MARKET POWER, COST EFFICIENCY AND PRICING STRATEGIES OF DOMESTIC AIRLINE INDUSTRY

A Dissertation
Presented to
The Academic Faculty

by

Ran Wang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Economics

Georgia Institute of Technology
August 2015

MARKET POWER, COST EFFICIENCY AND PRICING STRATEGIES OF DOMESTIC AIRLINE INDUSTRY

Approved by:

Dr. Levent Kutlu, Advisor
School of Economics
Georgia Institute of Technology

Dr. Vivek Ghosal
School of Economics
Georgia Institute of Technology

Dr. Byung-Cheol Kim
School of Economics
Georgia Institute of Technology

Dr. Usha C. Nair-Reichert
School of Economics
Georgia Institute of Technology

Dr. Jerry Thursby
College of Management
Georgia Institute of Technology

Date Approved: JUN 23rd, 2015

I dedicate my dissertation to my beloved family and my advisor. I have to thank my parents for their love and support throughout my life. A special feeling of gratitude to my husband, Weiwei Yang. He always supports every decision that I make during life. His enduring love and encouragement ring in my ears. Also, I always appreciate the encouragement from my brother, Yong Wang. He always cheers me up when I am unhappy. I will always appreciate what my advisor, Levent Kutlu has done for me. Levent is a nice and knowledgeable advisor and friend to me.

ACKNOWLEDGEMENTS

There are a number of people without whom this thesis might not have been written, and to whom I am greatly indebted. First, I am grateful to my committee members, Dr. Levent Kutlu, Dr. Vivek Ghosal, Dr. Byung-Cheol Kim, Dr. Usha C. Nair-Reichert and Dr. Jerry Thursby. I have to express my thanks and appreciation to Dr. Levent Kutlu for the guidance and support as my advisor throughout the three years for completing this research and writing the dissertation. Levent inspires me about the research idea and helps me with the research methodologies. His hard work and accomplishments in research always encourage me. Levent is also my friend who shares happiness with me. Dr. Vivek Ghosal, as my secondary advisor, provides me insights about the research from industrial perspectives. I would like to thank him for the comments from empirical aspects on my dissertation. Dr. Byung-Cheol Kim gives me a lot of constructive suggestions on the theoretical sections. His pure research spirit sets an example for me. Dr. Usha C. Nair-Reichert has helped me a lot since the first day I entered Economics Ph.D. program. She encourages me in the research and more importantly, she helps me a lot in life as a friend. Dr. Jerry Thursby, from College of Business, provides me many helpful suggestions on the theoretical framework, as well as the empirical parts. His suggestions improve my dissertation a lot.

Besides my committee members, I have to express my appreciations and deep gratitude to Dr. Emilson Caputo Delfino Silva. He is the first professor who collaborates with me. His knowledge and patience help me a lot build a systematic understanding of academic research. Also, I would like to thank Dr. Mikhail M. Klimenko for his suggestions on job market.

I must acknowledge the financial support and other supports from School of Economics in Georgia Institute of Technology. The school chair, Dr. David N. Laband, graduate director, Dr. Vivek Ghosal and undergraduate director, Dr. Usha C. Nair-Reichert provide us a good research environment by providing the Ph.D. students stable financial supports and teaching opportunities. Also, I would like to thank the advisors in School of Economics, Tony Gallego and Jyldyz Hughes. They provide me great help with my teaching assistant work. I highly appreciate the good research and life environment provided by School of Economics and the whole Georgia Institute of Technology. Without the good research environment, I could not make through this five years.

I also would like to give me appreciations to my dearest friends and classmates. I am hesitated to say goodbye to my friends in Georgia Tech, who have accompanied me for five year. Graduation happens and sends us to different places. However, friendship endures. I wish all of us a good future and a happy life.

My deepest gratitude goes to my family members: my mother, my father, my brother and my husband for their enduring love and encouragement. Every step that I take, my family members are always there for me. 2015 is a big year for me, getting married and graduating from Ph. D. program. 2015 will be in my memory forever.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis examines the market power, cost efficiency, price dispersion and revenue management (dynamic pricing behaviors) in the airline industry. Chapter II develops a theoretical framework to enable the estimation of cost efficiency without total cost data. Based on the estimates obtained from Chapter II, we continue to examine the price dispersion determinants in Chapter III, especially focusing on cost efficiency and market power. Chapter IV proceeds to examine the dynamic pricing strategy differences between high-efficiency and low-efficiency flights.

The estimation of cost efficiency without total cost data is impossible from previous literatures. Stochastic Frontier Analysis (SFA) obtains the cost efficiency from total cost data. Also, traditional market power literatures suffer from same problem that they require total cost data to estimate marginal cost and thus estimate market power such as Lerner Index in Lerner (1934). One of the advantages of conduct parameter games is that they enable estimation of market power without total cost data. In Chapter II, this study applies the conduct parameter framework in SFA estimations and develops a conduct parameter based model to estimate the firm specific implied marginal cost efficiency and conduct parameter without using total cost data. Our study is the first to relax the total cost data requirement for estimating cost efficiency, to the best of our knowledge. We testify the theoretical framework by estimating the conducts and marginal cost efficiencies of U.S. airlines. Also, we find support for QLH (Quiet Life Hypothesis) based on the estimated conduct and cost efficiency.

In Chapter III, we analyze the determinants of price dispersion for airline industry. We particularly concentrate on the conduct parameter and marginal cost efficiency. The effect of conduct on price dispersion seems to depend on the characteristics of the

market. For the big city routes, we observe a positive effect; however for the leisure routes we find a negative effect. Moreover, we find that marginal cost efficiency has a negative effect on price dispersion. Chapter III also sheds light on the potential estimation bias in previous studies.

In Chapter III, we find a large price dispersion in the U.S. airlines for Chicago based routes (Gini coefficient is 0.21). We are interested in the dynamics pricing strategies and revenue management strategies that would lead to such high price dispersion. In Chapter IV, by scripting the daily prices and seats information from priceline.com, we examine the revenue efficiencies of the airlines. Based on the revenue efficiencies of the flights, we further divide the flights into high-efficiency and low-efficiency flights. Then we compare the differences in dynamic pricing patterns between these two groups. Based on the findings, we give potential policy applications to the airlines in order to improve their revenue efficiencies.

CHAPTER I

INTRODUCTION

Market power and cost efficiency are two popular topics in Industrial Organization. However, these two literatures seem to be developed paralleled. There are very few studies that combine these two areas together. Chapter II in this thesis combines two classical frameworks in market power and cost efficiency to overcome the difficulties faced by both traditional market power and cost efficiency literatures.

Stochastic frontier analysis (SFA) literature (Kumbhakar and Lovell (2000)) relaxes the full efficiency assumption of neoclassical production theory by allowing firms to act suboptimally. There are many potential sources for inefficiency, one of which is the principle agent problem. The objectives of share-holders and managers are not fully aligned. For example, the managers may use extra staff to reduce managers workloads at the expense of higher costs. In SFA literature, cost efficiency is defined as the ratio of minimum (optimal) cost to its actual cost. A conventional stochastic frontier model estimates cost inefficiency from a log-log cost function, which treats the cost inefficiency as an unobserved non-negative error term. The SFA model has a composed error term consisting of a one-sided error term, which captures cost inefficiency, and a conventional two-sided error term. And the cost inefficiency is identified from the asymmetry of the one-sided error term. Hence, SFA also requires total cost data to obtain the cost efficiencies. However, the total cost data is most of the time not available. The firms are not willing to share the total cost data for many strategic reasons.

Lerner index (1934) is a widely used market power measure, whose calculation requires the total cost data as well.[1] Conduct parameter (or conjectural variations) game (Perloff, Karp, and Golan (2007) and Bresnahan (1989)) solves this problem using a demand and supply system, which allows the firm to form a conjecture about the variation in the other

[1] The calculation of marginal cost requires total cost data.

firms' strategies (e.g., output) in response to a change in its own strategy. Under conduct parameter framework, the estimation is based on "perceived marginal revenue" and marginal cost, which does not require total cost data.

Our motivation for Chapter II is to combine conduct parameter framework and SFA literature. By doing this, this study allows for the estimations of marginal cost efficiency and conduct parameter simultaneously and explicitly. The conduct parameter game enables us to estimate "implied marginal cost efficiencies" and conduct parameters jointly without using the total cost data. Hence, compared with the traditional SFA literature, which infers the cost efficiency from a cost function, we estimate the marginal cost efficiency from a supply-demand system, which is derived from a conduct parameter game. To the best of our knowledge, this is the first study that enables estimation of marginal cost efficiency that doesn't require the total cost data.

