27]. Our baseline estimation is founded on the basic fare equation of Morrison and Winston (1995) and includes the fare price as the dependent variable, the distance traveled as the measure of output and vector z product characteristics.

In the baseline equation defined below we estimate mean fare p for airline h on route j at time t as a function of route distance y and characteristics z. We include t as a quarter dummy to capture any time related effects and an error term,

$$\ln p_{hjt} = \beta_0 + \beta_y \ln y_j + \sum_i^n \beta_{zi} \ln z_{ihjt} + \beta_t t + \varepsilon_{hjt} \ . \qquad (7)$$

The first extension of baseline estimation (7) is the addition of variables to capture competition and market power. Following Morrison and Winston (1995) we add the number of route passengers, the number of route competitors, and the number of airport competitors. The number of passengers on the route measures the size of the market, we would expect an increase in market size to lead to reductions in price. Passenger numbers are noted per 1,000 passengers and include all passengers on the route flying in either direction for all airlines. Regarding competition, we account for both current and potential competition. The number of actual route competitors serves as the measure of current competitors. The number of potential competitors borrows from ideas presented by Bailey and Baumol (1984) on contestable markets. In this work they argue that the price dampening effect of perfect competition can arise solely due to the threat of entry and exit.

---

[26] Non-stop flights are flight with no intermediate stops. A direct flight may include a stop.

[27] https://www.bls.gov/ series CUUS0000SETG01 (accessed 20/09/2019)

To measure this, we use the average of the number of competitors at the endpoint airports. The logic being that firms that have a presence in the airport can freely begin to compete on a route. We would expect an increase in the number of airport competitors to lead to reductions in price.

We consider two other measures of competition, the potential effect of mergers or acquisitions, and the competitive effect of adjacent routes. Over the period under observation there were a significant number of mergers and acquisitions[28] which could potentially reduce competition. We assumed that a merger that occurred in a market would have an effect for four quarters beginning with the quarter in which the merger closed, and created a dummy variable for the occurrence of a merger in each market *j* at time *t*. The value of the dummy is equal to 1 in a market where two of the carriers in that market merged, and 0 otherwise.

We also looked at including potential competition from adjacent routes. In an article studying the effect of Southwest Airlines on competition Morrison (2001) notes three ways airfares on a route can be influenced by another airline. First, the airline could serve the route. Second, they could serve a route that would be considered a substitute. Third, the presence of the competitor airline in the same, or a nearby airport, could cause fare reductions on the route. The article defines nearby as being within a 75-mile radius. The article outline multiple combinations of these factors based on origin and destination point. We did not include this measure of competition, as we are already accounting for much of this within the model. In terms of substitutes, we are already combining the observations for airports that serve the same city, and we are accounting for presence of competitors in the origin or destination airports.

---

[28] In 2005 US Airways merged with America West, in 2009 Northwest and Delta merged; in 2010 Continental joined United; in 2011 Southwest acquired AirTran; in 2013 American and US Airways merged.

The three measures of competition, number of route passengers, number of route competitors, and number of airport competitors are added to estimation (7) at the route level as $Market_{ljt}$,

$$\ln p_{hjt} = \beta_0 + \beta_y \ln y_j + \sum_i^n \beta_{zi} \ln z_{ihjt} + \sum_{l=1}^3 \beta_{ml} \ln Market_{ljt} + \beta_t t + \varepsilon_{hjt} \ . \qquad (8)$$

A potential method to refine (8) and create a bridge to the next section, where we fully introduce costs into the model, would be to replace the scalar value of characteristics z with their predicted costs at values z_{ihjt} and estimate,

$$\ln p_{hjt} = \beta_0 + \beta_y \ln y_j + \sum_i^n \beta_{zi} \ln c(z_{ihjt}) + \sum_{l=1}^3 \beta_{ml} \ln Market_{ljt} + \beta_t t + \varepsilon_{hjt} \qquad (9)$$

In (9) we would interpret β_{zi} as the expected increase in price for a one percent increase in cost level $c\left(z_{ihjt}\right)$.

Recalling that the coefficients of a hedonic function are understood to define both customer valuation and producer cost, we can clearly see in (8) and (9) why this definition may cause some debate. While changes in output measure y product characteristics z can effect cost, it is less clear why the level of competition on a single route would affect cost. It is also not clear where the effect of any mark-up can be seen. We might expect β_{ml} on *Market* to reflect mark-up, but it also may be present in β_y on output and β_{zi} on characteristics. To sidestep these issues, and answer our remaining research questions, we now depart from the standard hedonic estimation and convert (3) to a hedonic mark-up function.

2.4.3. Hedonic Mark-Up Function

The first step in generating the hedonic mark-up function is defining a new dependent variable. Instead of fare p we now measure the effect on the ratio of the mean fare to the cost of the core product, or mark-up as a percentage. Given the log nature of the hedonic regression and the cost function we find this as,

$$M_{hjt} = \ln\left(\frac{p_{hjt}}{c_{hjt}\left(w_h, y_{hjt}, \bar{z}_{jt}\right)}\right) \tag{10}$$

conceptually, this is equivalent to moving $mc(\bar{z})$ to the left-hand side of equation (3).

As our equation now only measures mark-up over the cost of the core product, characteristic vector z is less relevant to the estimation. What is relevant are the characteristics the firm is providing beyond the level of the core product. To proxy this, we include an index of firm heterogeneity introduced in Chapter 1. This index measures how much a firm's product differs from others in the market based on the cost of producing characteristics.[29] Following the notation in this chapter, we redefine this index as

$$Q(w_h, y_h, z_h, \bar{z}) = \frac{c(w_h, y_h, \bar{z})}{c(w_h, y_h, z_h)};\ \ h\ = 1, .., l \tag{11}$$

The index value ranges from (0,1] and measures the difference in cost between the observed level of characteristics of firm h and the level in the core product. An index value of 1 indicates that the characteristics provided by the firm are not different from the core product, while a value less than

---

[29] The index follows the logic given by a Konüs price index. See Grifell-Tatjé and Lovell (2015, Ch 7) for a discussion of a Konüs price index and the use and decomposition of a cost frontier.

1 indicates differentiation. For example, a value of 0.90 would imply a core cost that is 10% less than the minimum cost with observed characteristics.

We retain the *Market* variables to capture market power and competition. To ease interpretation we use the inverse of $Q(w_h, y_h, z_h, \bar{z})$ found as $c(w_h, y_h, z_h)/c(w_h, y_h, \bar{z})$ and noted as Q^{-1} and have,

$$M_{hjt} = \beta_0 + \beta_y \ln y_j + \beta_Q \ln Q_{hjt}^{-1} + \sum_{l=1}^{3} \beta_{li} \ln Market_{ljt} + \beta_t t + \varepsilon_{hjt}. \qquad (12)$$

Similar to equation (8), we would expect increases in the *Market* variables to lead to reductions in mark-up. The index $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ now captures the effect of any differentiation beyond the core product and we would expect coefficient β_Q to be positive. Since the firm is differentiated from the core product it has a mark-up over the cost of the core product.

Coefficient estimates $\hat{\beta}_Q$ in (12) provide more information than just the direction and magnitude of change in mark-up. This is easier to see by reorganizing the estimation results of (12) to find the effect of a change in Q_{hjt}^{-1} with all else constant. We have :$M_{hjt} = \hat{\beta}_Q \ln Q_{hjt}^{-1}$, and the from the definition of definition of M_{hjt} in (10) and the definition of $Q(w_h, y_h, z_h, \bar{z})$ in (11) we can derivate from (12),

$$\ln\left(\frac{p_{hjt}}{c\left(w_{hjt}, y_{hjt}, \bar{z}_{jt}\right)}\right) = \hat{\beta}_Q \ln\left(\frac{c\left(w_{hjt}, y_{hjt}, z_{hjt}\right)}{c\left(w_{hjt}, y_{hjt}, \bar{z}_{jt}\right)}\right). \qquad (13)$$

From (13) we see that an increase in characteristics z_{hjt}, and therefore an increase in $c_{hjt}\left(w_h, y_{hjt}, z_{hjt}\right)$ leads to an increase in mark-up over core cost equal to the percentage change multiplied by $\hat{\beta}_Q$. A percentage change in mark-up though does not directly tell us how a cost increase is reflected in price increase. From (13) we can note that the denominator on the left-hand and right-hand sides are equal, meaning that the percentage change in $c(w_{hjt}, y_{hjt}, z_{hjt})$ is directly reflected as a percentage change in p_{hjt}. When $\hat{\beta}_Q > c(w_{hjt}, y_{hjt}, z_{hjt})/p_{hjt}$ that change results in a price increase that exceeds cost increase.

The effect captured in equation (12) for index $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ is the effect of all product characteristics combined. To move this to an estimation for individual characteristics we can take advantage of the log nature of the translog cost function. The value of $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ can be calculated from the cost function, expression (6), after parameters $\hat{\alpha} = \{\hat{\alpha}_k, \hat{\alpha}_{kj}, \hat{\alpha}_y, \hat{\alpha}_{yy}, \hat{\alpha}_{ky}, \hat{\alpha}_i, \hat{\alpha}_t\}$ are estimated, and noting the definition of $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ we have,

$$[Q(w_h, y_h, z_h, \bar{z})]^{-1} = \exp[\ln c(w_h, y_h, z_h) - \ln c(w_h, y_h, \bar{z})]$$

$$= \exp\left[\sum_{i=1}^{n} \hat{\alpha}_i \ln z_{iht} - \sum_{k=1}^{n} \hat{\alpha}_i \ln \bar{z}_{it}\right] \qquad (14)$$

$$= \exp\left[\sum_{i=1}^{n} \hat{\alpha}_i \ln\left(\frac{z_{iht}}{\bar{z}_{it}}\right)\right]. h = 1, \ldots, l\,, t = 1, .., T.$$

This allows us to replace $\ln[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ in estimation (12) with the definition from (14) generating an estimation that measures the effect of each characteristic separately,

$$M_{hjt} = \beta_0 + \beta_y \ln y_j + \sum_{i=1}^{n} \beta_i \left[\hat{\alpha}_i \ln\left(\frac{z_{iht}}{\bar{z}_{it}}\right)\right] + \sum_{l=1}^{3} \beta_{mi} \ln Market_{ljt} + \beta_t t + \varepsilon_{hjt}. \qquad (15)$$

We can interpret $\hat{\beta}_i$ on individual characteristics the same way as $\hat{\beta}_Q$. Rewriting (13) in terms of individual characteristics we have,

$$\ln\left(\frac{p_{hjt}}{c\left(w_{hjt}, y_{hjt}, \bar{z}_{jt}\right)}\right) = \sum\nolimits_{i=1}^{n} \beta_i \left[\, \hat{\alpha}_i \ln\left(\frac{z_{iht}}{\bar{z}_{it}}\right)\right]. \tag{16}$$

The expression $\hat{\alpha}_i \ln\left(\frac{z_{iht}}{\bar{z}_{it}}\right)$ in (16) can understood as the percentage change in minimum cost over minimum core cost, $\frac{c(w_{hjt}, y_{hjt}, z_{hjt})}{c(w_{hjt}, y_{hjt}, \bar{z}_{jt})}$, that occurs with a change in characteristic z_i. Variable $\hat{\alpha}_i$ comes from the cost function estimation (6) and relates changes in characteristics to change in total cost. In this way, similar to (13) the denominator on the left-hand and right-hand sides are effectively equal, meaning that the percentage change in z_i multiplied by $\hat{\alpha}_i$ is directly reflected as a percentage change in p_{hjt}. When $\hat{\beta}_i > c\left(w_{hjt}, y_{hjt}, z_{hjt}\right)/p_{hjt}$ that percentage change again results in a price increase that exceeds cost increase.

We estimate regressions (7, 8, 9, 12 and 15) first for the entire period 2002 – 2016, then separately for each annual period between 2002 and 2016. This provides fifteen separate sets of coefficient results for each estimation. Comparing results, explanatory power and stability will allow us to comment on our research questions.

2.5. The Data

We sourced all data for this study from the Bureau of Transportation Statistics (BTS), an independent statistical agency within the US Department of Transportation (DOT). The data is collected by the BTS via U.S. DOT Form 41 and relevant filings to the Security and Exchange Commission (SEC). The laws behind Form 41 require most passenger and cargo carriers to report financial and operating information to the DOT on a monthly, quarterly or semi-annual basis. We specify the unit of analysis as a route, which is defined as a non-stop, one way or round-trip flight between two airport pairs[30] for a single carrier. Routes are considered bi-directional, meaning a flight from Atlanta (ATL) to Denver (DEN) or the return trip, DEN to ATL comprise a single route. Non-stop flights are those that have no intermediate stop between the airport pairs. Descriptive characteristics of the data, including the mean, standard deviation and the 25th and 75th percentile, for the average route and quarter are outlined in Table I.

(Table I near here)

Fare data is highlighted in the first row of Table I and comes from the Airline Origin and Destination Survey (DB1B). The DB1B is collected quarterly by the BTS and is a 10% sample of all itineraries. Data provided includes information such as the ticketing carrier, operating carrier, origin, destination, fare, miles flown and service class. Fare shown are in 2002 dollars and prices are for a one-way ticket or half the price of a round-trip ticket.

We define a market as the combination of all the routes between two airport pairs. All market characteristics come from the T-100 Domestic Segment table, a 100% census of all non-stop segment data reported by US carriers. At an average of 2.17 competitors per market, many markets are highly concentrated. However, the larger number of airport competitors 8.93, should serve to provide some level of competition and hold down prices.

---

[30] Several cities have airports that are close enough together to be considered substitutes. In those cases, we have combined the airports. The following close-by airports are combined in the results: DFW(Dallas–FortWorth) and DAL (Love Field); LGA (La Guardia), EWR (Newark) and JFK (J. F. Kennedy); AZA (Phoenix-Mesa Gateway) and PHX (Phoenix Sky Harbor); TPA (Tampa) and PIE (St. Petersburg Clearwater); DCA (Reagan) and IAD (Washington Dulles); ORD (O'Hare) and (MDW) Midway

We include three product characteristics for vector *z,* measuring service level, performance and convenience. For performance we use on-time arrival percentage, a measure of the percentage of flights that arrive at their destination within fifteen minutes of scheduled arrival time. On-time arrival is valued by consumers, as reflected in a recent study by Gayle and Yimga (2018) who found that travelers would be willing to pay $1.56 per minute to avoid late arrival. In a study on the cost impact of operational performance Zou and Hansen (2012) found that delays increased cost. However, they also found that "buffering" or padding a schedule to increase on-time performance increases cost as well, indicating that committing more resources can increase on-time performance.

Convenience is measured in terms of flight frequency. For a similar level of volume, more frequent flights can affect cost through reductions in load factors (how full the plane is), aircraft size and landing fees. However, it is an attribute valued by consumers. Both Borenstein (1989) and Douglas and Miller (1974) note it as a product differentiator that increases brand value, particularly to more inelastic consumers who place a higher value on time. We measure flight frequency as the number of daily departures. The product characteristic of service level is measured as the percentage of tickets sold that are first-class or business class. Increasing the percentage of first-class passengers increases unit cost. This can be due both to an increase in the use of resources servicing these passengers directly (food, materials, labor) and indirectly in seating space used in the cabin.

Product characteristic data comes from a number of sources including the DB1B survey, the T-100 and BTS On-Time Performance Database, which includes actual and scheduled arrival times for all non-stop domestic flights by major air carriers. We show statistics for both the observed level and the level of the core product. Recalling that the core product is the lowest level of characteristic available in a market we see that on the average route, the percentage of first-class travelers was 4%, while the core level is 0.00%. This core level of 0.00% reflects the fact that several carriers (e.g. Southwest, JetBlue or Spirit) do not market first or business class tickets. Regarding flight frequency, we see that the average carrier provided almost two more flights per day then carriers providing only the core level of service. Finally, for on-time arrival performance, we see the average for all carriers at 81% on time, while the core level is 78%. All three characteristics exhibit a high degree of variation.

In terms of output (*y*) there have been a variety of measures used in the literature depending on the purpose of the study. Some of the more common have been revenue passenger miles (RPM), available seat miles (ASM), revenue ton miles (RTM) or an aggregate measure combining outputs. We have chosen to use RPM as it matches up best with our secondary use of cost in a hedonic regression. In their study of economies of scale Johnston and Ozment (2013) elect to use ASM noting that a cost function should depend on the total amount of output produced, not just the portion sold. However, differences between seat miles and passenger miles can arise not only from marketing failures, but as an operational decision. Leaving some amount of output unsold can ensure all customers are serviced and allows for the sale of higher priced last-minute tickets. In either case, recovery of costs can only come from seats sold, so we believe it is appropriate to use RPM as our unit of output.

The final section of Table I presents descriptive statistics for input prices *(w)*, similar to fares, all prices are shown in 2002 dollars. For inputs in the cost function we follow a standard KLEM model and include capital, labor, energy and materials, the same inputs used in most previous studies. This data is provided as part of Form 41 in various schedules. Fuel price and labor price are calculated as total expense over total gallons used or number of full-time equivalents per quarter. The price of capital has been measured a number of ways and is very dependent on how capital quantity is measured. We are following a process similar to Färe, Grosskopf, and Sickles (2007) and develop capital price based on the cost of capital and the total number of seats available. The cost of capital comes from two sources, leasing rates and capital depreciation while the number of seats available is based on the number of planes in service and the seat configuration utilized by the carrier. As a proxy for all other materials we use the producer price index (PPI) collected by the US Bureau of Labor Statistics. The PPI varies by quarter, but not by airline. All variable definitions and computations can be found in Appendix 2.1.

2.6. Results

2.6.1 Cost Function

As the focus of this chapter is on hedonic regressions, we limit our discussion of the cost function estimation to the main results. For full results of the estimation, share equations and fitted values see Tables IA and IIA in Appendix 2.2. We find the input price shares for capital, labor, fuel and other materials to be 7.1%, 29.7%, 22.8% and 40.4% respectively. These results are in line with other cost studies cited in this chapter. We use the average route for each carrier for each quarter between 2002 and 2016 to obtain coefficient estimates. The set of carriers included are those carry at least 1% of the domestic US passenger volume and are required to report to the DOT. Over the period studied this included sixteen different airlines. We find cost per passenger mile to be 14.7 cents on average, which is in line with industry reported values. Applying the elasticity coefficients at the observed route level data we find a mean cost of 11.5 cents per mile on the core product and 15.0 cents per mile based on observed characteristics. We also find the average value of the characteristic index $Q(w_h, y_h, z_h, \bar{z})$ to be 0.86, indicating that on average the minimum cost of the core product is 14% less than the minimum cost of characteristics that the carriers are providing.

2.6.2. Hedonic Price Function

In Table II we present the results of estimations (7) and (8). The coefficient on *Distance* is positive, relatively consistent and highly significant for both price estimations, indicating that all else held equal, fare price increases by roughly 0.30% for each 1% increase in miles traveled. These results are similar to the work previously cited (Morrison and Winston, 1995; Borenstein, 1989).