The relationship between market power and cost efficiency has been acknowledged by the economists since a long time ago. The "Quiet Life Hypothesis" (QLH) by Hicks (1935) and the "Efficient Structure Hypothesis" (ESH) by Demsetz (1973) are two well-known hypotheses that explain the potential relationship between market power and cost efficiency. QLH claims that higher competition pressure is likely to force managers to work harder, which in turn increases efficiencies of firms. ESH states that the firms with superior efficiency levels use their competitive advantages to gain larger market shares, which lead to higher market concentration and thus higher market power. Both hypotheses are supported by different empirical studies. Berger and Hannan (1998) and Kutlu and Sickles (2012) support the QLH for the banking and airline industries, respectively. Maudos and Fernandez de Guevara (2007) show evidence for ESH for the banking industry. Moreover, Delis and Tsionas (2009) find mixed results for two hypotheses. Hence, the relationship between market power and efficiency has long been acknowledged from both theoretical aspects and empirical aspects.

However, both the market power literatures and SFA literatures largely ignore this relationship. Until recently, some studies attempt to estimate market powers of firms in a framework where firms are allowed to be inefficient, such as Koetter, Kolari and Spierdijk

(2012), Delis and Tsionas (2009), Koetter and Poghosyan (2009), Kutlu and Sickles (2012) and Delis and Tsionas (2009). The relationship between market power and cost efficiency would lead to inconsistent estimates for both traditional conduct parameter and SFA studies. The reason behind this is quite intuitive. The traditional conduct parameter studies ignore the cost efficiency. For example, in a Cournot market, the best practicing firm may not really be fully efficient. Lee and Johnson (2012) show that, in the Cournot environment, inefficiency may in fact be a result of endogenous prices and the effect of output production on price. Similarly, ignoring inefficiencies of firms in a conduct parameter model would lead to omitted variable bias. In the traditional SFA studies, the studies do not control for the market power. The differences in market powers would lead to different firm behaviors. If the firm level conducts affect the performance of the best-practice units, the efficiency estimates would not be accurate without considering market power effect. Our methodology overcomes these difficulties by explicitly and simultaneously modeling a conduct parameter game, and under this framework the firms are allowed to be inefficient.

Also, the simultaneous estimation of conduct parameter and marginal cost efficiency can provide a more accurate calculation of dead-weight-loss (DWL). How to measure inefficiencies and market powers of firms when the firms face optimization constraints is an interesting question for the economists in recent years. Puller (2007) develops a conduct parameter model measuring market powers of firms in the California electricity market under capacity constraints. Inspired by Puller (2007), Chapter II also presents an extension of our framework to take capacity constraints into consideration. This extension allows the estimation of conduct parameter and marginal cost efficiency when facing capacity constraints and when total cost data is not available.

To testify our theoretical framework, we apply our methodology to estimate the firm-route-quarter specific conducts and marginal cost efficiencies of the U.S. airlines for routes that originate from Chicago. The time period that the data set in Chapter II covers is 1999I-2009IV. The airline data is suitable for our research because the available cost data set is for the entire U.S. system, but the route level total cost data is not available. Our results suggest that concentration ratio (measured by CR_4) and market share of airlines are negatively related to the marginal cost efficiency, so we find supports for the QLH from the airline data. In contrast to this, the concentration ratio and market share of airlines are positively related to the conducts. By doing this, we validate our theoretical framework in the empirical study in Chapter II.

In Chapter II, we find that there is large price dispersions in the airline industry. It means that for the Chicago based routes, the Gini coefficient has a mean value of 0.21. It means that for two randomly selected tickets, the absolute price differences as the ratio of mean price is 42%. The sources of price dispersion attract us and drive us to explore into this topic. There are many sources of price dispersion, as shown in Borenstein and Rose (1994). A variety of pricing strategies can lead to price dispersion such as peak-load pricing and stochastic demand pricing. Moreover, price discrimination is one important source of price dispersion. It is a popular topic to measure how market power affect price dispersion both theoretically and empirically. On the one hand, competition can limit the firms' price discrimination abilities and force them to charge a single price. Following traditional textbook theory, the extent of price discrimination is positively related to the market power of firms. On the other hand, different consumers have different elasticities of demand, so an increase in competition might also induce the airlines to charge higher prices to the high-end (first class and business class) consumers due to their low elasticities of demand, leading to higher price dispersion. Under this situation, this relationship is negative. So, the theoretical direction between market power and price dispersion is ambiguous.

The structure-conduct-performance (SCP) paradigm states that the structure of an industry, such as market concentration affects the conduct of firms and thus affects the performance of the industry. SCP paradigm shows an inverse relationship between competition and market concentration. The empirical studies mostly concentrate on the relationship between market concentration and price dispersion. Under SCP paradigm, we interpret such studies as relationship between market power and price dispersion. However, there are many reasons that encourage us to believe that market concentration measures such as Herfindahl-Hirschman Index (HHI) may not be sufficient to capture market power. HHI is a market level measurement, which does not include the information about elasticity of demand and cost efficiency information. In order to solve this issue, we include conduct parameter, instead of HHI as a measure for market power.

In this study the determinants of price dispersion that we consider are: conduct, marginal cost efficiency, other cost based factors, population attributes, and product attributes. In particular, we are interested in how conduct and marginal cost efficiency affect price dispersion. In Chapter II, we find evidence for QLH that market power and cost efficiency are negatively related. However, previous literatures in price dispersion largely ignore this relationship and do not control for cost efficiency when they estimate the competition's effect on price dispersion. One of our objectives of Chapter III is to identify whether there is bias in previous estimation due to ignorance of cost efficiency. We find a correlation between one instrumental variable called geometric market share used by Borenstein and Rose (1994) and Gerardi and Shapiro (2009) and the omitted variable, marginal cost efficiency. We suspect that this correlation might lead to bias in their estimations.

How cost efficiency affects price dispersion is rarely explored by price dispersion literature. The reason is that firm level cost efficiency data is not available. In Chapter II, we develop a theoretical framework that allows simultaneous estimation of firm level conduct parameter and marginal cost efficiency. And the empirical part in Chapter II provides us the estimates of firm-route-quarter level conduct parameter and marginal cost efficiency. Based on these estimates, we are able to explore the relationship between marginal cost efficiency and price dispersion in Chapter III. An increase in marginal cost efficiency does not necessarily lead to same amount of price changes for the high-end travelers and low-end travelers. For instance, if the market share for high-end consumers for an airline is larger than that of low-end consumers. This Airline would be reluctant to increase prices for high-end tickets, afraid of losing long term high profit from high-end consumers. However, the high-end consumers have low elasticity of demand and thus less sensitive to the price changes. This Airline would also take the differences in demand elasticities into account when making pricing decisions. So, the relationship could be positive or negative depending on the market characteristics.

Using the conduct and cost efficiency estimates from Chapter II, Chapter III sheds light on the determinants of price dispersion. The main findings are as follows. First, we find that omitted variable of cost efficiency may lead to overestimation of market power's effect on price dispersion. The overestimations are robust to different subsamples. However, the difference is not significant at 5% significance level. Second, same as in Gerardi and Shapiro (2009), we find different effects from conduct on price dispersion between big city routes and leisure routes. Conduct has a positive effect on price discrimination. Third, marginal cost efficiency has a negative effect on price dispersion, which means that an increase in marginal cost efficiency leads to larger amount of prices changes for the high-end consumers than low-end consumers. Fourth, we find a negative relationship between conduct and price dispersion for the whole sample.

Chapter III in this thesis explores the relationship between conduct, marginal cost efficiency and price dispersion. While working on price dispersion, we are attracted by the dynamics of the pricing behavior that lead to such high price dispersion. It is common to have a round trip ticket's price lower than an one-way ticket's price for the same city pairs. The pricing patterns of the flight tickets are always mysteries for the consumers. So, in Chapter IV, we are trying to examine how the airlines dynamically make the pricing decisions in order to maximize their revenue. In this study, we focus on revenue optimization. This is because once the flight schedule is determined, the fixed cost is large enough to allow us to ignore the variable cost.

Revenue management is also called yield management or seat inventory management or dynamic pricing. In the airline industry, airlines sell identical seats at different prices to maximize their revenues. Since American Airlines introduced the yield management technique in early 1980s, yield management becomes more and more popular and sophisticated. Davis (1994) and Smith et al. (1992) show that American Airlines benefit from Yield management (YM) and Smith et al. (1992) state that American Airlines made an extra \$1.4 billion between 1989 and 1991 because of its advanced yield management techniques. Pinder (1995), Sridharan (1998) and Barut and Sridharan (2004) state that the made-to-order (MTO) manufacturing industry is suitable for revenue management. Yield management is a general practice for perishable inventory control, such as hotel, rental car, cruises and flight tickets. There are two common characteristics of these products or service. First, the product/service expires at a certain point of time. Second, the capacity is fixed in advanced and the capacity constraint can only be extended at a very high marginal cost. These two characteristics make the dynamic pricing strategy highly important for these perishable goods/service.

There are mainly three sets of yield management theories. One is traditional price discrimination theory. Under this theory, the airlines utilize different ticket restrictions to segment the consumers. The consumers make purchase decisions based on their own preferences and budgets. The second one is capacity-based theories. Based on these theories, capacity is limited in the airline industry and the cost associated with augmented capacity is large, that is, airline capacity is costly but perishable. Prescott (1975), Eden (1990) and Dana (1999b) explain the relationship between fares and seat availability under the assumption of uncertain demand. Moreover, Dana (1999a) shows that price dispersion increases demand shifting and thus increases the social welfare by allocating the consumers into available seats. The third set of theory is time-based theories. Gale and Holmes (1992, 1993) model advance purchase discount in a monopoly market. The monopoly firm uses fare discounts to divert the consumers from "peak" to "off peak" flights. In their studies, it is assumed that the consumers can only learn their time preferences right before the departures and different customers have different opportunity costs of waiting. Dana (1998) states that in a perfectly competitive market, there might be advance purchase discounts.

However, there is not much empirical research about yield management in the airline industry. The main reason might be that the load factor data is not easy to get. Only a limited number of empirical studies try to testify and analyze these theories. Puller et al. (2009) use a census of ticket transactions from one computer reservation system to study the relationship between fares, ticket characteristics and flight load factors. They find mixed support for the scarcity pricing theories. Escobari and Gan (2007) employ a panel data analysis and find empirical support for the capacity-based theory. By developing an effective cost of capacity (ECC) model, they show that higher ECC would lead to higher prices. Also, they find that the effect from ECC on price is higher in competitive markets.

Escobari (2012) further shows that the fares decrease until about two weeks before departure and then increase, holding inventories constant. Chapter IV in this thesis overcomes this difficulty by scripting the online data from *priceline.com*. Using this unique dataset, we are able to have daily dynamic prices and its corresponding dynamic load factors, which enable us to analyze both revenue efficiency and dynamic pricing behaviors. Due to computational restrictions, we only script the data from top 10 Chicago based metropolitan city routes.

In Chapter IV, we first analyze the revenue efficiency of the flights using stochastic frontier analysis method. Then we divide the whole sample into high-efficiency flights group and low-efficiency flights group based on the revenue efficiency level. Last but not the least, we compare the high-efficiency flights' dynamic pricing patterns with the low-efficiency flights' patterns. Based on their differences, we give policy applications to the low-efficiency flights.

The main findings in Chapter IV are illustrated as follows. First, we find evidence of differences in revenue efficiency among different flights. Second, the dynamic pricing patterns of the high-efficiency flights are different from those of low-efficiency flights. Third, we find weak evidence for capacity-based theories and stronger evidence for time-based theories.

To sum up, we conduct a comprehensive study of the market power, cost efficiency and price dispersion, revenue efficiency and dynamic pricing strategies. In Chapter II, we build up a theoretical framework to allow for simultaneous estimations of marginal cost efficiency and conduct parameter without total cost data. Based on the estimates of conduct parameter and marginal cost efficiency, we analyze the factors that influence price dispersion in Chapter III. We mainly focus on the determinants of conduct parameter and marginal cost efficiency. Chapter II and Chapter III employ the data from DB1B, which is post sale data. In Chapter IV, we script the dynamic pricing data from *priceline.com* using *Perl* program, which helps us to understand better how the airlines make dynamic pricing decisions based on available capacity and advanced days purchased (ADP).

CHAPTER II

ESTIMATION OF COST EFFICIENCY WITHOUT COST DATA

2.1 Introduction

A widely used market power measure is the Lerner index (1934), which is the ratio of price-marginal cost mark-up and price. One potential difficulty for calculating the Lerner index is that the total cost data may not be available, which makes estimation of the marginal cost difficult. A potential solution to this problem is estimating a conduct parameter (or conjectural variations) game[1] in which the firm form a conjecture about the variation in the other firms' strategies (e.g., output) in response to a change in its own strategy. For given demand and cost conditions, the conjecture corresponding to the observed price-cost margins can be estimated "as-if" the firms are playing a conduct parameter game. In this setting the "implied marginal cost" can be estimated via a supply-demand system.

Stochastic frontier analysis (SFA) literature relaxes the full efficiency assumption of neoclassical production theory by allowing firms to act suboptimally. Among others, one potential reason for inefficiency is the principle agent problem that the objectives of share-holders and manager are not fully aligned. For example, the manager may use extra staff to reduce manager workloads at expense of higher costs. In SFA literature, cost efficiency is defined as the ratio of minimum cost to actual cost. A standard stochastic frontier model estimates cost inefficiency from a (log-transformed) cost function, which treats the cost inefficiency as an unobserved non-negative error term. The resulting model would have a composed error term consisting of a one-sided error term, which captures cost inefficiency, and a conventional two-sided error term. Hence, the SFA literature[2] suffers from the same

[1] See Perloff, Karp, and Golan (2007) and Bresnahan (1989) for more details on conduct parameter approach.

[2] See Kumbhakar and Lovell (2000) for a book-length survey on SFA and Sickles (2005) for a simulation study examining the performances of some estimators in the SFA literature.

problem that market power literature suffers. That is, it requires the total cost data in order to estimate the cost efficiencies of firms. We overcome this issue by introducing a conduct parameter game, which enables us to estimate “implied cost efficiencies” and conduct parameters jointly without using the total cost data. Hence, in contrast to the SFA literature, which infers the cost efficiency from a cost function, we estimate the cost efficiency from a supply-demand system, which is derived from a conduct parameter game. To the best of our knowledge, this is the first study that enables estimation of cost efficiency that doesn't require the total cost data.

The Quiet Life Hypothesis” (QLH) by Hicks (1935) and “the Efficient Structure Hypothesis” (ESH) by Demsetz (1973) are two well-known hypotheses that relate market power to efficiency. The former claims that higher competitive pressure is likely to force management work harder, which in turn increases efficiencies of firms. The latter states that the firms with superior efficiency levels use their competitive advantages to gain larger market shares, which leads to higher market concentration and thus higher market power. The findings of Berger and Hannan (1998) and Kutlu and Sickles (2012) support the QLH for the banking and airline industries, respectively. However, Maudos and Fernández de Guevara (2007) show evidence for ESH for the banking industry. Moreover, Delis and Tsionas (2009) are in favor of the QLH on average but they also mention that for the highly efficient banks the relationship reverses in favor of the ESH. Hence, the relationship between market power and efficiency has long been acknowledged by economists. However, the market power and SFA literatures largely ignore this relationship.[3] This can potentially cause inconsistent parameter estimates for both conduct parameter and SFA models. For example, consider a market in which the true efficiency levels of the firms are the same but the researcher does not control for firm specific market power when estimating efficiencies of the firms. The differences in market powers would lead to different firm behavior and this can be confused with the firm level cost inefficiency. Generally, efficiencies are measured by closeness of production

[3] Koetter, Kolari, and Spierdijk (2012), Delis and Tsionas (2009), Koetter and Poghosyan (2009), and Kutlu and Sickles (2012) exemplify some studies that attempt to estimate market powers of firms in a framework where firms are allowed to be inefficient. Except for Delis and Tsionas (2009) the market power estimates in these studies are conditional on efficiency estimates.

units to the best-practice units observed in the market. If the firm level conducts affect the performance of the best-practice units, then the efficiency estimates which do not take this into account would not be accurate. For example, in a market facing a Cournot competition the best practicing firm may not really be fully efficient. Lee and Johnson (2012) show that, in the Cournot environment, inefficiency may in fact be a result of endogenous prices and the effect of output production on price. Similarly, ignoring inefficiencies of firms in a conduct parameter model can lead to an omitted variable bias. Our methodology aims to overcome these difficulties by explicitly and simultaneously modeling a conduct parameter game in an environment where firms are allowed to be inefficient.

Another estimation problem is related to calculation of dead-weight-loss (DWL). Ignoring the inefficiencies of productive units may invalidate standard DWL calculations since DWL from collusive behavior depends on inefficiency levels.[4] If the productive units exhibit inefficiency that is misinterpreted as firm heterogeneity, then the standard calculations of DWL may not be valid. In such cases, Kutlu and Sickles (2012) recommend using what they call the efficient full marginal cost (EFMC) for the markup calculation.[5] Hence, the simultaneous estimation of conduct and efficiency may provide us more precise estimates for DWL.

A common problem is that of measuring inefficiencies and market powers of firms when the firms face optimization constraints.[6] For example, a firm which is seemingly inefficient may actually be relatively more efficient if we take the optimization constraints (e.g., capacity constraints) into account. Standard stochastic frontier models do not explicitly model such optimization constraints which may result in inaccurate efficiency estimates. In order to overcome this difficulty, we present an extension of our model which makes it possible to get efficiency and market power estimates jointly in the presence of capacity constraints.

[4] See Comanor and Leibenstein (1969) and Kutlu and Sickles (2012) for more details about calculation of DWL when firms are inefficient.

[5] They define EFMC as the sum of a shadow cost and efficient marginal cost calculated from stochastic frontier analysis techniques. The shadow cost reflects the constraints that the firms face such as capacity or incentive compatibility constraints. Efficient marginal cost is the marginal cost when the firm achieves full efficiency.

[6] See Puller (2007) for a conduct parameter model measuring market powers of firms in the California electricity market under capacity constraints.