(Table II near here)

Comparing equations (7) and (8) provides some insight into our first research question, can we improve the results of a hedonic regression by controlling for competition and market power. The coefficients on *market passengers* and *route competitors* are both significant and carry the expected negative sign, indicating that fares decrease as the market size or the number of direct competitors increases. The significant positive sign on *airport competitors* was not expected but might be explained by hub and spoke networks. A hub airport is likely to have a larger number of airlines and one of the findings by Borenstein (1989) was that at a hub the dominant airline can charge higher prices than it does throughout their system. In terms of explanatory power (adjusted R2) we see an increase, from 0.49 to 0.57 with the addition of *Market* characteristics. We also examined the effects of a merger or acquisition on price and find that in the four quarters following a merger of two of the airlines on a route, the mean fare price increase is 5.4%. As noted, we do not include this variable in our final model, so it is not in Table II.

However, we note some changes in the coefficients on *characteristics (z)*. Between (7) and (8) *First Class/Business Share* remains consistent, but the other two characteristics have significant changes. After we add controls for market power, the coefficient on both *Flight Frequency* and *On-Time Arrival* change. In addition, the coefficient on *Flight Frequency* is now significant. The change to *Flight Frequency* is probably best explained by the following concept. When a market is not competitive, a characteristic that differentiates a firm from its competitors is not effective. In other words, if the firm is a monopoly, providing more frequent flights doesn't increase their monopoly price. The negative sign on *On-Time Arrival* does not have an economic explanation, but instead is more likely the result of the complex combination of marginal cost and market power that Pakes (2003) references. We explore this further in following sections.

We do not report results for estimation (9) as the cost predictions required to do the estimations were not reliable. The cost elasticities estimated in (6) are only applicable to changes within the range of observations. For example, mean on-time arrival is 80.1%, and generally ranged from 70% to 90%. Predicting the cost of 80.1% on-time arrival requires also predicting the cost of arriving 0% on-time arrival, for which the results were unreliable. In the next section, the cost predictions used are for smaller increments, within the observed range.

2.6.3 Hedonic Mark-Up Function

Moving to the second part of our research question, can we improve the results by accounting for core product cost, we review Table III and estimations (12) and (15). In both estimations the coefficient on *Distance* is statistically significant and negative, indicating that mark-up over core cost reduces as distance increases. This may seem counterintuitive at first but recalling that mark-up is a ratio of price to core cost it makes sense. As distance increases, both core cost and price increase, indicating that the increase in the core cost cannot be completely translated to the price of the product. As a consequence, the resulting ratio value is lower.

(Table III near here)

The 1.085 estimate of $\hat{\beta}_Q$ on $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ is statistically significant and positive, indicating that carriers can increase mark-up over the core cost by differentiating from the core. Recalling that a $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ value of 1 indicates homogeneity with other carriers on the route, an increase in $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ indicates an increase in differentiation.[31]

In estimation (12) the signs on the variables for competition have reversed compared to (8). *Airport Competitors* now carries it´s expected negative sign, indicating that potential competition holds down the mark-up that carriers can expect. The positive and significant coefficient on *Route Competitors* is likely due to the way the dependent variable is constructed. Recall that we are measuring mark-up as price over core cost. For markets that have only a single carrier, the core cost is very much defined by that carrier. It is only in market with multiple competitors that the carrier can differentiate itself. We tested a version of (12) that includes a dummy for merger or acquisition and find that in the four quarters after a merger mark-up over the core cost increases by 9%. All other variables remain as reported.

Estimation (15) is similar to (12) except that $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ has been broken down into the effect of each characteristic separately. The interpretation of coefficients $\hat{\beta}_i$ are the same as discussed above for $\hat{\beta}_Q$. We see that all are significant and positive, indicating that increasing the characteristic can increase mark-up over core cost.

---

[31] An example of this can be constructed at our sample means. Mean fare price is \$174.72, mean cost $c_{hjt}(w_h, y_{hjt}, z_{hjt})$ is \$110.63, and, mean core cost $c_{hjt}(w_h, y_{hjt}, \bar{z}_{jt})$ is \$92.50, and mean mark-up over base cost M_{hjt} is equal to 1.89. At these values, a 10% increase in $c_{hjt}(w_h, y_{hjt}, z_{hjt})$, is an increase of \$11.06 in cost. The 1.085 estimate of $\hat{\beta}_Q$ indicates that the 10% increase in $[Q(w_h, y_h, z_h, \bar{z})]^{-1}$ results in a 10.85% increase in mark-up M_{hjt} to 2.09 which implies a fare price of \$193.79, which is a \$19.07 increase in price.

Mark-up estimations (12) and (15) provide insights that a standard hedonic regression does not, we also see a small increase in explanatory power compared to (7) and (8). Beyond explanatory power, the other benefit we want to explore is stability in coefficient value. To this end, we have run separate estimations of (8) and (14) for each year between 2002 and 2016, presenting the results in Table IV. As consumer tastes and preferences shift, or input prices change, we might expect a change in the coefficients over time. However, beyond shifts over time, coefficients should be relatively stable. To test this, we include the coefficient of variation (CV) of all coefficient estimates in Table IV. The CV in the mark-up estimation is smaller than that of the price regression coefficient for most variables, and for all our variables of primary interest. This would indicate more stability in the mark-up hedonic compared to a standard price hedonic.

(Table IV near here)

This table, and the results from Table III, allow us to answer our second research question as well. Is the effect on mark-up the same for all characteristics? These results clearly show that it is not. In fact, not only do we see differences in the effect on mark-up, but we can see that the effect on some mark-ups have changed over time. Mark-up on *First Class/Business Share* has seen a slight downward trend over time, indicating that the premium over core cost is reducing. *Flight Frequency* has been relatively consistent, averaging very close to one over this period. Another interesting change is the changing effect of *Percentage On-Time Arrival*. Prior to 2012, improving performance increased mark-up, however after 2012 this effect reversed and providing more than the core level of this characteristic reduced mark-up. An explanation for this could be the growth of ultra-low-cost-carriers (ULCC), and the conversion of Frontier Airlines from a low-cost-carrier (LCC) to a ULCC. These carriers placed a lower priority on arriving on time, reducing the cost of the core product.

Our third research question asks if firms can earn a profit providing characteristics above the core product. At a high level, based on the results in Table III we can state that they can increase mark-up over the core cost, however, this is not a clear indication that they are able to earn a price increase greater than the cost increase. Recalling the discussion in the methodology, we analyze the ability to earn additional profit, or increase price over the cost increase, based on the condition $\hat{\beta}_i > c_{hjt}(w_h, y_{hjt}, z_{hjt})/P_{hjt}$. The mean value, weighted by number of passengers, of $c_{hjt}(w_h, y_{hjt}, z_{hjt})/P_{hjt}$ of our sample is 0.77. With the $\hat{\beta}_i$ estimates reported in Table III we can see that at a high level, the expected price increase covers the cost increase.

We find that increasing the characteristic *First Class/Business Share* most consistently earns a price increase greater than the cost increase. Based on the annual estimates presented in Table IV an increase in this characteristic results in a price increase greater than the cost increase in all observations. We do observe a reduction in the level of this profit as the market becomes more competitive in terms of direct competitors, but not to the point where the price premium is less than the cost increase. This result is intuitive as this is the only characteristic included that allows the airline to price discriminate. The other two characteristics benefit all passengers equally and do not allow price discrimination between passengers but do create differentiation between carriers.

The second most consistently profitable characteristic is *Flight Frequency*. Based on comparing the estimates in Table IV to observation level values of $c_{hjt}(w_h, y_{hjt}, z_{hjt})/P_{hjt}$, we observe that price increases resulting from an increase in *Flight Frequency* exceeds the cost increase in approximately 80% of the observations. In contrast to *First Class/Business Share*, having more frequent flights is more profitable in markets where there is more direct competition. As noted above, *Flight Frequency* creates differentiation between carriers. Our results would indicate that this differentiation can be converted to a higher premium in price when there are more competitors present to be compared to. The characteristic that least consistently leads to an increase in profit is *Percentage On-Time Arrival*. In our observations we find that increasing *Percentage On-Time Arrival* only increases price over the cost increase in 60% of the observations. The pattern it follows is similar to that of *Flight Frequency*, but results are less conclusive.

In analyzing profit effects by market characteristics, we observe an interesting effect related to the number of market passengers. Although the number of market passengers has a negative effect on price and on mark-up over core cost, we find that in larger markets carriers are better able to convert characteristic driven cost increases into price increases. In almost all cases, as market size grows, profit earned from an increase in characteristics grows. We believe this is because carriers are better able capitalize on customer willingness to pay when there are other carriers to compare to. As stated before, having more competitors may reduce basic price, but it increases the value of differentiation.

2.7. Conclusion

In this chapter we have extended the index of firm heterogeneity introduced in Chapter 1 to define the concept of a core product and shown how it can be used to improve the results of a hedonic regression. This chapter opens up a new line of research based on the definition of an innovative concept of mark-up. By controlling for the cost of the core product and several market characteristics we generate a hedonic estimation that is more stable over time and can explain a higher percentage of variation than a standard hedonic equation. Applying this to the US domestic airline market we analyze the cost and price effects of product differentiation. We find that differentiating by providing a higher level of service, as measured by first/business class share, and more frequent flights an airline can consistently improve their mark-up over the core product. Furthermore, they can earn a premium over the cost increase due to the higher level of product characteristics. However, improvements in on-time arrival do not consistently result in increased mark-up over the core cost, nor are cost increases fully recouped by price increases.

These findings support those found in the business model literature. Franke (2007) notes that passengers have shown a preference for price and convenience over an extensive network presence, and that carriers offering a premium brand are able to command higher prices. We see both of these effects, with the differentiators of first/business share and flight frequency improving mark-ups. The effect of the ULCC model can be seen in the low to negative returns for improving on-time arrival as carriers identified as ULCC have consistently had the worst on-time performance. Although not analyzed in this chapter, a possible explanation for the low returns to on-time performance is the weakening of the price premium for the FSA hub-and-spoke model. Depending on connection, on-time performance is more important in a hub-and-spoke model than in a point-to-point model. This may argue that the value to consumers is not in arriving on time, but in not missing a connecting flight.

As a tool for firms and pricing practioners this model can assist in making business decisions at the margin, in comparison to products the consumer considers substitutes. It allows the firm to make a more informed and strategic business decision on additions, subtractions or changes to the characteristics of their product. By focusing on the mark-up, rather than on the full price, our hedonic model provides results that are more consistent from year to year, providing decision makers a degree of confidence in the guidance provided. With the inclusion of cost estimates, it also can provide guidance on whether the firm will be able to fully recover cost based on changes to the product.

Our empirical results provide specific strategic implications for airlines and generalizing the results to all firms providing differentiated products. The first strategic implication is that differentiation should not be a one size fits all decision. The rivals and substitute products in a market effect the resulting price and mark-up. When possible, the firm should adjust the service offering and emphasize, or de-emphasize, characteristics by market. For example, in markets with fewer competitors the firm should emphasize the characteristics that allow for price discrimination and differentiation between their own customers and deemphasize those that differentiate themselves from competitors. The opposite is true for more competitive markets. In our empirical example this would imply increasing first and business class seating in markets with few competitors but provide less frequent flights and expand less resources on ensuring on-time arrival.

A further strategic implication is that firms cannot be static or complacent in their choice of product characteristics. As competitors, substitute products, and consumer demands change the firm should adjust their product. This often means adding characteristics, but they should also carefully analyze existing characteristics to ensure they are still creating value. This is best seen in our results over time for on-time-arrival. From 2002 to 2010 the mark-up coefficient was positive, though declining over time. From 2011 onwards it became consistently negative, indicating that the increase in cost of higher on-time-arrival is not covered by an increase in price. The implication would be that consumers are not willing to pay for better on-time arrival. As a group, they may complain about service quality, but as individuals they are not willing to choice the higher fare price required to support that level of quality. This brings up potential policy implications.

The characteristics we have discussed that define a core product have some similarities to public goods in terms of non-rivalry and non-excludability. All consumers of a flight have access to the same on-time arrival and choice of flight frequency, whether or not they value it. In the aviation industry the needs of time-sensitive business travelers demands pushed airlines to meet scheduled arrivals. With all airlines striving to meet the standard of business customers all air travelers benefitted from a certain expected performance. However, with the entry of ULCC and their focus on leisure and discretionary travelers, that core product has been reduced, and some imbalance created. This opens up an area of potential research on the policy need to set some standards for the industry and define a minimally acceptable core product. These steps have already been taken in terms of seat size with legislation passed in 2018 that allows for regulating minimum seat dimensions.

Appendix 2.1. - Variables and Computations

Variable	**Definition**
Output Quantity *y*	Output quantity: Revenue passenger miles in the cost estimation and Stage length of flight, distance between two city pairs in the hedonic regressions
Output price *p*	Output price: Mean ticket price for the carrier on a given route and quarter
Input Quantity *x*	Input Quantities: Labor (Full-Time Equivalents), Gallons of Fuel, Capital (measured as quantity of aircraft seats), All Other Materials (quantity measured as cost divided by a price index)
Input Price *w*	Input Prices: Total expenditures for the input divided by quantity *x*
Characteristics *z*	Product Characteristics: Flight Frequency (number of daily departures), First/Business Class Share (percentage of total tickets sold that are first class or business class), On-Time Arrival (percent of flights that arrive within 15 minutes of scheduled time). All are measured per carrier, route and quarter
Core Characteristics $\bar{z}$	$\bar{z} = \{\overline{z_1}, \dots, \overline{z_n}\}$ and $\overline{z_\imath} = Min\left\{z_{ij}, j = 1,..,k\right\}, i = 1,..,n$
Differentiation Characteristics *d*	$d_{ij} = z_{ij} - \overline{z_\imath}+, i = 1,..,n;\ \ j = 1,..,k$
Market Power Variables	*Market:* The number of passengers on a route flying in either direction, The number of carriers serving the route, Potential Competitors (defined below). All are measured per route and quarter.

Variable	**Definition**
Potential Competitors	Assuming the route is between airports A and B, potential competitors = (number of carriers at A + number of carriers at B)/2
Merger/Acquisition dummy	Value set to 1 where two or more carriers on a route were part of an acquisition for the quarter the acquisition closed and the following three quarters.
Mark-Up over Core Cost	$M_{hjt} = \ln\left(\frac{p_{hjt}}{c_{hjt}\left(w_h, y_{hjt}, \bar{z}_{jt}\right)}\right)$
Differentiation Index	$Q(w_h, y_h, z_h, \bar{z}) = \frac{c(w_h, y_h, \bar{z})}{c(w_h, y_h, z_h)}; \;\; h = 1, .., l$

Appendix 2.2. – Translog Cost Estimation Results

Table IA
Full Estimation of Translog Cost Function and Cost Shares

Eq.	Variable	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Frontier							
	pl	-0.095	0.006	-15.270	0	-0.107	-0.083
	pk	0.494	0.003	190.830	0	0.489	0.499
	pf	2.857	0.006	443.990	0	2.845	2.870
	y	8.485	0.086	98.870	0	8.316	8.653
	plpl2	0.150	0.001	254.400	0	0.149	0.151
	pkpk2	0.055	0.000	334.440	0	0.055	0.056
	pfpf2	0.148	0.000	616.630	0	0.148	0.149
	yy2	-0.419	0.005	-87.150	0	-0.429	-0.410
	ply	-0.036	0.000	-119.780	0	-0.037	-0.036
	pky	-0.032	0.000	-267.690	0	-0.032	-0.031
	pfy	-0.077	0.000	-219.290	0	-0.077	-0.076
	plpk	-0.038	0.000	-185.500	0	-0.038	-0.037
	plpf	-0.103	0.000	-390.930	0	-0.104	-0.103
	pkpf	-0.020	0.000	-191.020	0	-0.020	-0.019
	lt	0.058	0.007	8.900	0	0.045	0.071
	lc	0.381	0.002	173.580	0	0.377	0.386
	lf	0.010	0.000	89.780	0	0.010	0.010
	time	-0.003	0.000	-106.290	0	-0.003	-0.003
	_cons	-69.105	0.765	-90.280	0	-70.605	-67.604
Labor Share							
	pl	0.150	0.001	254.400	0	0.149	0.151
	pk	-0.038	0.000	-185.500	0	-0.038	-0.037
	pf	-0.103	0.000	-390.930	0	-0.104	-0.103
	y	-0.036	0.000	-119.780	0	-0.037	-0.036
	_cons	-0.095	0.006	-15.270	0	-0.107	-0.083
Capital Share							
	pl	-0.038	0.000	-185.500	0	-0.038	-0.037
	pk	0.055	0.000	334.440	0	0.055	0.056
	pf	-0.020	0.000	-191.020	0	-0.020	-0.019
	y	-0.032	0.000	-267.690	0	-0.032	-0.031
	_cons	0.494	0.003	190.830	0	0.489	0.499
Fuel Share							
	pl	-0.103	0.000	-390.930	0	-0.104	-0.103
	pk	-0.020	0.000	-191.020	0	-0.020	-0.019
	pf	0.148	0.000	616.630	0	0.148	0.149
	y	-0.077	0.000	-219.290	0	-0.077	-0.076
	_cons	2.857	0.006	443.990	0	2.845	2.870

Table IIA
Mean Effective Coefficient Across Observations

Variable	Mean	Std. Dev.	Min	Max
Labor	0.30	0.06	0.12	0.43
Capital	0.07	0.03	0.02	0.18
Fuel	0.23	0.07	0.03	0.40
Other Materials	0.40	n.a.	n.a.	n.a.
Output	1.00	0.17	0.65	1.50
RTS	1.03	0.18	0.67	1.53
First/Business Class Share	0.01	0.00	0.01	0.01
Percentage On Time Arrival	0.06	0.00	0.06	0.06
Flight Frequency	0.38	0.00	0.38	0.38

Figures, Tables and Graphs

Table I

Quarterly Descriptive Statistics by Carrier and Route

Statistic	Mean	St. Dev.	Pctl (25)	Pctl (75)
Fares and Market				
Mean Fare	$174.72	$74.99	$121.71	$214.90
RPM (000's)	63,105.22	85,147.32	15,725.43	73,575.28
Distance	1,043.71	719.17	490.00	1,440.00
Route Passengers	65,063.00	69,916.00	21,170.00	82,810.00
Market Passengers	136,346.00	167,471.00	30,475.00	176,333.00
Airport Competitors	8.93	2.54	7.50	10.50
Route Competitors	2.17	0.14	1.00	3.00
Product Characteristic				
Business/First Class	0.04	0.06	0.00	0.07
Flight Frequency	6.49	6.42	2.02	8.30
On Time Arrival	0.81	0.09	0.76	0.87
Core Characteristic				
Business/First Class	0.00	0.00	0.00	0.00
Flight Frequency	4.78	4.67	1.97	6.03
On Time Arrival	0.78	0.10	0.73	0.85
Input Prices				
Labor Price	23,339.34	5,272.13	19,801.80	26,885.70
Fuel Price	2.04	0.84	1.42	2.87
Capital Price	12,767.02	3,750.19	10,200.65	13,838.20
Other Materials Price	147.90	15.08	137.40	160.67

Table II
Cross Section Estimates
Dep Var: Mean Route Fare

	P_{hjt}	P_{hjt}
Estimation Formula:	(7)	(8)
Distance (*y*)	0.285***	0.319***
	(0.001)	(0.001)
Market Passengers (*Market*)		-0.169***
		(0.002)
Airport Competitors (*Market*)		0.089***
		(0.004)
Route Competitors *(Market)*		-0.035***
		(0.002)
First/Business Class Share (*z*)	0.031***	0.032***
	(0.000)	(0.000)
Percentage On Time Arrival (*z*)	0.148***	-0.099***
	(0.010)	(0.009)
Flight Frequency (*z*)	-0.002	0.161***
	(0.001)	(0.002)
Constant	3.503***	3.520***
	(0.012)	(0.012)
Observations	103,980	103,980
Adjusted R2	0.49	0.57