We apply our methodology to estimate the firm-route-quarter specific conducts and marginal cost efficiencies of the U.S. airlines for routes that originate from Chicago. The time period that our data set covers is 1999I-2009IV. One of the difficulties that empirical researchers face is that the available cost data set is for the entire U.S. system. So, route level total cost data is not available. Kutlu and Sickles (2012) try to overcome this problem by incorporating a specific number of enplanements for each airline, a specific distance of each city-pair, and airline fixed effects when estimating the cost function. This enables them to calculate the route specific marginal costs from the cost function estimation. However, their efficiency estimates are still firm-quarter specific.[7] Moreover, their conduct estimates are conditional on efficiency estimates. That is, they first estimate the efficiencies using the standard stochastic frontier models; and use these efficiency estimates when estimating the supply relation. In contrast to their study, we jointly estimate the firm-route-quarter specific conducts and efficiencies of the U.S. airlines; and when doing so we do not need route specific total cost data.[8] Our results suggest that concentration ratio (measured by CR_4) and market share of airlines are negatively related to the marginal cost efficiency. In contrast to this, the concentration ratio and market share of airlines are positively related to the conduct.

The rest of Chapter II is structured as follows. In Section 2.2, we build up our theoretical model. In Section 2.3, we describe our data set. In Section 2.4, we present our empirical model. In Section 2.5, we present and discuss our results. In the Section 2.6, we make our concluding remarks for this chapter.

2.2 Theoretical Model

In this section we describe our theoretical framework, which enables us to estimate marginal cost efficiencies and conducts of firms without total cost data. The stochastic frontier

[7] Since Kutlu and Sickles (2012) estimate an aggregate model (i.e., route specific market power), they use market share weighted efficiency estimates in their estimations. That is why their route specific efficiency variables are not the same for different routes.

[8] Similar to our study, Delis and Tsionas (2009) simultaneously estimate bank conducts and efficiencies. However, their model requires total cost data. Hence, as it stands, their methodology is not applicable for our airline example.

literature relaxes full efficiency assumption of neoclassical production theory by allowing the firms to be inefficient. The inefficiency is treated as an unobserved component which is captured by a one-sided error term. The total cost of firm i at time t is given by:

$$C\left(q_{it}; X_{c,it}\right) = C^{*}\left(q_{it}; X_{c,it}\right) \exp\left(u_{it} + v_{it}\right) \tag{1}$$

where q_{it} is the quantity of firm i at time t; $X'_{c,it}$ is a row vector of variables related to cost; $u_{it} \geq 0$ is a term which is capturing the inefficiency; v_{it} is the conventional two-sided error term; and C^{*} is the deterministic component of cost when firms achieve full efficiency. In the conventional stochastic frontier framework the cost efficiencies of firms would be estimated by using the following model:

$$\ln C\left(q_{it}; X_{c,it}\right) = \ln C^{*}\left(q_{it}; X_{c,it}\right) + u_{it} + v_{it}. \tag{2}$$

Figure 2.1 shows a 2-input and 1-output example. The inputs in the figure are labeled as x_1 and x_2. The points in this figure represent input bundles. The curve labelled y^0y^0 is the efficient frontier of the input requirement set for producing output y^0. In this curve any further equi-proportionate reduction of inputs would make the output y^0 infeasible. The point P represents the actual input and the AB line is its corresponding iso-cost line. The points below this line are less expensive than input bundle P. The point Q is the cost-minimizing bundle and the CD line is its corresponding iso-cost line. The cost efficiency is defined as $\left|OR\right| / \left|OP\right|$.

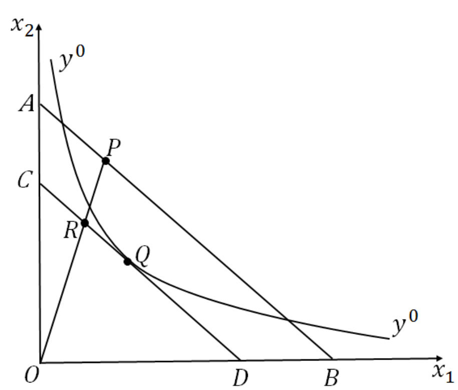

Figure 2.1. Cost efficiency

A variety of distributions is proposed for u_{it} including the half normal (Aigner, et al., 1977), exponential (Meeusen and van den Broeck, 1977), truncated normal (Stevenson, 1980), and gamma (Greene, 1980a, 1980b, 2003) distributions. The cost efficiency of a firm, EFF_{it}, is estimated by:

$$\begin{aligned} EFF_{it} &= \exp\left(-\hat{u}_{it}\right) \\ \hat{u}_{it} &= E\left[u_{it} \mid u_{it} + v_{it}\right]. \end{aligned} \tag{3}$$

The stochastic frontier approach that we presented above requires a detailed cost data set which many times is not available. We utilize the conduct parameter approach in order to overcome this issue. For this purpose, instead of modelling total cost as in the conventional SFA models, we directly model marginal cost, c, as follows:

$$\ln c\left(q_{it}; X_{c,it}\right) = \ln c^{*}\left(q_{it}; X_{c,it}\right) + u_{it} + v_{it}. \tag{4}$$

Here, rather than estimating a cost function, we estimate a supply-demand system, which enables us to calculate the implied cost efficiency. We call c^{*} efficient marginal cost (EMC), which is equal to the marginal cost when firms achieve full efficiency. Under constant marginal cost assumption[9] Equation (4) can directly be derived from Equation (1). Hence, in this case the theoretical efficiency values for these two approaches coincide. When the marginal cost is not constant, we directly model marginal cost efficiency through Equation (4) and remain silent about the way in which the inefficiency is modelled in the cost function. Therefore, the constant marginal cost assumption is not required in our theoretical model. Nevertheless, from the antitrust point of view, which is concerned with market power and DWL estimations, the marginal cost efficiency seems to be a more relevant efficiency concept.[10]

Let $P_t = P\left(Q_t; X_{d,t}\right)$ be the inverse demand function, Q_t be the total quantity, and $X'_{d,t}$ is a row vector of demand related variables at time t. The perceived marginal revenue

[9] Whenever we refer to constant marginal costs we also asssume that $\frac{\partial(u_{it}+v_{it})}{\partial q} = 0$, which would be consistent with the constant marginal cost assumption.

[10] The reason may be clearer when we introduce Figure 2.2 later on in this section.

(PMR) is given by:

$$
\begin{aligned}
PMR\left(\theta_{it}\right) &= P_t + \frac{\partial P_t}{\partial Q_t}\frac{\partial Q_t}{\partial q_{it}} q_{it} \\
&= P_t\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right)
\end{aligned} \tag{5}
$$

where $s_{it} = \frac{q_{it}}{Q_t}$ is the market share of firm i at time t; $E_t = -\frac{\partial Q_t}{\partial P_t}\frac{P_t}{Q_t}$ is the (absolute value of) elasticity of demand; $\theta_{it} = \frac{\partial Q_t}{\partial q_{it}}$ is the conduct parameter. Three benchmark values for $\theta_{it} = \left\{0, 1, \frac{1}{s_{it}}\right\}$ correspond to perfect competition, Cournot competition, and joint profit maximization, respectively. The supply relation is:[11]

$$
\begin{aligned}
P_t\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) &= c_{it} \Leftrightarrow \\
\ln P_t + \ln\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) &= \ln c_{it}
\end{aligned} \tag{6}
$$

where $c_{it} = c\left(q_{it}; X_{c,it}\right)$. After including the econometric error terms, the supply relation becomes:[12]

$$
\begin{aligned}
\ln P_{it} &= -\ln\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) + \ln c_{it} + \varepsilon_{it}^{s} \\
&= g\left(\theta_{it}, s_{it}, E_t\right) + \ln c_{it} + \varepsilon_{it}^{s}. \\
&= \ln c_{it}^{*} + g_{it} + u_{it} + \varepsilon_{it}^{s}
\end{aligned} \tag{7}
$$

where $g_{it} = -\ln\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) \geq 0$ is the market power term and $u_{it} \geq 0$ is the inefficiency term.[13] The E_t term is identified through the demand equation. We assume that the demand function and marginal cost functions are so that the conduct parameter and marginal cost can be separately identified.[14] Intuitively, Equation 5 suggests that if c_{it} and q_{it} are highly collinear, then the conduct parameter maybe identified through the variation in $\frac{\partial P_t}{\partial Q_t}$. A common approach to achieve identification is assuming a constant marginal cost.[15] Our

[11] Note that perceived marginal revenue must be positive so that the equilibrium makes sense. Hence, we assume that $1 - \frac{s_{it}}{E_t}\theta_{it} > 0$. So, we have $\ln\left(1 - \frac{s_{it}}{E_t}\theta_{it}\right) \leq 0$.

[12] The introduction of the error term enables us to deviate from a single market price. Also, the price may be considered to be a function of firm specific variables, $X_{d,it}$.

[13] Note that g_{it} is an increasing function of θ_{it}.

[14] For details about the identification conditions for conduct parameter models, we direct the reader to Bresnahan (1982), Lau (1982), Perloff, Karp, and Golan (2007), and Perloff and Shen (2012).

[15] A constant marginal cost function does not depend on quantity but it may still depend on variables other than the quantity.

model is different from the standard market power models due to the additional u_{it} term. This inefficiency term is identified by utilizing the asymmetric distribution of the variations of u_{it}. Intuitively, u_{it} is identified if the signal-to-noise ratio (the variance ratio of the inefficiency component to the composite error) is not small. Hence, the identification of model parameters requires the standard conduct parameter model and SFA identification assumptions. A standard conduct parameter model ignores u_{it}, which would likely be correlated with g_{it}. Hence, any conduct parameter model that is ignoring u_{it} risks getting inaccurate market power estimates.

Figure 2.2 aims to illustrate the underlying mechanism of our model and consequences of ignoring inefficiency when calculating DWL. This figure includes inverse demand function, perceived marginal revenue (PMR), marginal revenue (MR) that is corresponding to monopoly scenario, marginal cost (MC), and efficient marginal cost (EMC). For illustrative purposes we consider the same constant marginal costs, conducts, and efficiencies for each firm. P_θ and Q_θ are the equilibrium price and quantity at conduct level θ. Similarly, P_C and Q_C are price and quantity for the perfect competition scenario, in which conduct equals 0. In the figure it is assumed that under perfect competition there would be no inefficiency and thus the efficient marginal cost and marginal cost coincide for this case. If QLH holds, then as the market power, measured by θ, increases MC diverges from EMC. In our framework, the marginal cost efficiency is defined as EMC/MC. The social welfare loss at conduct level θ would be equal to the shaded area (sum of dark and light shaded regions). In Figure 2.2, the efficiency is roughly 60%, which is relatively low. As a consequence the social welfare loss due to inefficiency is substantial. The conventional DWL value, which is ignoring inefficiency, is given by the dark shaded triangular area; and is much smaller than the overall social welfare loss. In general, when both conduct and marginal cost efficiency are low, the conventional DWL values would be much lower than overall social welfare losses.

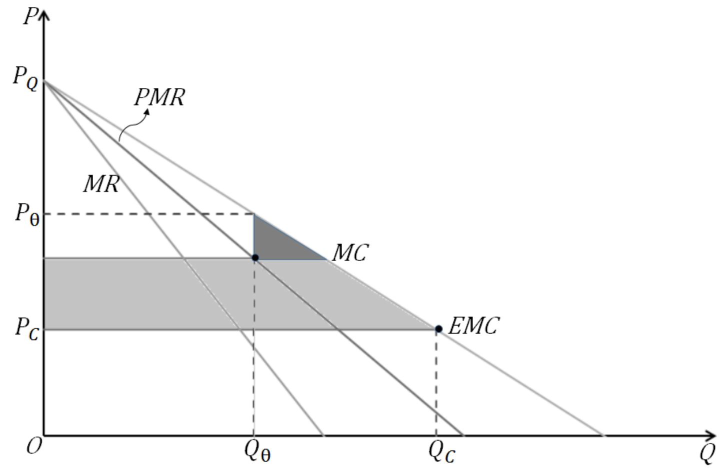

Figure 2.2: Conduct, marginal cost efficiency, and social welfare

Now, we describe how this conduct parameter game would be estimated. We assume that the conduct parameter θ_{it} is a function of variables, $X_{g,it}$, that affect firm specific market power such as market shares and concentration ratios. Modeling θ_{it} through this function may lead to computational difficulties. An arguably better way, which we prefer to follow, would be directly modeling g_{it} as a function of $X_{g,it}$ and solving for θ_{it} after getting the parameter estimates. That is:

$$\hat{\theta}_{it} = \frac{\hat{E}_t}{s_{it}} \left(1 - \exp\left(-\hat{g}_{it}\right)\right) \tag{8}$$

where $\hat{E}_t$ and $\hat{g}_{it}$ are the estimates for E_t and g_{it}, respectively. The market power term, g_{it}, is bounded by 0 and $B_{it} = -\ln\left(1 - \frac{1}{E_t}\right)$. Hence, the choice of functional form should be so that $g_{it} \in [0, B_{it}]$. In this study we use:

$$g_{it} = \frac{B_{it} \exp\left(X'_{g,it}\beta_g\right)}{1 + \exp\left(X'_{g,it}\beta_g\right)}. \tag{9}$$

One of the drawbacks of the standard stochastic frontier models is that if the regressors are correlated with v_{it} or u_{it}, then the parameter and efficiency estimates are inconsistent. Moreover, in this setting, v_{it} and u_{it} terms are assumed to be independent, which can be a questionable assumption in a variety of settings. We use a control function approach to handle the endogeneity issue that occurs when the v_{it} is correlated with the regressors or u_{it}.

For example, u_{it} can be a function of regressors (e.g., market shares of firms or concentration ratios) that are correlated with v_{it} term.[16] The idea is including a bias correction term in the model. Consider the following supply relation model with endogenous explanatory variables:

$$
\begin{aligned}
\ln P_{it} &= \ln c_{it}^{*} + g_{it} + u_{it} + \varepsilon_{it}^{s} \qquad (10)\\
X_{en,it} &= \zeta_{it}'\delta + v_{it}\\
\begin{bmatrix} \tilde{v}_{it} \\ \varepsilon_{it}^{s} \end{bmatrix} &\equiv \begin{bmatrix} \Sigma_v^{-1/2} v_{it} \\ \varepsilon_{it}^{s} \end{bmatrix} \sim \mathbf{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_m & \rho\sigma_{\varepsilon} \\ \rho'\sigma_{\varepsilon} & \sigma_{\varepsilon}^{2} \end{bmatrix} \right)\\
u_{it} &= h_{it}\tilde{u}_{it}\\
h_{it} &\geq 0\\
u_{it} &\sim \mathbf{N}^{+}\left(\mu_u, \sigma_u^2\right)
\end{aligned}
$$

where P_{it} is the price; $X_{en,it}$ is an $m \times 1$ vector of all endogenous variables used in modelling c_{it}^{*}, g_{it}, and u_{it}; $\zeta_{it} = I_m \otimes Z_{it}$ where Z_{it} is a $l \times 1$ (with $l \geq m$) vector of all exogenous variables. The irregular term ε_{it}^{s} is correlated with the regressors but conditionally independent from the inefficiency term u_{it} given $X_{en,it}$ and Z_{it}.[17] Note that ε_{it}^{s} and u_{it} may still be correlated unconditionally. By applying a Cholesky decomposition of the variance-covariance matrix of $\begin{bmatrix} \tilde{v}_{it}' & \varepsilon_{it}^{s} \end{bmatrix}'$, we can rewrite the supply equation as follows:

$$
\begin{aligned}
\ln P_{it} &= \ln c_{it}^{*} + g_{it} + \sigma_{\varepsilon}\rho'\tilde{v}_{it} + \tilde{\varepsilon}_{it}^{s} + u_{it} \qquad (11)\\
&= \ln c_{it}^{*} + g_{it} + \eta'(X_{en,it} - \zeta_{it}'\delta\ ) + \tilde{\varepsilon}_{it}^{s} + u_{it}
\end{aligned}
$$

where $\tilde{\varepsilon}_{it}^{s} \sim \mathbf{N}(0, (1-\rho'\rho)\sigma_{\varepsilon}^{2})$ and $\eta = \sigma_{\varepsilon}\rho'\Sigma_v^{-1/2}$. The parameters of this supply relation can be estimated in one stage using the maximum likelihood estimation method. However, sometimes it is simpler to get the consistent parameter estimates in two stages by first estimating the bias correction term $\eta'(X_{en,it} - \zeta_{it}'\delta\ )$; then including the estimate of bias correction term in the second stage in which we apply traditional SFA methods. For the

[16] See Kutlu (2010) and Karakaplan and Kutlu (2015) for control function solutions to the endogeneity problem. Also see Guan et al. (2009) and Tran and Tsionas (2013) for GMM based stochastic frontier approaches that aim to handle endogeneity.

[17] We may replace $u_{it} = h_{it}\tilde{u}_{it}$ assumption by $u_{it} = h_{it}\tilde{u}_{i}$ so that $\tilde{u}_{i}$ is a firm specific term. This would be in line with the panel data frameworks.

two-stage approach the standard errors need to be corrected, e.g., by a bootstrap procedure. In our empirical section we use the limited information maximum likelihood estimator that we presented in this section, i.e., the one stage method.

The model that we introduced can be extended to a setting in which firms have capacity constraints. This extension of our model is inspired by the conduct parameter model proposed by Puller (2007). In the presence of capacity constraints the optimization problem for firm i becomes:

$$\max P_{it}q_{it} - C_{it} \text{ s.t. } q_{it} \leq K_{it} \tag{12}$$

where K_{it} is the capacity constraint that firm i is facing at time t. Then, the corresponding supply relation becomes:

$$\ln P_{it} = \ln c^*_{it} + g_{it} + \lambda_{it} + u_{it} + \varepsilon^s_{it} \tag{13}$$

where $\lambda_{it} \geq 0$ is the shadow cost of capacity which can be estimated by including variables capturing extent of capacity constraints. For example, Puller (2007) uses a dummy variable which is equal to one when the constraint is binding.