Notes: All regressions include qtr dummies and all variables logged. Std error in parantheses.
* Significant at the 10% level, ** Significant at the 5% level, *** Significant at the 1% level

Table III
Cross Section Estimates
Dep Var: Mark-Up Over Core Cost

Estimation Formula:	M_{hjt} (12)	M_{hjt} (15)
Distance (y)	-0.515*** (0.002)	-0.554*** (0.002)
Market Passengers ($Market$)	-0.300*** (0.002)	-0.319*** (0.002)
Airport Competitors ($Market$)	-0.082*** (0.008)	-0.083*** (0.007)
Route Competitors ($Market$)	0.117*** (0.004)	0.129*** (0.004)
$[Q(w_h, y_h, z_h, \bar{z})]^{-1}$	1.085*** (0.005)	
First/Business Class Share ($\bar{z}/z$)		4.574*** (0.035)
Percentage On Time Arrival ($\bar{z}/z$)		1.073*** (0.301)
Flight Frequency ($\bar{z}/z$)		1.004*** (0.005)
Constant	5.461*** (0.022)	5.538*** (0.021)
Observations	103,980	103,980
Adjusted R2	0.54	0.58

Notes: All regressions include qtr dummies and all variables logged. Std error in parantheses.
* Significant at the 10% level, ** Significant at the 5% level, *** Significant at the 1% level

Table IV

Regressions 8 and 12 Estimated for Annual Periods

Price Hedonic (8)	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	CV
Distance (y)	0.33	0.30	0.28	0.29	0.31	0.33	0.34	0.36	0.36	0.33	0.34	0.31	0.30	0.30	0.30	0.08
Market Passengers $(Market)$	-0.22	-0.16	-0.16	-0.15	-0.15	-0.13	-0.17	-0.15	-0.17	-0.15	-0.15	-0.12	-0.14	-0.15	-0.24	0.19
Airport Competitors $(Market)$	-0.03	0.08	0.01	0.03	0.06	0.11	0.08	0.12	0.16	0.11	0.13	0.14	0.09	0.10	0.18	0.61
Route Competitors $(Market)$	-0.02	-0.07	-0.07	-0.06	-0.05	-0.04	0.00	-0.03	0.00	0.00	-0.03	-0.04	0.00	-0.10	-0.04	0.82
First/Business Class Share (z)	0.06	0.05	0.05	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.34
Percentage On Time Arrival (z)	0.09	-0.39	0.18	0.08	-0.23	-0.27	-0.04	-0.25	-0.32	-0.31	-0.01	-0.10	-0.08	0.47	0.52	6.40
Flight Frequency (z)	0.21	0.16	0.17	0.16	0.13	0.11	0.17	0.15	0.17	0.13	0.12	0.10	0.13	0.17	0.24	0.24
Constant	3.99	3.75	4.07	3.85	3.54	3.27	3.26	2.90	2.80	3.07	3.19	3.28	3.51	3.73	3.76	0.11
Adjusted R2	0.57	0.55	0.56	0.59	0.60	0.60	0.59	0.60	0.61	0.59	0.66	0.63	0.61	0.56	0.55	0.05

Mark-Up Hedonic (12)	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	CV
Distance (y)	-0.69	-0.68	-0.66	-0.62	-0.56	-0.53	-0.52	-0.52	-0.49	-0.50	-0.46	-0.48	-0.48	-0.56	-0.59	0.14
Market Passengers $(Market)$	-0.42	-0.39	-0.34	-0.31	-0.31	-0.31	-0.29	-0.33	-0.31	-0.30	-0.30	-0.30	-0.26	-0.29	-0.34	0.13
Airport Competitors $(Market)$	-0.31	-0.09	-0.23	-0.15	-0.10	-0.07	-0.11	0.05	0.10	0.05	-0.03	-0.08	-0.18	-0.15	-0.05	1.21
Route Competitors $(Market)$	0.19	0.16	0.07	0.06	0.06	0.09	0.09	0.09	0.11	0.12	0.10	0.16	0.21	0.18	0.21	0.41
First/Business Class Share $(z/\bar{z})$	5.53	4.95	5.12	5.09	5.17	5.31	4.72	4.06	3.80	3.85	4.25	4.22	4.33	4.81	4.53	0.12
Percentage On Time Arrival $(z/\bar{z})$	2.68	10.75	12.17	7.28	10.58	7.72	5.17	8.32	0.51	4.51	-1.32	-4.68	-4.59	-3.73	-3.94	1.76
Flight Frequency $(z/\bar{z})$	0.98	0.93	1.07	1.05	0.99	0.95	0.97	1.04	1.03	0.99	1.16	1.11	1.08	0.92	0.86	0.08
Constant	7.27	6.68	6.63	6.00	5.50	5.28	5.19	5.08	4.67	4.75	4.76	4.93	5.07	5.80	6.03	0.14
Adjusted R2	0.61	0.61	0.62	0.59	0.55	0.53	0.52	0.53	0.54	0.56	0.58	0.61	0.60	0.60	0.62	0.06

Chapter 3 - The Role of Product Differentiation in Profitability Change: An Analysis of the Domestic US Airline Industry[†]

Abstract

The US Airline industry swung from $31 billion in losses over the eight-year period 2002 to 2009 to $90 billion in profits over the seven-year period 2010 to 2016. This reversal of fortune was not driven by specific airlines but can be observed industry wide. To fully explore this change we decompose annual profitability change over the fifteen-year period and analyze the economic drivers, with a focus on the driver of product differentiation. The inclusion of product differentiation, an economic driver not typically included in the analysis of financial performance, is one of the primary contributions of this chapter. We introduce a novel method of decomposing profitability change and use the combination of a standard cost function paired with a non-standard revenue function. To fully understand this effect, we examine these financial performance measures for twenty individual carriers in the US airline industry between 2002 and 2016.

[†] This chapter was developed in collaboration with David Saal during my stay at Loughborough University.

3.1. Introduction

The US Airline industry swung from $31 billion in losses over the eight-year period 2002 to 2009 to $90 billion in profits over the following seven-year period 2010 to 2016. Furthermore, average profitability, defined as revenue over cost, rose from 0.98 in the first period to 1.10 in the second period. This reversal of fortune was not driven by specific airlines or segments but was observable industry wide. A handful of studies have looked at this shift and put forward some possibilities: Hazel (2018) points to an increase in capacity discipline, Barrows (2018) to recent mergers and airfare increases, and Mccartney (2018) attributes the profitability increase to baggage and other fees. The worldwide aviation industry has improved over this period as well, but not to the degree of the US industry, further motivating us to understand the change drivers in the US market. To explore this change in depth we decompose annual profitability change over the fifteen-year period and explore the economic drivers, with a focus on the driver of product differentiation. We analyze profitability change, rather than profit, as it allows for easier comparison between firms and periods of varying size.

There are academic studies that analyze profitability change in the airline sector (see Oum and Yu (1998); Färe, Grosskopf, and Sickles (2007); Scotti and Volta (2017)) and industry studies by the International Air Transport Association (IATA) and others. However, the effects of product differentiation are not accounted for in these studies, nor in the wider profitability change literature analyzing other industries. Therefore, one of the primary contributions of this chapter is the inclusion of product differentiation as an economic driver not typically included in the analysis of financial performance.

Product differentiation is usually considered from the side of revenue as a method of allowing for price discrimination. However, product differentiation also has an effect on cost. In this study we begin on the cost side and decompose cost change into its component drivers, product differentiation change and all other effects. While the other effects have received much attention in the literature, product differentiation is relatively unexplored. To measure this driver, we introduce a Konüs type index number to isolate the effect of product differentiation on cost and cost change. Since product differentiation can also affect revenue, we use a non-standard revenue function and decompose revenue change in the same way as cost. Combining the non-standard revenue function with a cost function is a primary contribution of this article.

Cost efficiency can be defined as the ratio of observed cost to minimum cost, given industry best practices. However, when measuring cost efficiency, the implicit assumption in standard analysis is that output is homogeneous between firms. When we consider characteristics that describe how the output is provided, this assumption is frequently not the case. In the airline example there is heterogeneity in characteristics such as on-time performance, density of seating or level of service. Producing output with higher levels of these characteristics can require more inputs at a higher cost. This means that a cost efficiency measure that excludes the intensity of characteristics will be biased for firms with a higher level of them. To explore this, we measure cost efficiency using a cost function that accounts for the level of product differentiation. In current efficiency literature there is often a recognition of heterogeneity in terms of the operating environment, but this does not extend to product differentiation. In the next section we expand on the difference between operating characteristics, which define the operating environment, and characteristics which define the product.

Moving to cost change over time we come to a second implicit assumption of standard analysis, that output is homogeneous between periods. Cost improvements are attributed to drivers such as input price reduction, better technology or better management methods, not changes in the product. However, in the airline industry, as with many other industries, the product changes over time. This raises our second research question, what have been the drivers of cost change over the observation period? How does the effect of change in product differentiation compare to other well-known effects such as productivity? Our third research question asks the same questions as the second but moves the focus to the side of revenue. With the results of this we move naturally to profitability change and our fourth question. Combining cost and revenue change, what have been the drivers of profitability change and how important is product differentiation relative to other effects?

We make three primary contributions in this chapter. First is the inclusion of product differentiation to the financial performance measures of cost efficiency, cost change and profitability change. These attributes vary between airlines and over time, and affect production cost, accounting for them in our analysis of these measures paints a clearer picture. Second, combining a cost function and revenue function together is a contribution and allows us to explain profitability change in a novel way. Finally, while US airline profitability has been the subject of prior research, none has focused on this recent period of prosperity. To fully understand this effect, we examine these financial performance measures for twenty individual carriers in the domestic US airline industry between 2002 and 2016, presenting aggregate results for the entire period and the two sub-periods, 2002-2009 and 2010-2016 to ease comparison.

In the next section we provide a brief review of the literature that forms the structure of the following sections. Our methodology is covered in section 3, the data and empirical application in section 4, results are presented in section 5 and conclusions in section 6.

3.2. Related Literature

In the introduction we note the current literature includes heterogeneity in the form of the operating environment. Expanding on this, Coelli, Perelman, and Romano (1999) research the question of whether to include characteristics that measure the environment as factors that shape the technology, or as factors that influence the degree of technical efficiency. They test these two alternatives using a parametric Stochastic Frontier Analysis (SFA) and three characteristics, average stage length, aircraft size and load factor. They find that the results provide similar rankings of airlines, but different degrees of technical efficiency.

In a more recent article, Coelli et al. (2013) take a different approach and include a characteristic as an input. In their study of electricity distribution, they consider the quality dimension of continuity of supply as an imperfect substitute for maintenance labor and capital allowing them to find the shadow price of quality improvement. Finally, in another article from the same year, Galán, Veiga, and Wiper (2014) include characteristics as a form of unobserved firm heterogeneity. Testing their model on airline data they find that including the factors of average stage length, points served, and load factor improves predictive performance.

In the works cited above, the characteristics included are what we would consider operating or environmental characteristics and not differentiating characteristics. The average stage length of a flight can effect firm cost, but from a consumer point of view it does not differentiate products between firms. In this article we focus on factors that allow firms to differentiate their product from others in the market. As an example, all carriers providing service between two city-pairs fly the same distance, but a carrier can differentiate it's product by having more frequent flights, or better on time performance. To prevent confusion, we follow Ray and Mukherjee (1996) who define these characteristics as "output attributes". With output attributes we also focus on differentiation between firms, rather than within the firm. The attribute is provided equally to all passengers. While differentiation between customers within a firm is an interesting question, the cost method we employ only allows us to measure differentiate between customers where it creates a cost difference. We contribute to the previous literature by our inclusion of output attribute as characteristics beyond the standard of just operating characteristics.

A number of papers have investigated the correlation between airline service levels and profitability, but not with the comprehensive method we are applying. In a non-parametric analysis, Merkert and Pearson (2015) develop an efficiency measure that incorporates perceived service levels, output and profitability. In a first stage test they find no correlation between service level and profitability. In a second stage they find that only crew size has an effect on their combined efficiency measure. In another study of quality and profitability, Kalemba and Campa-Planas (2017) come to a different conclusion. Combining four different measures of quality they find a positive relationship between quality and return on investment (ROI). However, they find no relationship between quality and revenues.

Turning the question around, Mellat-Parast et al. (2015) analyze the relationship between service failures and profitability. They also include the competitive strategy of the airline as a factor, separating the effects by “focused” and “non-focused” airlines. They find that the strategy moderates the effect of service failures. For example, arrival delays negatively affect profitability for a focused airline, but has inverted U shaped affect for a non-focused airline.

There is also a stream of literature that has studied the chronic lack of profitability in the airline sector. Some, such as Borenstein (2011) point to exogenous demand and cost shocks, while others, Wojahn (2012) points primarily to overinvestment and excess capacity as the primary causes. An interesting point made by Borenstein (2011) is that legacy carriers had turned to network expansion through mergers and alliances to help differentiate their product. The industry association IATA prepares an annual profitability report and commissioned a major study of the 2004 – 2011 period. The study stated the situation clearly, the aviation industry creates significant value for its customers but has found it difficult to make an adequate level of profit. They point to two major factors, excess profit in other sections of the value chain, and unconsolidated industry structure and ease of entry that has resulted in overcapacity. In the more recent 2018 report that highlights an improved situation in North America, consolidation, ancillary charges and low fuel prices are noted as the drivers of profitability improvement. We come to some of the same conclusions in this chapter.

Profitability and profitability change in the airline industry has also been the subject of academic research. In a study that focused on the 1983 to 2010 period Scotti and Volta (2017) examine profitability change in the global airline industry. Applying a Bayesian estimation of a cost function and a total factor productivity (TFP) approach they find increasing efficiency over the period, primarily due to technical change. However, this does not translate to higher profitability as much of this gain is transferred to consumers.

3.3. Methodology

We noted in the introduction that this article initially focuses on the effect of product differentiation on cost change in order to explain changes in profitability, so we begin this section with a cost approach. However, defining profitability as revenue over cost it is also possible to examine this effect on the side of revenue. Firms engage in product differentiation with the intent of increasing revenue through raising willingness to pay and price. Increasing the level of attributes is a lever that firms use to differentiate, therefore we should be able to capture the effect of attributes on revenue, and ultimately on profitability. In the second subsection we move the focus to decomposing revenue change and in the third subsection we combine them to analyze profitability change.

Our first research question, which explores a measure of efficiency that includes output attributes, could be measured as output efficiency using revenue, or input efficiency through cost. However, since airlines have more control over their input level than their output level, the measure of input efficiency through cost better fits our empirical application to answer our first question.

3.3.1. Cost Approach

Beginning with standard notation, we define the vector of output quantities as $y \in \mathbb{R}_{+}^{N}$, output prices as the vector $p \in \mathbb{R}_{++}^{N}$, input quantities as $x \in \mathbb{R}_{+}^{M}$, input prices as $w \in \mathbb{R}_{++}^{M}$, and output attributes as $q \in \mathbb{R}_{+}^{L}$. Ray and Mukherjee (1996) note that in many industries the scalar output produced may have qualitative output attributes, and that the maximum output that can be produced may depend on these attributes. With this in mind, we then define production possibility set T as the set of output quantities with associated attributes that can be produced for a given input set as $T = \{(x, y, q)\colon x\ can\ produce\ y\ with\ attributes\ q\}$. Given technology set T we can define the set of inputs required to produce a given output quantity with a given level of attributes as $L(y, q) = \{x\colon (x, y, q) \in T\}$.

Adding input prices *w*, we move from input quantities to a cost function and define the least expensive set of inputs required to produce a given output quantity and level of attributes. Letting total cost be C = $w^T x$ where T is the transposition of input quantity and price, the minimum cost to produce output quantity *y* with input prices *w* and output attributes *q* can be found as,

$$c(w, y, q) = min_x\{w^T x : x \in L(y, q)\}\,. \qquad (1)$$

From (1) we see that minimum cost will increase with either an increase in output quantity y or an increase in attributes q. We also see that $w^T x \geq c(w, y, q)$, observed cost will be equal to or greater than the minimum cost.

With expression (1) we can explore our first question and generate a measure of cost efficiency that accounts for the level of attributes. Expression $c(w, y, q)$ in (1) defines a cost frontier that is non-decreasing in y and q, concave, and homogeneous to the first degree in w. A measure of cost efficiency would be a function of,

$$CE_c(w, y, q, x) = w^T x / c(w, y, q) \geq 1. \qquad (2)$$

In expression (2), higher levels of inefficiency are associated with higher values of $CE_c(w, y, q, x)$. Figure I depicts a cost function and an observed cost $w^T x'$ associated with the level of attributes q'. Any observations in the shaded area above the curve would have a cost efficiency measure greater than one, as observed cost would be greater than minimum cost. On Figure I, inefficiency can be visually measured as the vertical distance between observed $w^T x'$ and the cost function $c(w, y, q')$.

(Figure I near here)

In Figure I we see that at as attributes q increases, minimum cost $c(w, y, q)$ increases, so if $q < q'$ then $c(w, y, q) < c(w, y, q')$. In the introduction we note that if attributes are not accounted for, and there are differences between firms, cost efficiency for firms that provide a higher level of attributes would be biased upward. For example, in a situation where a firm provides attributes q' where $q < q'$, and the assumed level of attributes is q, we would find that

$w^T x'/c(w,y,q) > w^T x'/c(w,y,q')$ and that efficiency is biased upward for the firm providing higher attributes if they are not accounted for. We see this situation depicted in Figure I where minimum cost $c(w,y,q)$ is less than $c(w,y,q')$, but observed cost $w^T x'$ is driven by the attributes q' not the assumed q.

From (1) we see that product differentiation is a determinant of costs and hence a determinant of cost efficiency as expressed in (2). Moving from static cost efficiency, to the dynamic measure of cost change, the effect of product differentiation then becomes a determinant of changes in cost. By definition, differentiation requires a form of comparison. We modify the Konüs type output attribute index of differentiation introduced in Chapter 1 as $Q(w_h, y_h, q_h^1, q^0) = c(w_h, y_h, q_h^1)/c(w_h, y_h, q^0)$ where q^0 is the minimum level of attributes in the market and q_h^1 the observed level of attributes of firm h. This static index of differentiation can be moved to a dynamic context to explain cost changes due to a change in the level of attributes.

In a dynamic context, we replace q^0 with q^t, which describes the initial situation or period *t*, and q_h^1 with q^{t+1} as the final situation in period *t*+1. In this context a firm moves from an initial level of attributes q^t to a final level q^{t+1}. When the level of attributes does not change: $q^t = q^{t+1}$, then $Q(w_h, y_h, q_h^{t+1}, q_h^t) = 1$. If the level of some attributes at the end of the period is lower: $q^{t+1} \leq q^t$, then $Q(w_h, y_h, q_h^{t+1}, q_h^t) \leq 1$. Finally, when the level of some attributes at the end of the period are higher: $q^{t+1} \geq q^t$, then $Q(w_h, y_h, q_h^{t+1}, q_h^t) \geq 1$. By changing their output attributes, a firm changes its level of differentiation, and potentially its position in the market. However, as shown above, this also affects their cost.

To maintain our focus on attributes, we decompose cost change first into the product differentiation effect, and then into the remaining effects. We follow a method outlined in Grifell-Tatjé and Lovell (2015: 282) and utilize a Konüs approach to first separate cost change into two components, a change in attributes defined by a Konüs type index and its implicit index as,

$$\frac{w^{t+1T}x^{t+1}}{w^{tT}x^{t}} = \frac{c^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{c^{t+1}(w^{t+1},y^{t+1},q^{t})} \times \frac{w^{t+1T}x^{t+1}/c^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{w^{tT}x^{t}/c^{t+1}(w^{t+1},y^{t+1},q^{t})}. \qquad (3)$$

In (3) the first expression on the right defines a Konüs product differentiation index, which measures the change in minimum costs due to a change in attributes holding the technology, output, and prices all equal to the final period *t*+1. The second expression is then an implicit index of all other cost change drivers. Decomposing the Konüs implicit index into its component drivers, our full cost decomposition is,

$$\begin{aligned}\frac{w^{t+1T}x^{t+1}}{w^{tT}x^{t}} = {} & \frac{c^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{c^{t+1}(w^{t+1},y^{t+1},q^{t})} \\ & \times \frac{w^{t+1T}x^{t+1}/c^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{w^{tT}x^{t}/c^{t}(w^{t},y^{t},q^{t})} \times \frac{c^{t+1}(w^{t},y^{t},q^{t})}{c^{t}(w^{t},y^{t},q^{t})} \\ & \times \frac{c^{t+1}(w^{t},y^{t+1},q^{t})}{c^{t+1}(w^{t},y^{t},q^{t})} \\ & \times \frac{c^{t+1}(w^{t+1},y^{t+1},q^{t})}{c^{t+1}(w^{t},y^{t+1},q^{t})}.\end{aligned} \qquad (4)$$

At this point we discuss the interpretation of the expressions in (4) and provide a simplified notation that will be useful in the rest of the chapter.

The first expression on the right in (4) is our measure of attribute driven cost change defined in (3), which we note as $Q_c^{t+1}(w^{t+1},y^{t+1},q^{t+1},q^{t})$. The second expression, cost efficiency change, measures the decrease (increase) in cost from a firm becoming more (less) efficient. This is measured as the change in the ratio of observed cost to minimum cost, in other words, how much closer (further) the firm is to the cost frontier. Following the notation introduced in expression (2) we note this as $\Delta CE_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t) = \frac{w^{t+1T}x^{t+1}/c^{t+1}(w^{t+1}, y^{t+1}, q^{t+1})}{w^{tT}x^t/c^t(w^t, y^t, q^t)}$. The third expression, in the second row, measures technical change, or a shift in the cost frontier. It can be understood as the introduction of new techniques or technology that allow production of the same quantity of output and attributes with less inputs. We note this technical change as $\Delta T_c(w^t, y^t, q^t) = \frac{c^{t+1}(w^t, y^t, q^t)}{c^t(w^t, y^t, q^t)}$.

The fourth expression in (4) is defined as the activity effect and is noted as $A_c^{t+1}(w^t, y^{t+1}, y^t, q^t) = \frac{c^{t+1}(w^t, y^{t+1}, q^t)}{c^{t+1}(w^t, y^t, q^t)}$. The activity effect captures cost increase (decrease) due to higher (lower) provision of product. It measures the cost variation associated with the movement in the level of output from y^t to y^{t+1}. This effect simply measures the increase on cost associated with the inputs needed to produce a higher quantity of output. For example, in the case of constant returns to scale and no other changes, a doubling of output would mean $A_c^{t+1}(w^t, y^{t+1}, y^t, q^t) = 2$, signaling that the cost in period *t+1* is double that of period *t*. At another level, this measure also captures any returns to scale effects and the firm's ability to adjust output mix based on the level of activity.[32] In this context, where the activity effect is analyzed in isolation, productivity change is given by the product of technical efficiency change and technical change $\Delta CE_c \times \Delta T_c$ in the second row of (4). The justification of this approach is the strong expansion in the industry of aviation activity over this period, a justification similar to Brea-Solís, Casadesus-Masanell, and Grifell-Tatjé (2015) who isolate the activity effect for the study of the intensive Walmart expansion

---

[32] For details on this activity effect, and its associated definition of productivity, see Grifell-Tatjé and Lovell (1999; 2015, p. 268-271)

policy since its foundation. The final expression in (4) is a Konüs input price index that measures cost change due to changes in input prices. We note this expression as $W_{kc}^{t+1}(w^{t+1}, w^t, y^{t+1}, q^t) = \frac{c^{t+1}(w^{t+1}, y^{t+1}, q^t)}{c^{t+1}(w^t, y^{t+1}, q^t)}$. Restating (4) with our simplified notation we have the notational form to be used going forward,

$$\begin{aligned} \frac{w^{t+1T}x^{t+1}}{w^{tT}x^t} = \; & Q_c^{t+1}(w^{t+1}, y^{t+1}, q^{t+1}, q^t) \\ & \times \Delta CE_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t) \times \Delta T_c(w^t, y^t, q^t) \\ & \times A_c^{t+1}(w^t, y^{t+1}, y^t, q^t) \\ & \times W_{kc}^{t+1}(w^{t+1}, w^t, y^{t+1}, q^t). \end{aligned} \tag{5}$$

There are two different possible decompositions associated with the Konüs implicit index from expression (3). The decomposition in (4) results in a "mixed" period in the factor measuring cost change due to input price change. Mixed in the sense that output *y* is relative to *t*+1 while attributes *q* are based on period *t*[33]. It is also possible to decompose (3) in a way that the mixed period appears in the measure of activity instead of the measure of change due to input price. As we don't have clear criterion to prefer one alternative over the other, we take the geometric mean of the two and find activity change as, $A_c^{t+1}(w^{t+1}, w^t, y^{t+1}, y^t, q^t) = \left[\frac{c^{t+1}(w^t, y^{t+1}, q^t)}{c^{t+1}(w^t, y^t, q^t)} \times \frac{c^{t+1}(w^{t+1}, y^{t+1}, q^t)}{c^{t+1}(w^{t+1}, y^t, q^t)}\right]^{1/2}$ and

---

[33] In a study of symmetric decompositions Balk and Zofío (2020) raise the topic of mixed periods.

input price change as $W_{kc}^{t+1}(w^{t+1}, w^t, y^{t+1}, y^t, q^t) = \left[\frac{c^{t+1}(w^{t+1}, y^{t+1}, q^t)}{c^{t+1}(w^t, y^{t+1}, q^t)} \times \frac{c^{t+1}(w^{t+1}, y^t, q^t)}{c^{t+1}(w^t, y^t, q^t)}\right]^{1/2}$. This mixed period only affects the factors of input price change and activity change.

Expression (4) is formed based on beginning the decomposition from the point of view of the technology *t*+1, the final period. We could also use the technology of period *t*, the initial period, to define the Konüs product differentiation index and begin the decomposition. In that decomposition attribute change would be $Q_c^t(w^t, y^t, q^{t+1}, q^t) = \frac{c^t(w^t, y^t, q^{t+1})}{c^t(w^t, y^t, q^t)}$. To weight change equally between the initial and final period we take the geometric mean of the indexes. This process is detailed in Appendix 3.1. Following this, the final cost decomposition is,

$$\begin{aligned}
\frac{w^{t+1T}x^{t+1}}{w^{tT}x^t} = {} & Q_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t) \\
& \times \Delta CE_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t) \times \Delta T_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t) \\
& \times A_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t) \\
& \times W_{kc}(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t).
\end{aligned} \tag{6}$$

3.3.2. Revenue Approach

The first expression on the right of (6) captures the change in cost driven by a change in attributes between the initial and final periods. However, our research question asks about the attribute driven changes in profitability resulting from changes in both revenue and cost. To get to this we include an analysis of revenue as the other component of profitability.

In a standard revenue function, firms are thought to maximize profit in a competitive setting by varying output quantities *y* for market determined output price vector *p* and fixed input vector

x. The standard indirect revenue function is given as $r(x, p)$. This is analogous to the cost minimization model discussed in the previous section.

However, as noted earlier, in the airline industry firms have more control over their level of inputs than over their outputs. In fact, during the period under observation there were two exogenous occurrences that significantly changed output levels, the Sept 11th terrorist attacks and the 2008 recession. Regarding output prices, most airline routes in the US are closer to an oligopoly setting than a competitive setting, and airlines have some control over prices[34]. Dynamic pricing systems, which play a large role in the management of every airline, are based on the airlines ability to manage prices. In recognition of this we look beyond the classical revenue function to the nonstandard maximum revenue function (NSRF) introduced by Berger, Humphrey, and Pulley (1996) and extend it to include output attributes. Their work specifies this alternative revenue function for the banking industry; however, we argue that the same conditions they cite to justify use of the NSRF also exist in the airline industry.[35]

The NRSF assumes that the firm has some degree of control over their output prices. Berger et al. (1996) support this by pointing to studies showing price dispersion for loan pricing to the same type of borrower as some proof of this control. They also cite studies showing that when banks have greater local concentration, they pay lower rates to depositors and charge higher rates to borrowers. In similar studies on the airline industry Gerardi and Shapiro (2009) and Dai, Liu, and Serfes (2014) both document higher price levels and higher levels of price dispersion in more concentrated markets, a clear indication of airlines exercising market power.

---

[34] In Chapter 1 we find that 73% of US non-stop routes are served by only one or two carriers.

[35] Use of the NSRF is most common in the banking industry, but it has been applied to sectors outside of banking, for example, railroads Cantos and Maudos (2001) and insurance Cummins et al. (2010)

The other assumption of the NSRF is the exogenous nature of output at the time of decision. Berger et al. (1996) note that outputs funded through deposits can only expand through the growth of the local market or through mergers or acquisitions, offering limited ability for managers to maximize profit by expanding output. Similarly, expansion of output in the airline markets is limited to growth in the city pairs served and airport capacity, leaving the decision maker with limited ability to effect output in the intermediate term.

The NSRF includes input prices as an argument in the revenue function. In the banking industry, Berger et al. (1996) argue that this is because an increase in input prices may provide a signal regarding willingness to pay, and that marking up the cost of funds is one pricing option taken by banks. In the airline industry our inclusion is primarily justified by the pass-through nature of fuel prices, and that these changes are a market level occurrence that signals industry wide price increases. Fuel surcharges, which would show up in output price in our model, are common in the industry. A similar argument can be made for airport charges for use of the facility. In our application this would be captured under the input price of other materials, changes in these charges are passed on to the customer in the form of fees or fare increases. As the NSRF maximizes price, instead of the level of output of the standard revenue function, another argument for including input prices would be the textbook description of the relation of price to marginal cost. The relationship between input prices and cost has already been established in the cost function.

In light of an exogenous output level, input prices, and some ability to alter prices, the NSRF introduced in their work defines revenue maximizing prices, p, as a function of y and w with the indirect revenue function as $p(y, w)^T y = r(y, w)$. To maximize revenue, a firm transforms its output y and input prices w to maximum possible prices p. Extending this, Humphrey and Pulley (1997) add to the model a vector of factors z that influence competitive position and willingness to pay. Combined, these factors form what they term a bank's *pricing* opportunity set for transforming given y, w and z into maximum output prices. This set contains all feasible combinations of output quantities, input prices, and factors z.

In an article that introduces the possibility of inefficiency to the NRF, Restrepo-Tobón and Kumbhakar (2017) assume that the pricing opportunity set is closed. They then note that a frontier can be defined of the highest feasible price for any given combination of input prices and other factors. Allowing for the existence of inefficiency, a firm may be operating inside that frontier, charging prices that are less than optimal. Since outputs are given, observed output prices that are lower than optimal implies revenue inefficiency. This article goes on to find profit efficiency change as a function of revenue efficiency change or cost efficiency change.

Building on this, we replace vector z with attributes q and formally define a price output frontier as $V = \{(p, w, y, q): p\ can\ be\ achieved\ with\ w\ and\ q\ at\ output\ level\ y\}$. Given V we can define price opportunity set $S(w, y, q) = \{p: (p, w, y, q) \in V\}$ as the set of prices achievable given w, y and q. Now letting total observed revenue be R = $p^T y$, the maximum revenue achievable with quantity y, input prices w and attributes q can be found as,

$$r(w, y, q) = max_p\{p^T y: p \in S(w, y, q)\}\,. \tag{7}$$

It is important to note at this point that the production technology defined by $T = \{(x, y, q): x\ can\ produce\ y\ with\ attributes\ q\}$ and the technology defined by $V = \{(p, w, y, q): p\ can\ be\ achieved\ with\ w\ and\ q\ at\ output\ level\ y\}$ are two distinct and different technology. The first relates to the possibilities in transforming input quantities into output quantities, and given input prices, a minimum cost. The second technology, V, represents the possibilities in transforming input prices and attributes into an output price, and given output quantities, a maximum revenue. It is this difference in technology that creates the distinction between (1) and (7) which are both functions of *w, y* and *q*.

These distinct technologies are referenced by Restrepo-Tobón and Kumbhakar (2017) in their study decomposing profit efficiency using the NSRF. In that work they point to five fundamental sources of profit efficiency, the first two being revenue and cost efficiency, the second two shifts of the revenue and cost frontiers, and finally a shift in the profitability frontier as a combination of the second two sources. In addition, they state that by explicitly modelling output price efficiency, which effects revenues, and input efficiency, which affects costs, we can interpret profit efficiency as an overall measure of cost and revenue efficiencies.

To facilitate understanding of the NSRF, and how output attributes *q* affect revenue, we have provided Figure II. This figure is conceptually similar to Figure I but depicts a revenue function rather than a cost function. In this figure revenue $p^T y$ is on the vertical axis, the level of output attributes *q* on the horizontal, and the revenue frontier is *r(w,y,q)*¸ All combinations within the shaded area are feasible in the technology, but only those that are on the curve would be considered efficient. Recall that in this revenue function we maximize revenue by maximizing price. As such, revenue level $p'^T y$ at attribute level q' would be an inefficient point because given all other factors, the firm could be charging a higher price. From this we can also see that higher levels of *q* allow for greater revenue through a higher price. In this figure the effects of technical change would be seen as a shift upward of the curve *r(w,y,q)*.

(*Figure II near here)*

Following the same method described in decomposing cost change, we can separate revenue change between initial period *t* and final period *t*+1. Maintaining the focus on output attributes, we begin with an index of revenue change based on change in output attributes as,

$$\frac{p^{t+1T}y^{t+1}}{p^{tT}y^{t}} = \frac{r^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{r^{t+1}(w^{t+1},y^{t+1},q^{t})} \times \frac{p^{t+1T}y^{t+1}/r^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{p^{tT}y^{t}/r^{t+1}(w^{t+1},y^{t+1},q^{t})}. \quad (8)$$

In expression (8) the first index on the right measures the change in maximum revenue due to a change in attributes holding the technology, output, and prices all equal to the final period. As with the cost decomposition, the second expression is then a Konüs implicit index of all other revenue change drivers. The implicit index in expression (8) can be decomposed in the same way the cost function was in (4). Maintaining the same notation as (6), with a change of subscript, our full revenue decomposition is,

$$\begin{aligned}\frac{p^{t+1T}y^{t+1}}{p^{tT}y^{t}} = {} & Q_r(w^{t+1},w^{t},y^{t+1},y^{t},q^{t+1},q^{t}) \\ & \times \Delta CE_r(w^{t+1},w^{t},y^{t+1},y^{t},q^{t+1},q^{t},x^{t+1},x^{t}) \times \Delta T_r(w^{t+1},w^{t},y^{t+1},y^{t},q^{t+1},q^{t}) \\ & \times r(w^{t+1},w^{t},y^{t+1},y^{t},q^{t+1},q^{t}) \\ & \times W_{kr}(w^{t+1},w^{t},y^{t+1},y^{t},q^{t+1},q^{t}).\end{aligned} \quad (9)$$

We include a full definition of each of the indexes in Appendix 3.1. Interpreting the components of the revenue change decomposition, the first expression on the right of (9) measures change driven by a change in the level of attributes. In terms of the price output frontier, this can be thought of as a higher level of attributes allowing a greater maximum price in the set. The second expression measures a change in revenue efficiency, here a value greater than one indicates that the firm's prices are closer to the optimal prices in period *t*+1 compared to period *t* and the third measures shifts in the frontier that alter the maximum revenue achievable for a given *w, y* and *q*.

This could arise from factors such as changes in consumer sentiment or the level of market power. The fourth expression is a measure of how changes in input prices *w* change revenue and the final a measure of how operating at a different activity level changes revenue.

3.3.3. Profitability

With the decompositions of change in revenue and change in cost we are ready to move to our third research question and combine cost change and revenue change into a measure of profitability change. Combining a cost function and revenue function into a profitability decomposition, is not standard. However, as noted previously, these two functions are based on two distinct and different technologies, which allows for a profitability decomposition that includes both functions but does not duplicate technology related effects. In their article measuring profit efficiency Restrepo-Tobón and Kumbhakar (2017) maximize profit as the difference between the NSRF and the standard cost function. They note that this is not possible in the standard neoclassical profit function which does not separate additively between the revenue and cost function.

Defining profitability as $\Pi = R/C$ we find initial period to final period profitability change as $\Pi^{t+1}/\Pi^{t} = (R^{t+1}/C^{t+1})/(R^{t}/C^{t})$ and $\Pi^{t+1}/\Pi^{t} = (R^{t+1}/R^{t})/(C^{t+1}/C^{t})$, which can be restated as,

$$\Pi^{t+1}/\Pi^{t} = \frac{p^{t+1T}y^{t+1}}{p^{tT}y^{t}} \Big/ \frac{w^{t+1T}x^{t+1}}{w^{tT}x^{t}}. \tag{10}$$

Replacing the final expression on the right in (10) with expressions (6) and (9), and shortening the notation, we decompose profitability change as,

$$\frac{p^{t+1T}y^{t+1}}{p^{tT}y^{t}}\Big/\frac{w^{t+1T}x^{t+1}}{w^{tT}x^{t}} = \frac{Q_r}{Q_c}\times\left[(\Delta RE\times\Delta CE^{-1})\times\left(\Delta T_r\times\Delta T_c{}^{-1}\right)\right]\times\frac{A_r}{A_c}\times\frac{W_r}{W_c}. \tag{11}$$

Interpreting (11) we understand the first expression on the right as the change in profitability that is driven by a change in output attributes. We define this index as $QPI(w^t, w^{t+1}, y^t, y^{t+1}, q^t, q^{t+1})$ or the attribute profitability index. The numerator captures the increase in maximum revenue achievable with the change in attributes, while the denominator measures the associated change in cost. When $QPI(w^t, w^{t+1}, y^t, y^{t+1}, q^t, q^{t+1}) > 1$ the change in output attributes contributes positively to profitability, $QPI(w^t, w^{t+1}, y^t, y^{t+1}, q^t, q^{t+1}) = 1$ would indicate that the increase in cost has been matched by the increase in revenue and when $QPI(w^t, w^{t+1}, y^t, y^{t+1}, q^t, q^{t+1}) < 1$ the variation in cost associated with the decision to change attributes is not compensated by the change in revenue.

The effects for the next four change expressions, revenue efficiency, cost efficiency, revenue technology and cost technology are presented as separate drivers that in combination measure the overall effect of productivity change on profitability. As the revenue and cost functions are based on different technologies there is no reason to believe that they should move together, or that one affects the other. A firm may be improving cost efficiency but be unchanged in revenue efficiency. Presenting the measure only as the results of the ratio would obscure that distinction. The interpretation of these four expressions in (11) thus are the same as the interpretation of their counterpart in (6) and (9), although cost changes are presented as the inverse, as an increase in cost reduce profitability.