Finally, a formal treatment of conduct parameter games in which the strategic interactions of the firms are dynamic is beyond the scope of this study. Following Puller (2009), we recommend including time fixed-effects which may condition out the dynamic effects in firms' optimization problems.[18] Note, however, that even though the estimates of parameters (including parameters of the conduct and efficiency) are consistent in this dynamic game scenario, we cannot separately identify the efficient marginal costs, c^*_{it}, and dynamic correction terms. The reason is that the time dummies not only capture cost related unobserved factors that change over time but also the dynamic correction terms.[19] Nevertheless, except the portion of time fixed-effects that contribute to c^*_{it}, the other parameters of c^*_{it} are identified. Moreover, many times c^*_{it} is not the main interest. In what follows we assume a static model.

[18] See Puller (2009) for further details about his model and restrictions. One particular assumption that Puller (2009) makes is that the firms play an efficient super-game equilibrium when they cooperate. That is, they maximize the joint profit subject to incentive compatibility constraints. Hence, the corresponding efficient super-game equilibrium values are benchmark for the full market power case.

[19] Of course, in the static setting we don't have this identification issue as the dynamic correction terms are zero.

2.3 The Data

In order to testify our theoretical framework, we use the U.S. domestic airline data. One of the main data sources that we use is the Passenger Origin-Destination Survey of the U.S. Department of Transportation (DB1B data set). This data set is a 10% random sample of all tickets that originate in the U.S. on domestic flights. In our data set a market is defined as a directional city-pair (route). Calculation of prices and quantities are based on the tickets that have no more than three segments in each direction. About 1% of tickets are eliminated during the elimination of tickets with more than 3 segments. We only focus on coach class tickets due to the differences in demand elasticities and other characteristics between coach class and high-end classes (first class and business class).

Our data set covers the time period from the first quarter of 1999 to fourth quarter of 2009. During this time period, the U.S. airlines face serious financial problems. As pointed out by Duygun, Kutlu, and Sickles (2014), the financial losses for the domestic passenger airline operations in this time period is substantially higher than their losses between 1979 and 1999. Increase in taxes and jet fuel prices, relatively low fares, and sharp decrease in demand are some of the challenging properties of this period for the U.S. airlines. In this time period, we observe dramatic increases in load factors. Borenstein (2011) argue that such an increase might be attributed to improved yield management techniques.

Now, we provide the details about data construction process. First, all multi-destination tickets are dropped as it is difficult to identify the ticket's origin and destination without knowing the exact purpose of the trip. Second, any itinerary that involves international flights is eliminated. Third, we adjust the fare class for high-end carrier. That is, for some airlines, due to marketing strategies, only first class and business class (high-end) tickets are provided to consumers on all routes. However, the quality should be taken as coach class. So, we consider all tickets as coach class tickets if there is no coach class tickets from certain carrier in given quarter. In different time periods, due to changes in the pricing strategy, sometimes high-end-only carrier switches to a regular carrier which sells both coach class tickets and high-end tickets. For instance, Sun Country Airlines does not provide coach class tickets in 2001 but provides coach tickets in 2005 and years after. So, we treat the tickets in each quarter separately when considering this adjustment. That is, we treat high-end tickets from Sun Country Airlines as coach class tickets in 2001, not in year 2005. Fourth, tickets that have high-end segments and unknown fare classes are dropped.

We followed Borenstein (1989) and Brueckner, Dyer, and Spiller (1992) by using ticketing carrier as our airline as an observation unit. After further elimination of multi-ticketing-carrier tickets, firm specific average segment numbers (SEG) and average stage length (SL) on a given route are calculated as indicators of quality and costs. Moreover, our data set includes a distance variable which is the shortest directional flight distance ($DIST$). A ticket is online when one-way ticket does not involve change of airplanes. The online variable is the percentage of online tickets.

For the price variable, we use the average price of all tickets for a given airline on a given route in given quarter. All tickets with incredible prices are dropped from our data set. Following Borenstein (1989) and Ito and Lee (2007), we eliminate the open-jaw tickets since it would be difficult to distribute the ticket price into outbound and inbound segment for open jaw tickets. We drop the tickets that have a price less than 25 dollars or higher than 99 percentile or more than 2.5 times standard deviation from the mean for each airline within a route. The tickets that have price less than 25 dollars are considered as frequent flyer program tickets and the tickets that have prices higher than 99 percentile are considered to be input (key punch) errors for the data set. For the round trip tickets, we divided the total price by two to get the one-way price.

The cost data set is constructed from the firm level data of DOT's airline production data set (based on Form 41 and $T100$). We control for three types of important costs: labor price (LP), energy price (EP), and capital price (KP). The salaries and benefits for five main types of personnel are provided in Form 41/P6. Annual employee number is given in Form 41, P10. We interpolated the annual employee data to get the quarterly values. For energy price, we only capture the cost based on aircraft fuel. The energy input is developed by combining information on aircraft fuel gallons used with expense data per period. Flight capital is described by the average size (measured in number of seats) of the fleet. The number of aircraft that a carrier operated from each different model of aircraft in airline's fleet is collected from DOT Form 41. For each quarter, the average number of aircraft in service is calculated by dividing the total number of aircraft days for all aircraft types by the number of days in the quarter. This serves as an approximate measure of the size of fleet.

In order to estimate the demand, we also include the city specific demographic variables: per capita income (PCI) and population (POP). We get the city level per capita income and population data from Bureau of Labor Statistics. We interpolate the annual data to get the quarterly PCI and population variables for each city. For each origin-destination city-pair, we use the population weighted PCI as the route-specific PCI measure. Similarly, city-pair population is the average population of origination and destination cities. In order to get the real prices, we deflated the nominal prices by Consumer Price Index (CPI) and use the first quarter of 1999 as the base time period. Since only the metropolitan areas have the demographic information but some airports are located in small cities, the number of the city-pairs is further reduced in our final database.

We apply our theoretical method on the routes that originate from Chicago, which is a popular choice because of its relatively large airport and wide selection of airline firms. For instance, Brander and Zhang (1990) use 33 Chicago-based routes in their studies. In our final data set, we further eliminate the small firms and small routes. On a given route, firms with market shares less than 0.01 are eliminated. For a given quarter, any route with enplanements less than 1800, i.e., 20 passengers per day, is dropped. Also, we eliminate routes with less than 30 observations. Table 2.1 presents the summary statistics for Chicago-based routes.

Table 2.1. Summary for Chicago Originating Routes

Variables: ***Name in Estimation***	**Mean**	**S.D.**	**Min**	**Max**
Low Cost Carrier Dummy: LCC	0.11225	0.31569	0	1
ln(Population): ln(POP)	15.50922	0.18392	15.33637	16.46417
ln(Per Capita Income): ln(PCI)	10.4226	0.03828	10.33827	10.61397
ln(Stage Length): ln(SL)	6.40837	0.59661	4.69135	8.34378
ln(Distance): ln(DIST)	6.63774	0.65719	4.69135	8.35303
ln(Average Fleet Size): ln(SIZE)	4.97262	0.08467	4.80562	5.33165
ln(Labor Price): ln(LP)	9.23198	0.7346	5.58453	10.00725
ln(Capital Price): ln(KP)	7.27984	0.76111	3.6658	8.1754
ln(Fuel Price): ln(FP)	3.87272	0.58681	2.37945	4.83215
Number of Firms: NF	8.2814	3.14474	1	24
ln(Price Per Ticket): ln(P)	4.96474	0.33975	3.72431	7.1672
ln(Number of Passengers): ln(Q)	9.79688	1.43261	7.49554	13.1454
ln(Number of Passengers for Other Routes): ln(OQOTH)	15.29761	0.10457	14.98133	15.47699
Market Share: s	0.22057	0.22458	0.01001	1
Geometric Market Share: GEOS	0.19418	0.12814	0.00682	1
Geometric Market Share*Number of Firms: GEONF	1.08933	0.66467	0.03784	4.27413
Online Rate: ONLINE	0.66744	0.36844	0	1
Top 4 Concentration Ratio: CR4	0.9195	0.11333	0.07516	1
ln(Average Number of Segments): ln(SEG)	0.3698	0.32171	0	1.09861
Number of Observations	18209			

Low cost carrier is a dummy variable that equals 1 if the ticketing carrier is a low cost carrier, otherwise it is 0. Number of firms represents the total number of firms that operate on the route. Total number of passengers is the total tickets sold on a given route by all airlines together in a given quarter. Total number of passengers for other routes ($OQOTH$) variable is the total number of tickets that are sold on the other routes that share the same origination city. We use the geometric market share ($GEOS$) variable of Gerardi and Shapiro (2009) as an instrument.[20] Another instrumental variable that we use is $GEONF$. This variable is the product of $GEOS$ and $nf^* = \sqrt{n_o * n_d}$, where n_o denotes to the mean value of number of firms for all routes that share the same origination as route of interest while the n_d refers to the mean value of number of firms for all routes that share the same destination city.