The next expression in (11) defines how operating at a different level of activity affects profitability. It can be thought of as the contribution of “scale economies” to profitability with the numerator measuring the change in revenue associated with the move from y^t to y^{t+1} and the associated cost for the same change in output quantity in the denominator. When the value of $A_r/A_c > 1$ the change in output quantity contributes positively to profitability, when $A_r/A_c = 1$ the increase in cost has been matched by the increase in revenue and when $A_r/A_c < 1$ the variation in cost associated with output quantity change is not compensated by the revenue change. The final expression on the right measures the effect of a change in input prices on profitability. The denominator indicates how this affects minimum cost, and the numerator how the input price change is translated to output prices. The resulting value can be interpreted in a similar way to the description of activity effects.

Our inclusion of the NSRF to measure revenue change allows for measures of profitability change not typically captured and is one of the primary contributions of this chapter. A standard decomposition of profitability change usually contains a combination of input and output price change factors and a factor that measures productivity change. For example, Scotti and Volta (2017) explain profitability change in the airline industry by output price change, output quantity change, input price change and total factor productivity change.[36]

3.4. Data and Method

3.4.1. Data

The Bureau of Transportation Statistics (BTS), an independent statistical agency within the US Department of Transportation (DOT), is the source of all the data for this article. The data used is collected through U.S. DOT Form 41 and associated schedules. By law, most passenger and cargo

---

[36] See Grifell-Tatjé and Lovell (2015, Chap 2 and 3) for a review and discussion of the concept of profitability and its decompositions.

carriers are required to report operational and financial information to the DOT monthly, quarterly, or semi-annually.

We analyze annual operating profitability change by carrier for the years 2002 to 2016. We choose annual rather than quarterly profitability to reduce variations that are purely due to seasonality, even though it does reduce the potential number of observations. We have only included domestic US operations, excluding any international flights. This allows for a more equal comparison of firms and product as the average stage length of an international flight is much longer. Over the period under observation we have 195 observations of profitability change, representing sixteen different carriers, descriptive statistics are presented in Table I.

(*Table I near here)*

To highlight the disparity in profitability between the periods we present first profitability and profitability change for the entire period, then for the two periods separately. As can be seen, profitability in the latter period is 11.2 percentage points higher, moving from no profit to gain. At 1.029, profitability change was greater as well, and in the percentiles, we can see that the move is industrywide. In terms of firm size, operating revenue and expense reveal a wide range of size. The maximum revenue of ~$19 billion was recorded by Delta in 2016 and the minimum of $367 million by Spirit Airlines in 2002.

To estimate the cost and revenue functions we use a standard model, with capital, labor, fuel and other materials as inputs, and passenger miles and freight transported as outputs. The inputs of labor and fuel are relatively straightforward to calculate. We use the number of full-time equivalents as the quantity of labor, and total salaries and benefits divided by labor quantity as the price. As we analyze operating profitability, both the quantity and salary of general management has been removed from the labor variable. Similarly, total domestic gallons as the quantity of fuel, with total domestic fuel cost divided by gallons as the price. Since fuel is a significant share of airline total cost, and has the most volatile price over time, we provide some historic detail on this input. In Graph I we can see that fuel prices were steadily growing over the entire first period 2002 – 2008, rising from under $1 a gallon to peak at over $4 a gallon in mid- 2008. In the second period, we can observe three separate trends, a growth period, followed by a period of stability, and finally a steep decline. Due to these price swings, fuel as a share of total operating expense has ranged from a low of 10% to a high of 35% measured as the annual industry average.

(*Graph I near here*)

For the measure of capital we follow Färe et al. (2007) and define capital quantity as the number of seats available based on planes in service and configuration. The price of capital is found as actual firm leasing costs and depreciation divided by capital quantity. The final input, other materials is found as total operating costs less all other identified inputs. Quantity is calculated by deflating the total by the bureau of labor statistics (BLS) producer price index of air transport activities and the producer price index defines the associated price.

With our focus on product differentiation we use a primary measure of output directly related to product choice, revenue passenger miles (RPM). The second output is freight and mail carried, measured as ton miles. Passenger miles is by far the most important output, generating over 98% of revenue on average, a number of carriers do not carry freight or mail. Recognizing that airline revenue increasingly comes from sources other than the base fare, we find the price per RPM as all operating revenue not identified as freight or mail, divided by the quantity of RPM. This measure then includes any ancillary charges for services such as baggage fees, cancellation fees, food sales, among others. In short, all revenue derived from transporting passengers. The price for the second output is found as freight and mail revenue divided by the quantity.

We use four output attributes that can differentiate the product provided. The first of these is first class share, which as measured as the percentage of tickets sold that are in first or business class. It serves as a measure of overall service level and is sourced from the DB1B, a 10% sample of all tickets sold. The second attribute, frequency of flight, is a measure of convenience and has been noted by Borenstein (1989) and Douglas and Miller (1974) as a product differentiator that increases the value of the brand to consumers. Since flight frequency is only relevant at the point of departure, we measure it as an average of departures per day, per airport-to-airport route.

The last two attributes, density and load factor, can be thought of as a physical attribute of a service product. Density measures the number of seats per plane and load factor measures how many of the seats were sold. Over the period under observation, density increased from an average 151 seats per plane to almost 170 seats. This corresponds with reports that show that seat pitch, the distance between seat rows, has declined from 89 to 79 cm, and that seat width has dropped from 46 to 43 cm over roughly the same period[37]. These all indicate that the amount of space allocated to a passenger has seen a 10% reduction. Density is calculated as the available seat miles divided by aircraft miles. Load factor, calculated as passenger miles divided by available seat miles, is a measure of how full the plane is. Over the period under observation, average load factor has grown from an average of 0.71 to 0.85. From the viewpoint of cost, increases in density and load factor reduce cost per passenger, spreading fixed cost over a higher output quantity. From the standpoint of the passenger, increases in either lessens the product, meaning less space, more delay in boarding and disembarking, and a higher likelihood of denied boarding.

---

[37] Marchitelli, Rosa (30 May 2016). "Air Canada passenger suffers 'horrible pain' after being stuck in cramped seat", www.cbc.ca

In the applied portion we orient the variables so that all indicate increases in q. To that end we measure density as the inverse and the load factor as (1-load factor). As can be seen in Figure III, which indexes all measures[38] to 2002 values, all except first class share have drifted downwards over the period. All input and output prices are deflated by the consumer price index.

(*Figure III near here)*

3.4.2. Empirical Method

Our final profitability decomposition in (11) requires estimating minimum cost $c(w, y, q)$ in (1) and maximum revenue $r(w, y, q)$ defined in (7). The methods used to estimate the frontiers needed fall into one of two general classes, parametric and non-parametric techniques. While either of these methods could be used to develop the frontiers we apply the well-known non-parametric data envelopment analysis (DEA) introduced by Charnes, Cooper, and Rhodes (1978). Unlike parametric methods, DEA does not require assumptions on the functional form of the technology. This feature is especially attractive in modelling the effect of attributes on revenue with the NSRF, a relatively unexplored factor with no well-established form.

While DEA does not require an assumption on functional form it does require making an assumption on returns to scale. The two most commonly used are constant returns to scale (CRS) and variable returns to scale (VRS). The difference between the two being whether the increase or decrease in inputs or outputs results in a proportional change in outputs or inputs respectively. Under CRS, all changes are proportional, while under VRS the model assumes increasing, constant and then decreasing returns to scale. Under this assumption the production frontier envelopes the existing observations more closely. However, this creates a problem in the type of cross period

---

[38] Average measures are weighted by passenger miles

modelling we are doing and results in infeasible solutions[39]. To counter this, but not enforce the proportionality assumptions of CRS, we assume the generalized returns to scale (GRS) outlined in Podinovski (2004). Under GRS the technology is allowed to alternate between sections of increasing and decreasing returns to scale. The result being that some observations could improve productivity by increasing, or by decreasing their scale. This form also fits well with our application to the airline industry where there are roughly three clusters of scale; large national network carriers, medium size national carriers and smaller regional carriers. GRS allows each cluster to have an increasing constant and decreasing return portion within an overarching technology.

In DEA, the best practice frontier is created as an envelope of observed input and output sets. This envelope can be created from contemporaneous observations, meaning that only observations from the current period are used, or using a sequential method, where all current and previous observations are used. In this article we employ the sequential method[40]. This implies the absence of technical regress, which is reasonable for this industry. It also creates a wider pool of observations which allows us to include more explanatory variables. To establish a base technology, we include observations beginning from 1998, but only report results for the period 2002 to 2016. To operationalize this in the DEA linear program below we define the 1998 to 2001 period as $t = 0$ and 2002 as $t = 1$, 2003 as $t = 2$, and so on.

To generate the cost and revenue estimations required for (11) we solve the following two DEA linear program problems for each observation of firm h with w, y and q values of the required periods. To implement GRS in DEA we constrain intensity vector λ between a lower bound U and

---

[39] See Ray and Mukherjee (1996) for a full discussion of this issue.

[40] See Tulkens and Vanden Eeckaut (1995) for a discussion of this method and Alam and Sickles (2000) for an application to the airline industry.

an upper bound L. The linear program on the left solves for minimum cost given input price vector *w*, attributes q and output vector *y* by minimizing inputs *x*. The program on the right solves for maximum revenue given input price vector *w*, attributes q and output vector *y* by maximizing output price p.

$$
\begin{array}{ll}
c(w_h, y_h, q_h) = \min_{x,\lambda} w_h^T x & r(w_h, y_h, q_h) = \max_{P,\lambda} y_h^T p \\
s.t. \sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t x_{km}^t \leq x \quad m = 1, \dots, M & s.t. \sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t p_{kn}^t \geq p \quad n = 1, \dots, N \\
\sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t q_{kl}^t \geq q_{hl} \quad l = 1, \dots, L & \sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t w_{km}^t \leq w_{hm} \quad m = 1, \dots, M \\
\sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t y_{kn}^t \geq y_{hl} \quad n = 1, \dots, N & \sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t q_{kl}^t \leq q_{hl} \quad l = 1, \dots, L \\
U \leq \sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t \leq L, \quad \lambda_k \geq 0. & U \leq \sum_{t=0}^{T} \sum_{k=1}^{K} \lambda_k^t \leq L, \quad \lambda_k \geq 0.
\end{array}
\tag{12}
$$

3.5. Results

3.5.1. Cost Efficiency

To answer our first research question, we generate the measure of cost efficiency defined in (2) that includes the effect of product differentiation. Results for this can be seen in Table II summarized by year in Table III summarized by carrier. Recall that the closer this value is to one the higher the level of efficiency. For comparison purposes we have also included the more common measure of cost efficiency as $w^T x / c(w, y)$. As was discussed in the methodology section, and highlighted in Figure I, the measure that includes the effect of output attributes will always be lower than or equal to a measure that excludes there effect. This method of measuring efficiency is new to the literature and is one of our primary contributions. As such, we cannot compare our results to previous studies, however the results for the measure that excludes attributes are similar to those found by Scotti and Volta (2017) in the study of profitability and by Assaf (2009) in a study of technical efficiency in the years where our works overlap.

(Tables II and III here)

Interestingly, the gap between the two measures has declined over time. One possible explanation of this is a reduction in product differentiation. As full-service carriers (FSC) have tried to reduce cost and emulate LCC the differences between carriers has diminished. This explanation is supported by Bitzan and Peoples (2016) who find evidence of cost convergence between LCC and FSC.

3.5.2. Cost Change

In Table IV we present industry change results. Following a format from De Witte and Saal (2010) we show cumulative industry results for the entire period 2002 to 2016 and then cumulative by the two subperiods. The values shown are the geometric average of each airline for the period and are standardized to the 2001 – 2002 change, providing an industry cumulative growth effect. For example, the value of 2.11 in the first column of "Cost Change" indicates that cost more than doubled, growing by 111% over the period. To highlight differences between the two subperiods each is measured separately, with the second subperiod standardized to the 2009 – 2010 change. Overall cost grew by 32% in the first period, and by 41% in the second period. In a similar format we present results by carrier in Table V. To conserve on space, we only present the 2002 to 2016 aggregate. For clear comparisons, only carriers that were active over the entire period are shown.

(Table IV and V here)

Our driver of primary interest, attribute change, adds 5% to industry cost in the first period, but is flat in the second period. Analyzing annual details for each carrier we can observe the impact of attribute change and business strategy. As an example, Alaska Airlines saw a 6% cost increase due to attribute changes. Much of this was driven by a rise in first class share when Alaska was implementing its "Alaska 2010" plan[41]. A plan that had a key goal of maintaining differentiation. This example highlights the discontinuous feature of attribute change in that they are more discrete business strategy driven. In addition, once the change occurs the firm tends to stay at the new level.

Productivity change has reduced industry cost overall by 21%, with a 13% reduction in the first period and 7% in the second. However, examining the drivers of productivity change we observe a change in the effect. In the first period improvements in the technology reducing costs by 25%, meaning that changes in the best practice frontier reflected a lower minimum cost for the same level of output. In contrast, over the same period efficiency change raised cost by 16%. In combination these two would describe an industry where the best performers were pushing the frontier and reducing minimum cost, while other firms were improving, but not quite catching up. In the second period, both these effects are muted, with technical change reducing cost by 8% and efficiency having almost no effect. With the carrier detail in Table V, we can see that in terms of productivity Spirit Airlines was the top performer, reducing cost by 47% over the period, followed closely behind by Alaska Airlines and Southwest. For a number of carriers, American, Delta and United, efficiency change added to cost over the period. One potential explanation for this is the series of mergers that occurred between 2010 and 2015[42] and the associated problems consolidating, or "digesting" the acquisitions.

We can observe that in all periods that change in industry activity level is the most

---

[41] https://aviationstrategy.aero/newsletter/Dec-2003/2/Alaska%3A_the_smallest_Major,_the_biggest_turnaround

[42] In 2005 US Airways merged with America West, In 2009 Northwest and Delta merged; in 2010 Continental joined United; in 2011 Southwest acquired AirTran; in 2013 American and US Airways merged.

important driver in cost, highlighting the expansion of the industry. Comparing the two subperiods we see some significant differences. In the first period total costs rose more than activity change, driven by input price and efficiency change. In the second period we observe the reverse with total costs growing slower than activity. These results are supported by the underlying data which shows average unit cost per passenger mile rising from $0.139 in 2002 to $0.175 in 2008, then declining to $0.136 in 2016.

Although not the focus of our study, we do find some evidence to support the findings of Bitzan and Peoples (2016) who document cost convergence between full-service and low-lost carriers. Looking at the standard deviation of unit cost between carriers we see a 10% drop in standard deviation, indicating a narrowing of difference between them.

3.5.3. Revenue Change

Revenue change is presented in the middle section of Tables IV and V, similar to cost change the largest driver of revenue change is the level of activity. The effect coming from a change in output attributes is stronger on revenue than on cost, with change reducing industry revenue by 24% overall, 15% in the first period and 7% in the second. In the carrier detail in Table V we can observe strategy driven effects. The biggest single change from attributes is Frontier Airlines at a 37% reduction. This can be explained in part by its conversion to an Ultra-Low-Cost-Carrier (ULCC) in 2014 and accompanying reduction in first class seating and frequency of flights. Another example is JetBlue, which positions itself as providing features and benefits not provided by LCCs, but at a lower price than full-service carriers. While this differentiation has increased cost by 2% over the period, it has also raised revenues by 3%.

We can observe in Table IV that productivity plays a larger role in revenue change than it does in cost change with productivity adding to revenue by 76% over the entire period. There are macro level industry effects that can also be seen in the unpresented annual detail[43], and in its decomposed drivers. The first is a productivity decline in 2001 – 2002, the result of a drop in revenue efficiency following the September 11th 2001 terrorist attacks. A similar downturn in revenue efficiency can be seen during the period of the 2007 – 2009 recession. The addition of baggage fees noted by Mccartney (2018) can also be seen in the productivity measure. The effect of baggage fees on total price paid was the subject of research by Brueckner et al. (2015). Their work suggests that while fares declined somewhat with the addition of baggage fees, the overall price rose. We see this effect first in 2006 and 2007 in the technical change component when carriers such as Allegiant and Spirit first introduced baggage fees, and later in 2010 in efficiency change when the practice was widely adopted by almost all carriers.

The effect of the 2009 – 2011 mergers and consolidation can also be seen, first in technical change and then a few years later in efficiency change as airlines consolidated routes and managed capacity. Managing capacity is the subject of research by Hazel (2018), who document that between 2010 to 2014 year over year capacity increases were less than they had been in the past, this effect can be seen in our results as well. In the carrier detail both the fees effect and the consolidation effect can be seen in the productivity driven revenue increases of American, Delta and United.

Overall, the effect of input price change reduced revenue by 20%, by 10% in the first period and 13% in the second. In the first period, 2002 to 2009, it appears that carriers were not able to translate input price driven cost changes into increased prices. In this period input price changes led to a 14% increase in cost, yet a 10% decrease in revenue. In the second period this appears to have changed somewhat and carriers were able to recoup a portion of the input driven cost change.

---

[43] Full annual results are available on request

Beginning in 2012, input prices, primarily driven by fuel, led to several years of input price driven reductions in cost. This was matched by a reduction in revenue, but at a lower rate of change than the cost increase, leading to the profitability growth we analyze in the next section.

3.5.4. Profitability Change

In the previous two subsections we discuss annual cost change and revenue change. In this section we bring the two together to understand the effect on profitability, results are presented in the lower part of Table IV. Over the entire period, profitability grew by 29%, growing from a revenue over expenses profitability of 0.95 in 2002 to 1.23 in 2016. We see that although profitability rose in both periods, the change in the second period is almost double that of the first period. In this table we present the effects of attribute change, activity change and input price change as the ratio of revenue change over cost change. We also present productivity change as a ratio but will discuss this presentation and its separate drivers later in this section.

Concerning output attributes, we see that in both periods attribute change has reduced profitability. This is due to attribute change driving a larger reduction in price than in cost, the effect is most evident in the earlier period at a 20% reduction but moderated in the second period to 7%. An explanation for this might be found in the research by Bitzan and Peoples (2016) who document some convergence of cost of full service carriers and low cost carriers. We also see this in Figure III with the rate of change of output attributes slowing after 2014.

In Table IV we note the combined effect of cost and revenue change due to efficiency and technology as “Productivity”, however this measure cannot be interpreted the same way as classical productivity change. Typically, the contribution of productivity to profitability is thought of as a reduction in required inputs, or an expansion of outputs. In our application, productivity measures two different effects, first the cost change due to a change in the inputs needed to provide a fixed output, and second a change in the ability of the carrier to convert that output into revenue.

We can observe that productivity adds 121% to profitability overall and that the effect is significant in both periods, but weaker in the second period. Digging deeper, we see that in the second period productivity change, and profitability growth, is largely driven by changes in revenue technical change and revenue efficiency. The switch between the importance of revenue technology and cost technology between the two periods also stands out, in the first period cost reductions were the primary technological driver, while in the second revenue technology is much more important. We discuss the policy implications of this in the next section.

Airline expansion over the total period drove a growth in cost of 133% and growth in revenue of 153%, taken together, these changes resulted in a growth in profitability of 9%. This measure, the ratio of activity driven revenue charge over activity driven cost change can be thought of as the contribution of scale economies to profitability. As with the other measures, this value differs between the two subperiods and reflects a contribution to profitability in the first period, but almost no change in the second. There is a fair amount of variation in this measure in the carrier detail. The two airlines that expanded the most, JetBlue and Spirit had contrasting results, with the activity increase adding 6% to profitability growth for Spirit but reducing it by 2% for JetBlue.

At the industry level, input price changes have reduced profitability by 26%, almost all of this in the first period. With most measures of profitability change moderating in the second period, this reduction of loss due to input price change, combined with continued productivity growth, is largely the main profitability driver in the second period.

3.6. Conclusions

In this article we have decomposed profitability change in the domestic US airline industry between 2002 and 2016. We focus on explaining the shift in profitability over the period and on measuring the effects of changes in product differentiation through output attributes. Beginning by introducing a measure of cost efficiency that includes the effect of output attributes we find that average efficiency is 5% higher. In other words, by excluding changes to the product, and differences in the product between firms, the standard measures of efficiency are biased downward by almost 5%. In addition, we find that this bias has reduced over time, an effect most likely driven by a convergence of product and service levels in the industry.

Next, we introduce the effect of output attributes to a measure of cost change and decompose cost change into its component drivers. We find that activity level and productivity have been the two largest drivers of change. We also find that input prices drove significant cost increases in the first part of the period, and cost declines in the second part, a shift that plays a big part in explain the change in profitability. The effect of attribute change on cost change has been a 5% increase in cost over time, almost all of that occurring in the first period.

In order to introduce the effect of attribute change to revenue change we use a non-standard revenue function, a function most commonly used in the banking industry, but applicable to airlines as well. Our application of this function to airlines is on contribution of this chapter. In terms of revenue, we find activity level to be the largest driver, followed by productivity and attribute change. Defining revenue productivity as the carrier’s ability to convert output into revenue we see that this has grown revenue by almost 76% over the entire period. This is likely due to the imposition of baggage and other fees, and an increase in market power following a period of mergers. Output attribute change has led to reductions in revenue, essentially meaning that product changes and degraded service levels have led to reductions in the price level that can be achieved.

Examining how changes in output attributes have affected profitability we introduce an attribute profitability index and we find that changes in attributes account for a 27% reduction in profitability over the period, primarily driven by revenue change. These findings have strategic and business implications. As previously noted, there has been some convergence in costs between low-cost carriers and full-service carriers as the difference between their business models begins to blur. As a result, the product they offer becomes more homogeneous, potentially making it harder to price discriminate. This effect can be seen in a reduction over time of the average passenger miles price, mostly driven by a reduction in maximum prices. For managers, these findings would encourage being cautious of chasing cost reductions through reductions in service levels as they can result in output price reductions that are greater than the cost reductions.

Between the two periods we see a shift in the source of profitability growth from technical change driven cost reductions in the first period to revenue increases in the second. This is very likely due the wave of mergers and consolidations that occurred between 2005 and 2013 and the associated increase in market power. For policymakers and regulators, a healthy airline industry is important, but they also must keep in mind welfare of the consumers. In that case the move from cost reductions to price and revenue increases to grow profitability should be a warning signal to policymakers that the industry may have become too concentrated. Any future mergers, or actions taken to enact barriers to entry should be carefully scrutinized.

Comparing profitability change between the periods we see that the main drivers are revenue productivity and input price change. In the first period, gains from revenue productivity were lost to reductions from changes in input prices and output attributes. In the second period gains from revenue productivity were still positive, and there was no counterbalancing reduction due to input price change, resulting in a sustained period of profitability growth. Combined, these effects do not characterize a healthy industry, but one that is dependent on input price change. Our results show that the industry does raise price enough to cover input drives cost increases and have a tendency to compete away gains when input prices drop. This highlights both a policy and strategy implication. To better handle fuel price changes airlines need to develop better strategies to pass on these input price increases, or strategies to become less vulnerable to input price driven cost swings. A possibility is to cut capacity and more aggressively rise prices when fuel prices rise. This solution however has adverse societal impacts, which brings potential policy implications.

Our results also support the findings of industry groups that report on profitability. A report by IATA in early 2011 that looked back over the previous ten years highlighted a fragmented industry, legal restrictions on cross-border investment and government subsidy of failing airlines as key drivers of perennially low profitability. By 2016 this same groups report on the industry highlighted strong financial performance, at least in North America, with gains underpinned by consolidation. Interestingly, as this report highlights, other areas of the world which have not seen the same kind of industry changes and consolidation are not nearly as profitable. Reporting net post tax profit as a percent of revenue, they show North America at 9.8%, Europe at 3.8%, Asia-Pacific at 3.9% and the Middle East at 1.8%. Latin America and Africa are both recording losses.

Appendix 3.1. – Cost and Revenue Function Decompositions

Cost change decompositions from the final period part of view are developed in the main text. Below we define a decomposition with the initial period as the base.

Attribute Change: $Q_c^t(w^t, y^t, q^{t+1}, q^t) = \frac{c^t(w^t, y^t, q^{t+1})}{c^t(w^t, y^t, q^t)}$

Cost Efficiency: $\Delta CE_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t)$

Technical Change: $\Delta T_c(w^{t+1}, y^{t+1}, q^{t+1}) = \frac{c^{t+1}(w^{t+1}, y^{t+1}, q^{t+1})}{c^t(w^{t+1}, y^{t+1}, q^{t+1})}$

Activity: $A_c^t(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}) = \left[\frac{c^t(w^{t+1}, y^{t+1}, q^{t+1})}{c^t(w^t, y^{t+1}, q^t)} \times \frac{c^t(w^{t+1}, y^t, q^t)}{c^t(w^t, y^t, q^t)}\right]^{1/2}$

Input Prices: $W_{kc}^t(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}) = \left[\frac{c^t(w^{t+1}, y^t, q^{t+1})}{c^t(w^t, y^t, q^{t+1})} \times \frac{c^t(w^{t+1}, y^{t+1}, q^{t+1})}{c^t(w^t, y^{t+1}, q^{t+1})}\right]^{1/2}$

To weight change equally between the initial and final period we use the geometric mean of the initial period and the final period indexes defining,

Attribute Change: $Q_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$

$$= [Q_c^{t+1}(w^{t+1}, y^{t+1}, q^{t+1}, q^t) \times Q_c^t(w^t, y^t, q^{t+1}, q^t)]^{1/2}$$

Cost Efficiency Change: $\Delta CE_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t)$

Technical Change: $\Delta T_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^{t+1})$

$$= [\Delta T_c(w^t, y^t, q^t) \times \Delta T_c(w^{t+1}, y^{t+1}, q^{t+1})]^{1/2}$$

Activity Change: $A_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$

$$= [A_c^{t+1}(w^{t+1}, w^t, y^{t+1}, y^t, q^t) \times A_c^t(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1})]^{1/2}$$

Input Price Change: $W_{kc}(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$

$$= [W_{kc}^{t+1}(w^{t+1}, w^t, y^{t+1}, y^t, q^t) \times W_{kc}^t(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1})]^{1/2},$$

In the text we present the final revenue change decomposition. The component changes are defined below as,

Attribute Change: $Q_r(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$

$$= [Q_r^{t+1}(w^{t+1}, y^{t+1}, q^{t+1}, q^t) \times Q_r^t(w^t, y^t, q^{t+1}, q^t)]^{1/2}$$

Revenue Efficiency Change: $\Delta RE_r(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t)$

$$= \frac{p^{t+1T} y^{t+1} / r^{t+1}(w^{t+1}, y^{t+1}, q^{t+1})}{p^{tT} y^t / r^t(w^t, y^t, q^t)}$$

Technical Change: $\Delta T_R(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$

$$= [\Delta T_r(w^t, y^t, q^t) \times \Delta T_r(w^{t+1}, y^{t+1}, q^{t+1})]^{1/2}$$

Activity Change: $A_r(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$

$$= [A_r^{t+1}(w^{t+1}, w^t, y^{t+1}, y^t, q^t) \times A_r^t(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1})]^{1/2}$$

Input Price Change: $W_{kr}(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$

$$= [W_{kr}^{t+1}(w^{t+1}, w^t, y^{t+1}, y^t, q^t) \times W_{kr}^t(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1})]^{1/2}$$

Appendix 3.2. – Table of Variables and Computations

Variable	**Definition**
Output Quantity *y*	Output quantities: Passenger miles and Freight Ton Miles
Output price *p*	Output prices: Total Revenues for the output divided by quantity *y*
Input Quantity *x*	Input Quantities: Labor (Full-Time Equivalents), Gallons of Fuel, Capital (measured as quantity of aircraft seats), All Other Materials (quantity measured as cost divided by a price index)
Input Price *w*	Input Prices: Total expenditures for the input divided by quantity *x*
Output Attributes *q*	Output Attributes: Density (average seats per plane), Load Factor (passenger seat miles divided by available seat miles), Flight Frequency (average of daily flights per location served), First Class Share (percentage of total tickets sold that are first class)
$Q_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$	$\left[\frac{c^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{c^{t+1}(w^{t+1},y^{t+1},q^t)} \times \frac{c^t(w^t,y^t,q^{t+1})}{c^t(w^t,y^t,q^t)}\right]^{1/2}$
$\Delta CE_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t)$	$\frac{w^{t+1T}x^{t+1}/c^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{w^{tT}x^t/c^t(w^t,y^t,q^t)}$
$\Delta T_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^{t+1})$	$\left[\frac{c^{t+1}(w^t,y^t,q^t)}{c^t(w^t,y^t,q^t)} \times \frac{c^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{c^t(w^{t+1},y^{t+1},q^{t+1})}\right]^{1/2}$

Variable	**Definition**
$A_c(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$	$\left[\frac{c^{t+1}(w^t,y^{t+1},q^t)}{c^{t+1}(w^t,y^t,q^t)} \times \frac{c^{t+1}(w^{t+1},y^{t+1},q^t)}{c^{t+1}(w^{t+1},y^t,q^t)} \times \frac{c^t(w^{t+1},y^{t+1},q^{t+1})}{c^t(w^t,y^{t+1},q^t)} \times \frac{c^t(w^{t+1},y^t,q^t)}{c^t(w^t,y^t,q^t)}\right]^{1/4}$
$W_{kc}(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$	$\left[\frac{c^{t+1}(w^{t+1},y^{t+1},q^t)}{c^{t+1}(w^t,y^{t+1},q^t)} \times \frac{c^{t+1}(w^{t+1},y^t,q^t)}{c^{t+1}(w^t,y^t,q^t)} \times \frac{c^t(w^{t+1},y^t,q^{t+1})}{c^t(w^t,y^t,q^{t+1})} \times \frac{c^t(w^{t+1},y^{t+1},q^{t+1})}{c^t(w^t,y^{t+1},q^{t+1})}\right]^{1/4}$
$Q_r(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$	$\left[\frac{r^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{r^{t+1}(w^{t+1},y^{t+1},q^t)} \times \frac{r^t(w^t,y^t,q^{t+1})}{r^t(w^t,y^t,q^t)}\right]^{1/2}$
$\Delta RE_r(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t, x^{t+1}, x^t)$	$\frac{p^{t+1T}y^{t+1}/r^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{p^{tT}y^t/r^t(w^t,y^t,q^t)}$
$\Delta T_r(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^{t+1})$	$\left[\frac{r^{t+1}(w^t,y^t,q^t)}{r^t(w^t,y^t,q^t)} \times \frac{r^{t+1}(w^{t+1},y^{t+1},q^{t+1})}{r^t(w^{t+1},y^{t+1},q^{t+1})}\right]^{1/2}$
$A_r(w^{t+1}, w^t, y^{t+1}, y^t, q^{t+1}, q^t)$	$\left[\frac{r^{t+1}(w^t,y^{t+1},q^t)}{r^{t+1}(w^t,y^t,q^t)} \times \frac{r^{t+1}(w^{t+1},y^{t+1},q^t)}{r^{t+1}(w^{t+1},y^t,q^t)} \times \frac{r^t(w^{t+1},y^{t+1},q^{t+1})}{r^t(w^t,y^{t+1},q^t)} \times \frac{r^t(w^{t+1},y^t,q^t)}{r^t(w^t,y^t,q^t)}\right]^{1/4}$

Variable	Definition
$W_{kr}(w^{t+1}, w^{t}, y^{t+1}, y^{t}, q^{t+1}, q^{t})$	$\left[\frac{r^{t+1}(w^{t+1},y^{t+1},q^{t})}{r^{t+1}(w^{t},y^{t+1},q^{t})} \times \frac{r^{t+1}(w^{t+1},y^{t},q^{t})}{r^{t+1}(w^{t},y^{t},q^{t})} \times \frac{r^{t}(w^{t+1},y^{t},q^{t+1})}{r^{t}(w^{t},y^{t},q^{t+1})} \times \frac{r^{t}(w^{t+1},y^{t+1},q^{t+1})}{r^{t}(w^{t},y^{t+1},q^{t+1})}\right]^{\frac{1}{4}}$
Π^{t+1}/Π^{t}	$\frac{p^{t+1T}y^{t+1}}{p^{tT}y^{t}} \Big/ \frac{w^{t+1T}x^{t+1}}{w^{tT}x^{t}}$
$QPI(w^{t}, w^{t+1}, y^{t}, y^{t+1}, q^{t}, q^{t+1})$	$Q_{r}(w^{t+1}, w^{t}, y^{t+1}, y^{t}, q^{t+1}, q^{t})/$ $Q_{c}(w^{t+1}, w^{t}, y^{t+1}, y^{t}, q^{t+1}, q^{t})$

Figures, Tables and Graphs

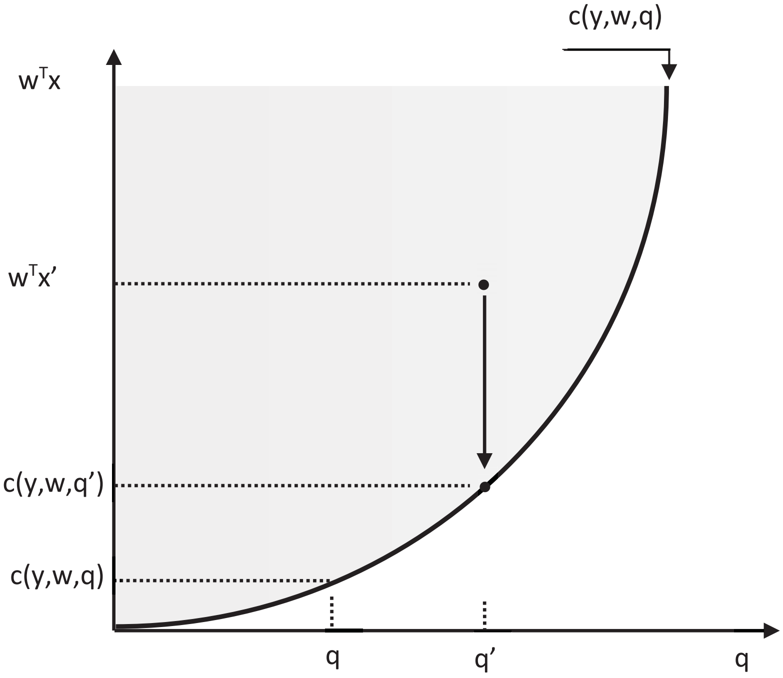


Figure I: Cost Function with Output Attributes

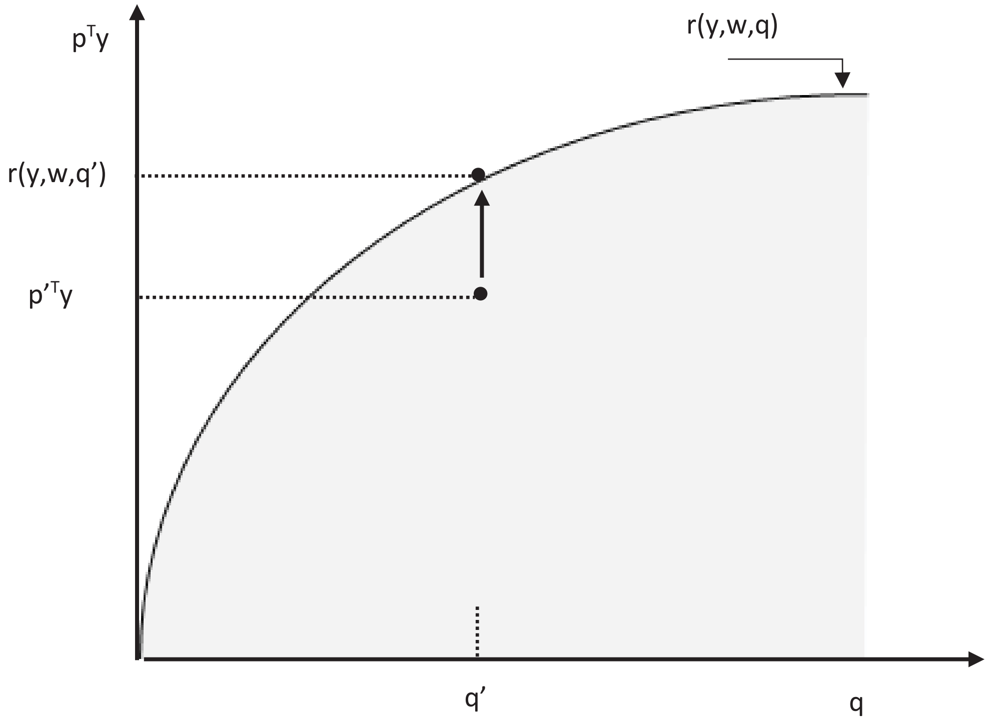
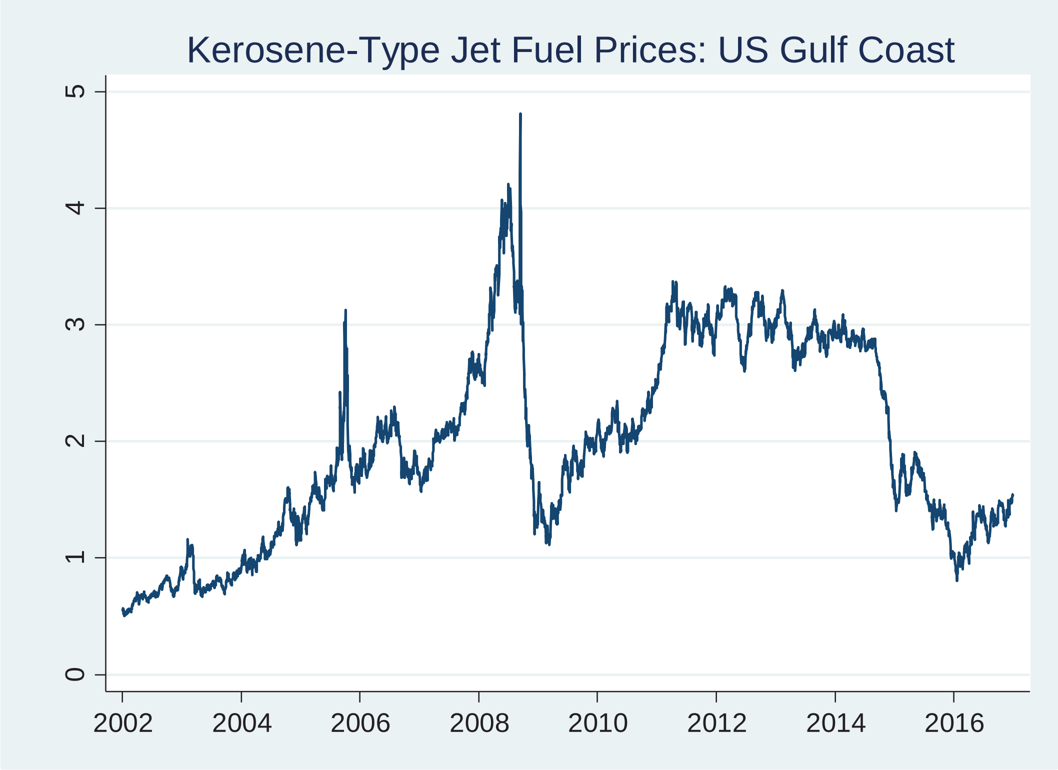