The final data set contains 108 routes that originate from Chicago and 14 carriers. The low cost carriers are Frontier Airlines, JetBlue Airways, Southwest Airlines, and Spirit Airlines. The rest of the carriers are: Alaska Airlines, American Airlines, Continental Airlines,

[20] $GEOS$ is the $GENSP$ variable that is used in Gerardi and Shapiro (2009). $GENSP$ is similar to the $GEOSHARE$ variable of Borenstein and Rose (1994). The difference is that Borenstein and Rose (1994) use average daily enplanements while we use average quarterly enplanements.

Delta Airlines, Northwest Airlines, United Airlines, US Airways, America West Airlines, ATA Airlines and Trans World Airways.

2.4 Empirical Model

In this section, we want to shed some light on the market powers and efficiencies of the U.S. airlines. Moreover, we aim to illustrate our methodology using this empirical example. In particular, we estimate time-varying firm-route-specific conducts and marginal cost efficiencies of the U.S. airlines. Similar to Brander and Zhang (1990) and Oun, Zhang, and Zhang (1993), we only consider coach class tickets. Our city-pair markets consist of one-way or round-trip directional trips in each direction for three-segment (up to three segments in each direction) data set. We divide the total ticket price by 2 to get the one-way fare for round-trip tickets. The demand and supply equations are estimated separately. The demand equation is given by:

$$\ln P_{itr} = \beta_0 + \beta_1 \ln Q_{tr} + \beta_2 \ln PCI_{tr} \ln Q_{tr} + f_d\left(X_{d,itr}\right) + \varepsilon^d_{itr} \tag{14}$$

where f_d is a function of demand related variables, Q_{tr} is the total quantity at time t for route r, $X'_{d,itr}$ is a row vector of demand related variables, and ε^d_{itr} is the conventional two-sided error term. We assume that $\ln Q_{tr}$ and $\ln PCI_{tr} \ln Q_{tr}$ are endogenous. Along with the exogenous variables included in the demand model, our instrumental variables are $GEOS_{itr}$, $GEONF_{itr}$, $\ln OQOTH_{rt}$, logarithm of labor price ($\ln LP_{it}$), logarithm of capital price ($\ln KP_{it}$), and logarithm of energy price ($\ln EP_{it}$).

The supply equation is given by:

$$\ln P_{itr} = \ln c^*_{itr} + g_{itr} + u_{itr} + \varepsilon^s_{itr} \tag{15}$$

where c^*_{itr} is the marginal cost when firms achieve full efficiency, $g_{itr} = -\ln\left(1 - \frac{s_{itr}}{E_{tr}}\theta_{itr}\right)$ is the market power term, $u_{itr} \geq 0$ is the inefficiency term, and ε^s_{itr} is the conventional two sided error term. The parameters of the E_{tr} term is identified through the demand equation. We assume that the efficient marginal cost, c^*_{itr}, is constant, i.e., it is not a function of quantity but maybe a function of exogenous cost shifters. Hence, as we described in the theoretical model section, the theoretical values for cost and marginal cost efficiencies coincide in this model. While constant marginal cost is a relatively strong assumption, it is commonly used in the conduct parameter models. Iwata (1974), Genesove and Mullin (1998), Corts (1999), and Puller (2007) exemplify some papers that use this assumption in a variety of conduct parameter settings. We use this simplifying assumption to illustrate our methodology. Nevertheless, we approximate the efficient marginal cost function by a fairly flexible function of input prices and other cost related exogenous variables. These cost related variables include year, quarter, and firm dummy variables, which capture time-firm-specific unobserved factors.[21] We model g_{itr} as in the theoretical model section and assume that $X'_{g,itr} = [s_{itr}, CR_{4,tr}, \ln DIST_r, t, E_{tr}, 1]$ where $CR_{4,tr}$ is the concentration ratio for largest four firms on route r at time t. We assume that $u_{itr} = h_{itr}\tilde{u}_{itr}$ and $\tilde{u}_{itr} \sim \mathbf{N}^+\left(0, \sigma_u^2\right)$ where $\sigma_u^2 = \exp\left(X'_{g,itr}\beta_u\right)$; and $\varepsilon_{itr}^s \sim \mathbf{N}\left(0, \sigma_\varepsilon^2\right)$ where $\sigma_\varepsilon^2 = \exp\left(\beta_\varepsilon\right)$. For the supply side, s_{itr} and $CR_{4,tr}$ are assumed to be endogenous. Our instrumental variables are $GEOS_{itr}$, $GEONF_{itr}$, $\ln POP_{tr}$, and $\ln PCI_{tr}$. The estimations of the supply relations are done by using the one-stage version of the control function approach that we described in our theoretical model section.

2.5 Results

In this section, we present our estimation results. The demand parameter estimates for the routes originating from Chicago are given in Table 2.2.[22] We estimate the (inverse) demand equation by 2SLS. Our demand model controls for year, quarter, and firm dummies. The demand elasticities are negative at each observation, i.e., $E_{tr} > 0$.

[21] If the airlines are playing a version of dynamic conduct parameters game that is suggested by Puller (2009), route-specific time dummy variables would capture dynamic factors that enter the airlines' optimization problems as well. In this case, the parameter estimates of the model would still be consistent including the conduct parameters and efficiencies. However, the marginal cost estimates may be biased (downwards) as the prediction of marginal costs include these dynamic factors.

[22] The estimation includes a constant term.

Table 2.2. Estimation for Demand Function

Dependent Variable: Price	Estimates	Std. Err.
ln(Q)	2.91356***	(0.50629)
ln(Q)*ln(PCI)	-0.29270***	(0.04861)
ONLINE	0.01940*	(0.00880)
ln(DIST)	7.47912***	(0.95064)
ln(SEG)	-10.03281***	(0.97631)
ln(SIZE)	-15.81152***	(2.31769)
ln(SL)	-8.24091***	(0.97745)
ln(DIST) Square	-0.22445**	(0.07828)
ln(SEG) Square	0.53698***	(0.11695)
ln(SIZE) Square	1.50167***	(0.22554)
ln(SL) Square	-0.62278***	(0.08377)
ln(DIST)* SEG	-0.65458**	(0.19646)
ln(DIST)* ln(SIZE)	-1.93435***	(0.18417)
ln(DIST)* ln(SL)	0.90929***	(0.15754)
ln(SEG)* ln(SIZE)	2.44755***	(0.19223)
ln(SEG)* ln(SL)	0.24996	(0.19551)
ln(SIZE)* ln(SL)	1.98114***	(0.18886)
ln(POP)	0.30783***	(0.01673)
ln(PCI)	4.19634***	(0.56805)
LCC	-0.21325***	(0.00862)
Quarter Dummies	Yes	
Year Dummies	Yes	
Firm Dummies	Yes	
Centered R Square	0.4976	
Number of Observations	18209	

Note: + p<0.1, * p<0.05, ** p<0.01, *** p<0.001.
Robust standard errors are given in bracket.

For the supply function, as we described earlier, we use the one-stage control function approach to deal with endogeneity. We estimate two supply models: the first one allows inefficiency (Benchmark model) and the second one assumes full efficiency so that $u_{itr} = 0$ (Full efficiency model). The full efficiency model is a standard conduct parameter model, which helps us to compare our benchmark estimates with the standard models. Table 2.3 shows the estimation results.

The bias correction terms (η) are jointly significant at any conventional significance level, which is an indication of endogeneity. The median of conduct estimates is 0.63. This value is lower than the theoretical conduct value for Cournot competition, which equals 1.[23] Hence, at the median, the extent of competition lies somewhere between perfect competition and Cournot competition. The median of conduct estimates from the full-efficiency model is

[23] The median of theoretical conduct values for joint profit maximization scenario is 6.77, which is the median of $\frac{1}{s_{itr}}$.

somewhat lower and suggests a competitive market. Low cost carriers, due to their special operating style[24], tend to have lower marginal costs compared to other airlines, which helps them to reduce price. Hence, it is worthwhile to examine the decomposition of conducts based on LCCs and non-LCCs. We observe that the conduct estimates for LCC and non-LCC carriers are 0.24 and 0.74, respectively. Therefore, while the conducts of LCCs are closer to perfectly competitive values, the conducts of non-LCCs are closer to Cournot competition values. Finally, the median of the conduct estimates from the full efficiency model is 0.22, which is somewhat lower than that of the benchmark model.