Figure II: Revenue Function with Output Attributes



Graph I: Kerosene-Type Jet Fuel Prices: US Gulf Coast

Table I
Descriptive Statistics

Variable	Mean	Std. Dev	Min	Pctl (25)	Pctl (75)	Max
2002 - 2016						
Profitability	1.056	0.112	0.793	0.986	1.124	1.440
Annual Profitability Change	1.017	0.065	0.819	0.972	1.056	1.239
2002 - 2008						
Profitability	0.998	0.084	0.793	0.933	1.058	1.211
Annual Profitability Change	1.004	0.066	0.819	0.961	1.043	1.176
2009 - 2016						
Profitability	1.112	0.108	0.878	1.036	1.180	1.440
Annual Profitability Change	1.029	0.062	0.915	0.983	1.070	1.239
2002 - 2016						
Operating Rev ($000's)	5,203,751	4,848,301	367,585	978,634	8,600,920	19,187,954
Operating Exp ($000's)	5,034,215	4,605,127	355,092	922,497	8,946,895	17,047,062
Capital Qty (Seats #)	33,919	28,722	2,029	8,200	54,368	109,108
Capital Price ($000's)	11.61	4.44	3.49	8.59	14.39	36.02
Labor Qty (FTE #)	20,874	18,922	642	3,705	32,724	74,821
Labor Price ($)	62.55	14.35	35.96	51.84	70.38	127.40
Fuel Qty (Gal 000's)	724,626	607,382	71,173	171,147	1,248,132	2,312,267
Fuel Price ($)	1.56	0.56	0.59	1.05	2.09	2.70
Other Materials Qty	19,634	19,963	495	2,807	32,692	79,574
Other Materials Price ($)	112	2	110	111	114	115
Passenger Miles (000's)	38,200	31,963	3,361	9,637	64,898	127,868
Passenger Miles Price ($)	121.43	27.21	70.06	100.88	139.86	181.40
Property Tons (000's)	153.43	179.20	0.00	7.16	245.35	662.40
Property Tons Price ($)	442.51	320.65	0.00	272.14	578.01	1,297.28
First Class Share	0.02	0.02	0.00	0.00	0.02	0.12
Density	156.25	27.91	113.35	140.98	161.14	257.38
Load Factor	0.81	0.06	0.59	0.78	0.85	0.92
Flight Frequency	5.25	2.47	0.46	4.01	6.18	14.90

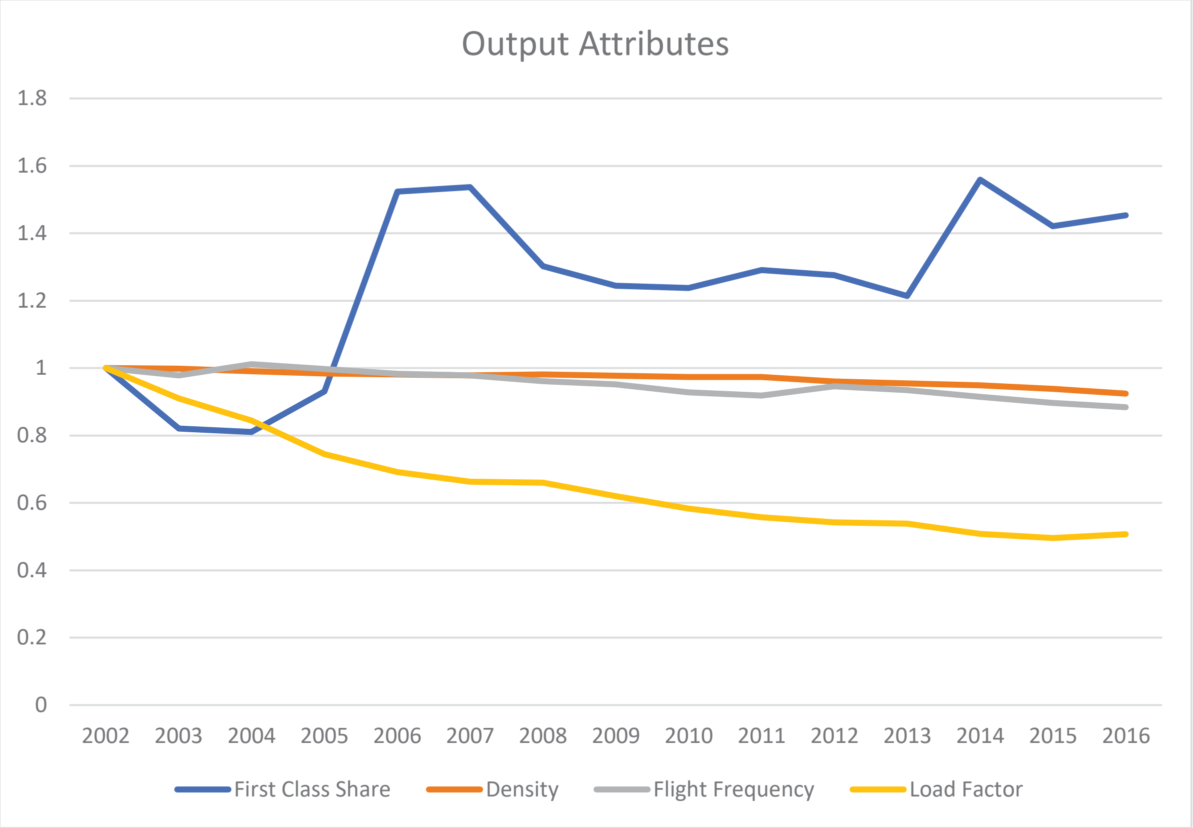


Figure III: Output Attributes Indexed to 2002

Table II
Mean by Year - Cost Efficiency

Year	$w^T x/c(w,y,q)$	$w^T x/c(w,y)$
2002	1.177	1.244
2003	1.255	1.323
2004	1.303	1.370
2005	1.305	1.359
2006	1.313	1.387
2007	1.317	1.376
2008	1.327	1.382
2009	1.341	1.403
2010	1.281	1.341
2011	1.289	1.336
2012	1.260	1.278
2013	1.247	1.261
2014	1.243	1.261
2015	1.245	1.264
2016	1.256	1.276
Total	1.277	1.323

Table III
Mean by Carrier - Cost Efficiency

Carrier	$w^T x/c(w,y,q)$	$w^T x/c(w,y)$
AirTran Airways	1.056	1.483
Alaska Airlines	1.392	1.451
Allegiant Air	1.215	1.217
America West Air	1.394	1.396
American Airline	1.261	1.287
Continental Air	1.584	1.604
Delta Air Lines	1.451	1.454
Frontier Airline	1.384	1.431
Hawaiian Airline	1.018	1.060
JetBlue Airways	1.146	1.148
Northwest Airline	1.165	1.165
Southwest Airline	1.074	1.074
Spirit Air Lines	1.097	1.133
US Airways Inc.	1.913	1.968
United Air Lines	1.424	1.442
Virgin America	1.148	1.238
Total	1.277	1.329

Table IV
Cumulative Growth Effect

	2002 - 2016	2002 - 2009	2010 - 2016
Cost Change	2.11	1.32	1.41
Attributes	1.05	1.05	1.00
Productivity	0.79	0.87	0.93
Efficiency	1.15	1.16	1.01
Technical	0.69	0.75	0.92
Activity	2.33	1.26	1.71
Input Price	1.08	1.14	0.89
Revenue Change	2.72	1.44	1.64
Attributes	0.76	0.85	0.93
Productivity	1.76	1.36	1.20
Efficiency	1.15	1.05	1.04
Technical	1.53	1.30	1.16
Activity	2.53	1.38	1.69
Input Price	0.80	0.90	0.87
Profitability Change	1.29	1.09	1.16
Attributes	0.73	0.80	0.93
Productivity	2.21	1.57	1.29
Efficiency - Cost	0.87	0.86	0.99
Efficiency - Rev	1.15	1.05	1.04
Technical - Cost	1.45	1.34	1.08
Technical - Rev	1.53	1.30	1.16
Activity	1.09	1.09	0.99
Input Price	0.74	0.79	0.98

Table V

Cumulative Growth Effect (2002 - 20016)

	American	Alaska	JetBlue	Delta	Frontier	Hawaiian	Spirit	United	Southwest
Cost Change	1.24	1.79	6.25	1.69	2.11	1.60	3.20	1.27	2.45
Attributes	1.04	1.06	1.02	1.00	0.95	1.10	0.99	1.01	1.00
Productivity	0.86	0.59	0.97	0.99	0.68	0.68	0.53	1.16	0.62
Efficiency	1.25	0.96	1.36	1.36	0.88	1.03	0.80	1.58	0.98
Technical	0.69	0.62	0.71	0.73	0.77	0.66	0.66	0.74	0.63
Activity	1.35	2.71	5.39	1.39	3.61	2.06	4.53	1.33	2.79
Input Price	1.04	1.06	1.17	1.24	0.91	1.02	1.35	0.82	1.42
Revenue Change	1.82	2.45	6.28	2.17	2.66	2.33	3.95	1.83	2.72
Attributes	0.80	0.69	1.03	0.68	0.63	0.80	0.78	0.74	0.82
Productivity	2.03	1.39	1.47	2.64	1.41	1.47	1.35	2.15	1.63
Efficiency	1.51	1.23	1.00	1.33	0.96	1.15	0.85	1.50	1.04
Technical	1.34	1.13	1.47	1.99	1.46	1.28	1.59	1.44	1.57
Activity	1.45	2.57	5.31	1.54	4.38	2.03	4.79	1.43	2.63
Input Price	0.78	0.99	0.78	0.79	0.68	0.97	0.79	0.81	0.78
Profitability Change	1.46	1.37	1.00	1.28	1.26	1.46	1.23	1.44	1.11
Attributes	0.77	0.66	1.01	0.68	0.67	0.73	0.79	0.73	0.82
Productivity	2.37	2.35	1.52	2.66	2.07	2.15	2.54	1.85	2.63
Efficiency - Cost	0.80	1.05	0.74	0.74	1.14	0.97	1.25	0.63	1.02
Efficiency - Rev	1.51	1.23	1.00	1.33	0.96	1.15	0.85	1.50	1.04
Technical - Cost	1.46	1.61	1.40	1.37	1.29	1.51	1.51	1.36	1.58
Technical - Rev	1.34	1.13	1.47	1.99	1.46	1.28	1.59	1.44	1.57
Activity	1.07	0.95	0.98	1.11	1.21	0.98	1.06	1.08	0.94
Input Price	0.75	0.94	0.67	0.64	0.75	0.95	0.58	0.98	0.55

Summary of the Thesis Contributions and Future Lines of Research

Main Findings

In this thesis we have introduced a Konüs type measure of a product differentiation that is based on measuring the difference in cost, or revenue, of a change in the level of attributes. In each chapter we extend the measure, to vary the type of difference being analyzed, and apply it to an economic model, price dispersion in chapter 1, the hedonic model in chapter 2, and profitability decomposition in chapter 3. Our focus is on product differentiation based on differences in the transaction characteristics or the services provided, rather than the more common type based on physical characteristics. Product differentiation of this form is more elusive and challenging to measure in its effect on cost and on price.

The first chapter introduces and defines the methodological contribution. In this chapter the method is applied to measuring the level of market heterogeneity based on output attribute. Creating a scalar measure that can be used to compare and rank markets based on their level of differentiation. As a component of market power, understanding the level of differentiation in a market is essential to understanding the competitive environment, and potentially the need for regulation. This addition to the regulator's toolkit is a further contribution of this chapter. In an empirical application we use the measure of market heterogeneity to contribute to the debate on the effects of competition on price dispersion. We find that the level of market heterogeneity has a significant effect on the relationship between competition and price dispersion, resolving contradictions found in previous research on the subject.

The second chapter extends the methodology to measure the cost of producing the core product in a specific market. The core product is defined as the version of the product with the minimally acceptable level of characteristics, the level absolutely necessary for it to function.

Operationalizing this concept to allow for measurement of core cost is a primary contribution of this chapter. We measure the core cost of the product, then for each firm in the market measure their added cost based on the characteristics they are providing beyond the minimum. This is done for the added cost for all characteristics combined and the added cost for individual characteristics. In an empirical application we use the core cost, plus the cost beyond to the core, to introduce a modified hedonic mark-up model, the second contribution of this chapter. By including the cost of the core product and variation from the core as separate measures, we improve the results of the hedonic model in terms of stability and interpretability. The modified version also provides pricing practitioners better information regarding the effects of a change in differentiation. In the empirical results we find that increases in the characteristics of percentage of first class, and frequency of flight, consistently allow carriers to cover the added cost, while improving on-time performance does not.

In the third chapter we adapt the measure to analyze firm level changes in differentiation over time, rather than between firms at a point in time. We also extend the cost-based measure developed in the first two chapters to a revenue-based measure. This methodological contribution uses an alternative revenue function to measure the change in revenue that is due to a change in output attributes. Combining this revenue change measure with a cost change measure, we introduce a novel profitability decomposition, where profitability is defined as revenues over cost. Since the concept of differentiation is not typically considered in financial measurements, including them is one of the contributions of this chapter. In an empirical analysis we apply this decomposition to a period of significant growth in airline profitability to understand the change drivers. Our empirical application and findings contribute to the literature on airline profitability. We find that over the period of analysis, changes in differentiation have reduced profitability. We also find that the largest driver of profitability change in the US domestic market has been productivity change, due mostly to technical change in cost and revenue.

Future Research Directions

As noted in the introduction, empirical studies of differentiation are relatively rare because of the challenges in definition and measurement. The methodology introduced in this thesis opens up the possibility for further empirical research on differentiation. In addition, any research that uses controls for market structure and market power, such as structure-conduct-performance research, could benefit by including our measure of differentiation.

In the field of productivity and efficiency analysis, the effect of firm heterogeneity is currently an active area of research. The methods developed in this thesis have direct application to that line of research. First, in the empirical estimation of efficiency that includes output heterogeneity, and second, helping decision makers understand the cost implications of changing output attributes. I am currently working on a project where we include these measures as part of a study of the sewage and water industry in Japan.

The core product concept, operationalized in chapter two, can be applied to research on market entry and product positioning and marketing. Redefining what is considered as the core product in the eyes of consumers can be a method of product entry. In many ways, the entry of LCC and ULCC in aviation markets was through redefining the core product. We are currently working on a line of research to examine changes in the core product, and how it can create entry opportunities. Along this idea, we believe the tools developed in this thesis open a research line into operationalizing the product concept developed by Kotler (1967) and depicted in figure 1 of the introduction.

In the first two chapters our focus was on cost-based measures of differentiation, in the third chapter we introduce a revenue-based measure which opens up new areas of research. The cost-based method has a drawback in that it only works for output attributes that create a cost difference, however, a revenue-based index allows measuring differentiation based on changes in the level of revenue generated. Allowing the measurement of a wider range of attributes. These two measures used in conjunction can provide a very clear picture of differentiation in a market. In the profitability decomposition, where we pair the two as a change driver, we are able to measure changes due to technology. We are investigating whether the technology effect can be used as a proxy measure for market power, similar to the way a Lerner index is used. Applying our novel profitability decomposition to other industries of interest is another potential research line.

Final Remarks

I would like to begin my final remarks with a return to the motivation that begin this work. When considering changes, additions or deletions to the sales and service attributes that surround a product, what quantitative methods and measures can guide the firm in making that decision. With that motivation in mind, I believe we have made significant contribution and headway in developing these methods. With ongoing work and research, these methods can be further tested in other industries and developed as a set of practical tools for both management and academic research.

References

Abrate, G. and Viglia, G., 2016, “Strategic and Tactical Price Decisions in Hotel Revenue Management,” *Tourism Management,* 55(August), pp. 123–32.

Ahn, S.C. and Sickles, R.C., 2000, “Estimation of Long-Run Inefficiency Levels: A Dynamic Frontier Approach,” *Econometric Reviews,* 19(4), pp. 461–92.

Alam, I.M. and Sickles, R.C., 2000, “Time Series Analysis of Deregulatory Dynamics and Technical Efficiency: The Case of the U.S. Airline Industry,” *International Economic Review*, 41(1), pp. 203–18.

Anstine, D.B., 2001, “How Much Will Consumers Pay? A Hedonic Analysis of the Cable Television Industry,” *Review of Industrial Organization*, 19(2), pp. 129–47.

Arellano, M., and Bover, O., 1995, “Another Look at the Instrumental Variable Estimation of Error-Components Models,” *Journal of Econometrics*, 68(1), pp. 29–51.

Arellano, M., and Bond, S., 1991, “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,” *The Review of Economic Studies*, 58(2), pp. 277–97.

Agarwal, S.; Chomsisengphet, S.; Mahoney, N. and Stroebel, J., 2015, “A Simple Framework for Estimating Consumer Benefits from Regulating Hidden Fees,” *Journal of Legal Studies*, 43(S2), pp. S239–52.

Arjomandi, A. and Seufert, J.H., 2014, “An Evaluation of the World’s Major Airlines’ Technical and Environmental Performance,” *Economic Modelling*, 41(August), pp. 133–44.

Armknecht, P.A. and Ginsburg, D.H., 1992, “Improvements in Measuring Price Changes in Consumer Services: Past, Present, and Future,” In *Output Measurement in the Service Sectors*, 109–57. (University of Chicago Press).

Arrondo, R.; Garcia, N. and Gonzalez, E., 2018, “Estimating Product Efficiency through a Hedonic Pricing Best Practice Frontier,” *BRQ Business Research Quarterly*, 21(4), pp. 215–24.

Assaf, A., 2009, “Are U.S. Airlines Really in Crisis?,” *Tourism Management,* 30(6), pp. 916–21.

Assaf, A., 2011, “A Fresh Look at the Productivity and Efficiency Changes of UK Airlines,” Applied Economics, 43(17), pp. 2165–75.

Bailey, E.E. and Baumol, W.J., 1984, “Deregulation and the Theory of Contestable Markets,” *Yale Journal on Regulation*,. 1(2), pp. 111-137.

Balk, B.M. and Zofío, J.L., 2020, “Symmetric Decompositions of Cost Variation,” *European Journal of Operational Research,* 285(3), pp. 1189–98.

Barros, C.P.; Liang, Q. Bin and Peypoch, N., 2013, “The Technical Efficiency of US Airlines,” *Transportation Research Part A: Policy and Practice,* 50(April), pp. 139–48.

Barrows, S.D., 2018, “After the Recent US Airline Mergers , Did Domestic Airfares Increase ?,” *International Review of Management and Marketing*, 8(1), pp. 1–8.

Berger, A.N.; Humphrey, D.B. and Pulley, L.B., 1996, “Do Consumers Pay for One-Stop Banking? Evidence from an Alternative Revenue Function,” *Journal of Banking and Finance*, 20(9), pp. 1601–21.

Bitzan, J. and Peoples, J., 2016, “A Comparative Analysis of Cost Change for Low-Cost, Full-Service, and Other Carriers in the US Airline Industry,” *Research in Transportation Economics*, 56(August), pp. 25–41.

Blundell, R. and Bond, S., 1998, “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics*, 87(1), pp. 115–43.

Borenstein, S., 1985, “Price Discrimination in Free-Entry Markets,” *The RAND Journal of Economics,* 16(3), pp. 380–97.

Borenstein, S., 1989, “Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry,” *The RAND Journal of Economics*, 20(3), pp. 344–65.

Borenstein, S. and Rose, N.L., 1994, “Competition and Price Dispersion in the U.S. Airline Industry,” *Journal of Political Economy*, 102(4), pp. 653–83.

Borenstein, S., 2011. “Why Can't US Airlines Make Money?” *American Economic Review,* Vol. 101(3), pp. 233-37

Bourassa, S.C.; Cantoni, E. and Hoesli, M., 2016, “Robust Hedonic Price Indexes,” *International Journal of Housing Markets and Analysis*, 9(1), pp. 47–65.

Brea-Solís, H.; Casadesus-Masanell, R. and Grifell-Tatjé, E., 2015, "Business Model Evaluation: Quantifying Walmart's Sources of Advantage," *Strategic Entrepreneurship Journal*, 9(1), pp. 12–33.

Brueckner, J.K.; Lee, D.N.; Picard, P.M. and Singer, E., 2015, "Product Unbundling in the Travel Industry: The Economics of Airline Bag Fees," *Journal of Economics and Management Strategy*, 24(3), pp. 457–84.

Cantos, P. and Maudos, J., 2001, "Regulation and Efficiency: The Case of European Railways," *Transportation Research Part A: Policy and Practice*, 35(5), pp. 459–72.

Carew, R., 2000, "A Hedonic Analysis of Apple Prices and Product Quality Characteristics in British Columbia," *Canadian Journal of Agricultural Economics*, 48(3), pp. 241–57.

Caves, D.W.; Christensen, L.R. and Tretheway, M.W., 1984, "Economies of Density versus Economies of Scale: Why Trunk and Local Service Airline Costs Differ," *The RAND Journal of Economics*, 15(4), pp. 471.

Chandra, A. and Lederman, M., 2018, "Revisiting the Relationship between Competition and Price Discrimination," *American Economic Journal: Microeconomics*, 10(2), pp. 190–224.

Charnes, A.; Cooper, W.W. and Rhodes, E., 1978, "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research*, 2(6), pp. 429–44.

Christensen, L.R. and Greene, W.H., 1976, "Economies of Scale in U.S. Electric Power Generation," *Journal of Political Economy*, 84(4, Part 1), pp. 655–76.

Christensen, L.R.; Jorgenson, D.W. and Lau, L.J., 1973, "Transcendental Logarithmic Production Frontiers," *The Review of Economics and Statistics*, 55(1), pp. 28–45.

Chwelos, P.D.; Berndt, E.R. and Cockburn, I.M., 2008, "Faster, Smaller, Cheaper: An Hedonic Price Analysis of PDAs," *Applied Economics*, 40(22), pp. 2839–56.

Coelli, T.; Gautier, A.; Perelman, S. and Saplacan-Pop, R., 2013, "Estimating the Cost of Improving Quality in Electricity Distribution: A Parametric Distance Function Approach," *Energy Policy*, 53(February), pp. 287–97.

Coelli, T.; Perelman, S. and Romano, E., 1999, “Accounting for Environmental Influences in Stochastic Frontier Models: With Application to International Airlines,” *Journal of Productivity Analysis*, 11(3), pp. 251–73.

Cotteleer, G.; Gardebroek, C. and Luijt, J., 2008, “Market Power in a GIS-Based Hedonic Price Model of Local Farmland Markets,” *Land Economics*, 84(4), pp. 573–92.

Court, A.T., 1939, “Hedonic Price Indexes with Automotive Examples,” In: Roos CF., (ed), *The Dynamics of Automobile Demand*, 99–117. (General Motors Corporation, New York, NY).

Cummins, J.D.; Weiss, M.A.; Xie, X. and Zi, H., 2010, “Economies of Scope in Financial Services: A DEA Efficiency Analysis of the US Insurance Industry,” *Journal of Banking and Finance*, 34(7), pp. 1525–39.

Dahlby, B. and West, D.S., 1986, “Price Dispersion in an Automobile Insurance Market,” *The Journal of Political Economy*, 94(2), pp. 418–38.

Dai, M.; Liu, Q. and Serfes, K., 2014, “Is the Effect of Competition on Price Dispersion Nonmonotonic? Evidence from the U.S. Airline Industry,” *Review of Economics and Statistics*, 96(1), pp. 161–70.

De Witte, K. and Saal, D.S., 2010, “Is a Little Sunshine All We Need? On the Impact of Sunshine Regulation on Profits, Productivity and Prices in the Dutch Drinking Water Sector,” *Journal of Regulatory Economics*, 37(3), pp. 219–42.

Douglas, G.W. and Miller, J.C., 1974, “American Economic Association Quality Competition, Industry Equilibrium, and Efficiency in the Price-Constrained,” *The American Economic Review*, 64(4), pp. 657–69.

Eden, B., 2018, “Price Dispersion and Demand Uncertainty: Evidence From U.S. Scanner Data,” *International Economic Review*, 59(3), pp. 1035–75.

Fabra, N. and Reguant, M., 2014, “Pass-through of Emissions Costs in Electricity Markets,” *American Economic Review*, 104(9), pp. 2872–99.

Färe, R.; Grosskopf, S. and Sickles, R.C., 2007, “Productivity? Of US Airlines After Deregulation,” *Journal of Transport Economics and Policy*, 41, pp. 93–112.

Färe, R.; Grosskopf, S. and Lovell, C.A.K., 1985, *The Measurement of Efficiency of Production*. (Springer Netherlands, Dordrecht).

Feenstra, R.C., 1995, “Exact Hedonic Price Indexes,” *The Review of Economics and Statistics*, 77(4), pp. 634–53.

Feng, G. and Wang, C., 2018, “Why European Banks Are Less Profitable than U.S. Banks: A Decomposition Approach,” *Journal of Banking and Finance*, 90(May), pp. 1–16.

Forsund, F.R. and Hjalmarsson, L., 1979, “Generalised Farrell Measures of Efficiency: An Application to Milk Processing in Swedish Dairy Plants,” *The Economic Journal*, 89(354), pp. 294–315.

Gaggero, A., and Piga, C., 2011, “Airline Market Power and Intertemporal Price Dispersion.” *Journal of Industrial Economics*, 59 (4), pp. 552–77.

Galán, J.E.; Veiga, H. and Wiper, M.P., 2014, “Bayesian Estimation of Inefficiency Heterogeneity in Stochastic Frontier Models,” *Journal of Productivity Analysis*, 42(1), pp. 85–101.

Gayle, P.G. and Yimga, J.O., 2018, “How Much Do Consumers Really Value Air Travel On-Time Performance, and to What Extent Are Airlines Motivated to Improve Their on-Time Performance?,” *Economics of Transportation*, 14(June), pp. 31–41.

Gerardi, K.S. and Shapiro, A.H., 2009, “Does Competition Reduce Price Dispersion? New Evidence from the Airline Industry,” *Journal of Political Economy*, 117(1), pp. 1–37.

Gillen, D.W.; Oum, H. and Tretheway, M.W., 1990, “Airline Cost Structure and Policy Implications: A Multi-Product Approach for Canadian Airlines,” *Journal of Transport Economics and Policy*,. 24(1), pp. 9-34

Good, D.H.; Sickles, R.C. and Weiher, J.C., 2008, “A Hedonic Price Index for Airline Travel,” *Review of Income and Wealth*, 54(3), pp. 438–65.

Grifell-Tatjé, E. and Lovell, C.A.K., 1999, “Profits and Productivity,” *Management Science*, 45(9), pp. 1177–93.

Grifell-Tatjé, E. and Lovell, C.A.K., 2015, *Productivity Accounting, The Economics of Business Performance*. (Cambridge University Press, New York).

Griliches, Z., 1961, “Hedonic Price Indexes for Automobiles: An Econometric of Quality Change,” In *The Price Statistics of the Federal Goverment*, 173–96. (NBER, Cambridge, MA and London, England).

Hanson, T., 2016, “Efficiency and Productivity in the Operational Units of the Armed Forces: A Norwegian Example,” *International Journal of Production Economics*, 179, pp. 12–23.

Hazel, R., 2018, “Airline Capacity Discipline in the U.S. Domestic Market,” *Journal of Air Transport Management*, 66(January), pp. 76–86.

Heckman, J.J.; Matzkin, R.L. and Nesheim, L., 2010, “Nonparametric Identification and Estimation of Nonadditive Hedonic Models,” *Econometrica*, 78(5), pp. 1569–91.

Hill, R.J., 2013, “Hedonic Price Indexes for Residential Housing: A Survey, Evaluation and Taxonomy,” *Journal of Economic Surveys*, 27(5), pp. 879–914.

Holmes, T.J., 1989, “The Effects of Third-Degree Price Discrimination in Oligopoly,” *The American Economic Review*, 79(1), pp. 244–50.

Hulten, C.R., 2003, “Price Hedonics: A Critical Review,” *Economic Policy Review*, 9(3), pp. 5–15.

Humphrey, D.B. and Pulley, L.B., 1997, “Banks’ Responses to Deregulation: Profits, Technology, and Efficiency,” *Journal of Money, Credit and Banking*, 29(1), pp. 73.

IATA, 2013, “Profitability and the Air Transport Value Chain,”, available at https://www.iata.org/en/iata-repository/publications/economic-reports/profitability-and-the-air-transport-value-chain/

IATA, 2018, “Economic Performance of the Airline Industry. End year 2018 report,” available at https://www.iata.org/publications/economics/Reports/Industry-Econ-Performance/IATA-Economic-Performance-of-the-Industry-end-year-2018-report.pdf

Ippolito, R.A., 1981, “Estimating Airline Demand with Quality of Service Variables,” *Journal of Transport Economics and Policy*, 15(1), pp. 7–15.

Johnston, A. and Ozment, J., 2013, “Economies of Scale in the US Airline Industry,” *Transportation Research Part E: Logistics and Transportation Review*, 51(May), pp. 95–108.

Judd, C.M., and Kenny, D. A., 1981, “Process Analysis: Estimating Mediation in Treatment Evaluations, ” *Evaluation Review,* 5(5), pp. 602–619.

Kalemba, N. and Campa-Planas, F., 2017, “The Quality Effect on the Profitability of US Airline Companies,” *Tourism Economics,* 24(3), pp. 251–69.

Kaplan, G.; Menzio, G.; Rudanko, L. and Trachter, N., 2016, “Relative Price Dispersion: Evidence and Theory,” *National Bureau of Economic Research. Working Paper Series,* No. 21931, pp. 1–72.

Konüs, A.A., 1939, “The Problem of the True Index of the Cost of Living,” *Econometrica*, 7(1), pp. 10–29.

Kotler, P., 1967, *Marketing Management: Analysis, Planning, and Control,* (Prentice-Hall, Englewood Cliffs, N.J.)

Lach, S. and Moraga-González, J.L., 2017, “Asymmetric Price Effects of Competition,” *The Journal of Industrial Economics,* 65(4), pp. 767–803.

Lancaster, K.J., 1966, “A New Approach to Consumer Theory,” *Journal of Political Economy*, 74(2), pp. 132–57.

Lee, B.L. and Worthington, A.C., 2014, “Technical Efficiency of Mainstream Airlines and Low-Cost Carriers: New Evidence Using Bootstrap Data Envelopment Analysis Truncated Regression,” *Journal of Air Transport Management*, 38(June), pp. 15–20.

Liu, J.S.; Lu, L.Y.Y.; Lu, W.M. and Lin, B.J.Y., 2013, “A Survey of DEA Applications,” *Omega The International Journal of Management Science*, 41(5), pp. 893–902.

Loy, J.P. and Weiss, C., 2019, “Product Differentiation and Cost Pass-Through,” *Journal of Agricultural Economics*, 70(3), pp. 840–58.

Mccartney, S., 2018, February 14, “How Much of Your $355 Ticket Is Profit for Airlines?” *Wall Street Journal*.

Mellat-Parast, M.; Golmohammadi, D.; Mcfadden, K.L. and Miller, J.W., 2015, “Linking Business Strategy to Service Failures and Financial Performance: Empirical Evidence from the U.S. Domestic Airline Industry,” *Journal of Operations Management*, 38, pp. 14–24.

Merkert, R. and Pearson, J., 2015, "A Non-Parametric Efficiency Measure Incorporating Perceived Airline Service Levels and Profitability," *Journal of Transport Economics and Policy*, 49(April 2014), pp. 261–75.

Morrison, S. and Winston, C., 1995, *The Evolution of the Airline Industry*. (The Brookings Institution Press, Washington, DC)

Mussa, M. and Rosen, S., 1978, "Monopoly and Product Quality," *Journal of Economic Theory*, 18(2), pp. 301–17.

Nerlove, M., 1963, "Returns to Scale in Electricity Supply," In *Measurement in Economics: Studies in Mathematical Economics and Econometrics*. (Stanford University Press, Stanford, CA), pp. 167-98

Obermeyer, A.; Evangelinos, C. and Püschel, R., 2013, "Price Dispersion and Competition in European Airline Markets," *Journal of Air Transport Management*, 26, pp. 31–34.

Oum, T.H. and Yu, C., 1998, "An Analysis of Profitability of the World's Major Airlines," *Journal of Air Transport Management*, 4(4), pp. 229–37.

Oum, T.H.; Fu, X. and Yu, C., 2005, "New Evidences on Airline Efficiency and Yields: A Comparative Analysis of Major North American Air Carriers and Its Implications," *Transport Policy*, 12(2), pp. 153–64.

Pakes, A., 2003, "A Reconsideration of Hedonic Price Indexes with an Application to PC's," *American Economic Review*, 93(5), pp. 1578–96.

Podinovski, V. V., 2004, "Bridging the Gap between the Constant and Variable Returns-to-Scale Models: Selective Proportionality in Data Envelopment Analysis," *Journal of the Operational Research Society*, 55(3), pp. 265–76.

Ray, S.C. and Mukherjee, K., 1996, "Decomposition of the Fisher Ideal Index of Productivity: A Non-Parametric Dual Analysis of US Airlines Data," *The Economic Journal*, 106(439), pp. 1659–78.

Restrepo-Tobón, D.A. and Kumbhakar, S.C., 2017, "A New Method to Decompose Profit Efficiency: An Application to US Commercial Banks," *Journal of Productivity Analysis*, 48, pp. 117–32.

Roodman, D., 2003, “XTABOND2: Stata Module to Extend Xtabond Dynamic Panel Data Estimator”, *Statistical Software Components S435901, Boston College Department of Economics*.

Rosen, S., 1974, “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition,” *Source Journal of Political Economy*, 82(1), pp. 34–55.

Salop, S., 1979, “Monopolistic Competition with Outside Goods,” *The Bell Journal of Economics*, 10(1), pp. 141–56.

Savioli, M. and Zirulia, L., 2020, “Does Add-on Presence Always Lead to Lower Baseline Prices? Theory and Evidence,” *Journal of Economics*, 129, pp. 143-72.

Seim, K., 2006, “An Empirical Model of Firm Entry with Endogenous Product-Type Choices,” *The RAND Journal of Economics*, 37(3), pp. 619–40.

Schefczyk, M., 1993, “Operational Performance of Airlines: An Extension of Traditional Measurement Paradigms,” *Strategic Management Journal*, 14(4), pp. 301–17.

Scotti, D. and Volta, N., 2017, “Profitability Change in the Global Airline Industry,” *Transportation Research Part E: Logistics and Transportation Review*, 102, pp. 1–12.

Sharp, B. and Dawes, J., 2001, “What Is Differentiation and How Does It Work ?,” *Journal of Marketing Management*, 17(7-8), pp. 739–59.

Shepard, A., 1991, “Price Discrimination and Retail Configuration,” *Journal of Political Economy*, 99(1), pp. 30–53.

Shephard, R.W., 1953, *Cost and Production Functions*. (Princeton University Press, Princeton, N.J.).

Spady, R.H. and Friedlaender, A.F., 1978, “Hedonic Cost Functions for the Regulated Trucking Industry,” *The Bell Journal of Economics*, 9(1), pp. 159-79.

Stahl II, D.O., 1989, “Oligopolistic Pricing with Sequential Consumer Search,” *American Economic Review*, 79(4), pp. 700–712.

Stavins, J., 2001, “Price Discrimination in the Airline Market: The Effect of Market Concentration,” *The Review of Economics and Statistics*, 83(1), pp. 200–202.

Stigler, G.J., 1964, “A Theory of Oligopoly,” *Journal of Political Economy*, 72(1), pp. 44–61.

Suzuki, Y., 2000, “The Relationship between On-Time Performance and Airline Market Share: A New Approach,” *Transportation Research Part E: Logistics and Transportation Review*, 36(2), pp. 139–54.

Taylor, L.O. and Smith, V.K., 2000, “Environmental Amenities as a Source of Market Power,” *Land Economics*, 76(4), pp. 550.

Triplett, J.E., 1991, “Hedonic Methods in Statistical Agency Environments: An Intellectual Biopsy,” In *Fifty Years of Economic Measurement: The Jubilee of the Conference on Research in Income and Wealth*, (University of Chicago Press, Chicago), pp. 207-38

Triplett, J.E., 2004, “Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes.” OECD Science, Technology and Industry. Working Papers, No. 2004/09. (Paris).

Tsionas, E.G., 2003, “Combining DEA and Stochastic Frontier Models: An Empirical Bayes Approach,” *European Journal of Operational Research*, 147(3), pp. 499–510.

Tulkens, H. and Eeckaut, P. V., 1995, “Non-Parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects,” *European Journal of Operational Research*, 80(3), pp. 474–99.

Wang, K.; Fan, X.; Fu, X. and Zhou, Y., 2014, “Benchmarking the Performance of Chinese Airlines: An Investigation of Productivity, Yield and Cost Competitiveness,” *Journal of Air Transport Management*, 38, pp. 3–14.

Wojahn, O.W., 2012, “Why Does the Airline Industry Over-Invest?” *Journal of Air Transport Management*, 19(1), pp. 1–8.

Wu, Y.; He, C. and Cao, X., 2013, “The Impact of Environmental Variables on the Efficiency of Chinese and Other Non-Chinese Airlines,” *Journal of Air Transport Management*, 29, pp. 35–38.

Yu, M.M.; Chen, L.H. and Hsiao, B., 2018, “A Performance-Based Subsidy Allocation of Ferry Transportation: A Data Envelopment Approach,” *Transport Policy*, 68, pp. 13–19.

Zou, B. and Hansen, M., 2012, “Impact of Operational Performance on Air Carrier Cost Structure: Evidence from US Airlines,” *Transportation Research Part E: Logistics and Transportation Review*, 48(5), pp. 1032–48.