Table 2.3. Estimation for Supply Function

Supply Function:	**Inefficiency Allowed**		**Full Efficiency**	
Price	**Estimates**	**Std. Err.**	**Estimates**	**Std. Err.**
ln(KP)	-0.72825***	(0.13523)	-0.63778***	(0.15070)
ln(FP)	0.02501	(0.11625)	0.07255	(0.12892)
ln(LP)	-0.34062*	(0.14423)	-0.23348	(0.16666)
ln(KP) Square	0.02449	(0.02310)	0.01674	(0.02643)
ln(FP) Square	0.01010	(0.01061)	0.01876	(0.01164)
ln(LP) Square	0.02082	(0.01929)	0.01612	(0.02175)
ln(KP)* ln(FP)	0.02793+	(0.01550)	0.02603	(0.01717)
ln(KP)* ln(LP)	0.02922	(0.04268)	0.03147	(0.04818)
ln(FP)* ln(LP)	-0.04628**	(0.01768)	-0.05404**	(0.01982)
ONLINE	-0.11235***	(0.00769)	-0.13278***	(0.00853)
ln(DIST)	3.25569***	(0.96877)	1.68620	(1.05468)
ln(SEG)	-3.81246***	(0.95276)	-1.65408	(1.01118)
ln(SIZE)	-15.47651***	(3.14511)	-6.51047+	(3.37444)
ln(SL)	-3.46404***	(1.01353)	-1.90943+	(1.10320)
ln(DIST) Square	0.41059***	(0.07352)	0.35812***	(0.07470)
ln(SEG) Square	0.88206***	(0.11481)	0.81164***	(0.11475)
ln(SIZE) Square	1.63611***	(0.31323)	0.75112*	(0.33609)
ln(SL) Square	0.19687*	(0.08723)	0.11232	(0.09075)
ln(DIST)* ln(SEG)	-2.11495***	(0.18315)	-1.75739***	(0.17727)
ln(DIST)* ln(SIZE)	-0.76913***	(0.19136)	-0.55407**	(0.21109)
ln(DIST)* ln(SL)	-0.51857**	(0.15756)	-0.37967*	(0.16206)
ln(SEG)* ln(SL)	1.93489***	(0.18700)	1.50137***	(0.18227)
ln(SEG)* ln(SIZE)	0.94664***	(0.18911)	0.60553**	(0.20198)
ln(SIZE)* ln(SL)	0.64006**	(0.20052)	0.40393+	(0.22051)
LCC	-0.76658***	(0.07363)	-0.76248***	(0.08575)
Quarter Dummies	Yes		Yes	
Year Dummies	Yes		Yes	
Firm Dummies	Yes		Yes	

[24] For example, some of them operate only on certain routes in order to reduce costs.

Table 2.3. Estimation for Supply Function (continued)

Nonlinear Function: g	**Inefficiency Allowed**		**Full Efficiency**	
	Estimates	**Std. Err.**	**Estimates**	**Std. Err.**
s	17.62885***	(1.77711)	17.86643***	(1.87244)
CR4	8.82176***	(2.35791)	16.89555***	(3.57813)
ln(DIST)	0.38011	(0.33349)	0.47333+	(0.27124)
t	-0.22842***	(0.02998)	-0.23597***	(0.03529)
Elasticity of Demand	-0.13075	(0.28139)	-0.89407**	(0.29823)
Constant	-10.53598*	(5.06577)	-13.87330**	(4.77015)
σ_w				
Constant	-3.46254***	(0.03206)	-2.70253***	(0.01051)
σ_u				
s	0.41774***	(0.08396)		
CR4	3.95924***	(0.28411)		
ln(DIST)	-0.94727***	(0.04206)		
t	-0.03700***	(0.00233)		
Elasticity of Demand	-0.15413***	(0.04344)		
Constant	1.94036***	(0.53855)		
η (bias correction term)				
s	-0.33883***	(0.02233)	-0.33863***	(0.01947)
CR4	0.09003**	(0.02838)	0.58208***	(0.02438)
Log-likelihood	30772.7323		30033.13415	
Number of Observations	18209		18209	

Note: + p<0.1, * p<0.05, ** p<0.01, *** p<0.001.

The conduct parameter estimates show that an airline with higher market share tends to have higher market power. In markets with high CR_4 values, it may be easier for airlines (those with higher market share) to cooperate. The positive coefficient of CR_4 in conduct verifies this. For the time period that we examine, the U.S. airlines seem to be losing market power over time. For longer flight distances the alternative transportation means (e.g., bus or car) are likely to become less attractive to the consumers. This reduction in outside competition suggests a positive relationship between market power and distance. The positive sign of distance variable for the market power term is in line with this intuition.

In our benchmark estimates, the medians of efficiency estimates for the whole sample and non-LCC carriers are 82.6% and 84.4%, respectively. Hence, the efficiencies of LCCs and non-LCCs are similar. The parameter estimates for inefficiency term show that an airline with higher market share, tends to have higher inefficiency. Similarly, higher CR_4 values lead to higher inefficiency. This is in line with the QLH that more market power leads to lower efficiency levels.[25] The medians of price-marginal cost markups, price-efficient marginal cost markups, and prices are \$4.63, \$30.30, and \$142.70, respectively. Historically, the airlines are known to complain about not being able to make much money. These markup values indicate that the airlines may partially be responsible for the financial difficulties that they face. Hence, the airlines can achieve reasonable profit levels if they work harder to improve their efficiency levels.

In Table 2.4, we decompose the sample into two regions based on market shares of airlines: 1) Airlines with market shares smaller than s^* and 2) Airlines with market shares greater than s^*. We consider three different values for $s^* = \{0.05, 0.25, 0.50\}$. The values in the table are the medians of conduct and efficiency estimates from the benchmark model, which correspond to these subsamples. Based on this decomposition, it seems that the airlines with market shares greater than 0.05 act similar to Cournot competitors. On the other hand, the airlines with market shares smaller than 0.05, act more like perfectly competitive firms. We also observe that the airlines with high market shares are much less efficient compared to airlines with smaller market shares. For instance, in the third column the medians of efficiencies are 83.92% and 68.46% for small market share and large market share groups, respectively.

Table 2.4. Conduct and Efficiency by Market Share

	Conduct Parameter			Efficiency		
Market Share	**0.05**	**0.25**	**0.5**	**0.05**	**0.25**	**0.5**
Smaller	0.2	0.2	0.37	85.61	84.71	83.92
Greater	0.98	1.4	1.42	81.18	77.44	68.46

2.6 *Summary and Concluding Remarks*

In this chapter, we provided a conduct parameter based framework to estimate market powers and efficiencies of firms simultaneously. Our methodology enables us to relax the total cost data requirement for the stochastic frontier models. The total cost data may not be available for a variety of reasons. For example, firms might not want to reveal such a strategic information whenever it is possible. Even when some form of total cost data

[25] Note that since u has a half normal distribution its mean depends on σ_u^2. In particular, the mean of u is an increasing function of σ_u^2.

is available, the data may not reflect the total cost of the relevant unit that we want to examine. For instance, for the U.S. airlines only firm specific total cost data is available for the whole U.S. airline system. Hence, the conventional stochastic frontier models cannot estimate route specific efficiencies as this would require route-specific total cost data.

Besides relaxing a vital data requirement, our methodology aims to overcome some estimation issues. Efficiencies are generally measured by the distance between the units of production and the best practice units observed in the market. If the performance of the best-practice units depends on their market powers, then the efficiency estimates not taking this into account would not be accurate. We overcome this difficulty by explicitly modeling a conduct parameter game in an environment where firms are allowed to be inefficient. Moreover, we provided a simple extension of our model which allows the firms to have capacity constraints.

Researchers may be interested in estimating the market powers and efficiencies of firms in an environment where firms interact repeatedly so that they play a dynamic game.[26] While we did not provide an explicit solution to this problem, it is possible to extend our model to a dynamic setting in which firms play an efficient supergame equilibrium as in Puller (2009). Finally, an extension of our conduct parameter model so that the firms price discriminate is not complicated. Such an extension would enable us to understand the connection between price discrimination, market power, and efficiency better.[27] Hence, our theoretical model serves as a guideline as to how conduct parameter and efficiency can be estimated simultaneously without requiring total cost data; and this guideline can be applied to a variety of conduct parameter settings.

The U.S. airline industry is a good example to apply our methodology due to the unavailability of route specific total cost data. Hence, we applied our methodology to estimate the conducts and marginal cost efficiencies of the U.S. airlines for the time period 1999I-2009IV. We found that the market concentration and market share of airlines are negatively

[26] See Corts (1999) for a simulation study examining the performance of the static version of conduct parameter method in a dynamic environment.

[27] See Borenstein and Rose (1994), Stavins (2001), Gerardi and Shapiro (2009), and Chakrabarty and Kutlu (2014) for works examining the relationship between market power and price discrimination.

related to the marginal cost efficiency, which is in line with the quiet life hypothesis.

CHAPTER III

PRICE DISPERSION, COMPETITION AND EFFICIENCY IN THE U.S. AIRLINE INDUSTRY

3.1 Introduction

Competition can limit the firms' price discrimination abilities and force them to charge a single price. In line with this idea it appears that the extent of price discrimination may be positively related to the market powers of firms. Since different consumer segments can have different demand elasticities, a negative relationship is possible as well. For example, in the airline context, frequent flyer programs play a central role in keeping the business travelers loyal to a particular airline or a group of airlines in the affiliated program. As a consequence, an increase in competition would potentially decrease the price for the discount tickets more compared with the regular tickets that target the business travelers. Therefore, an increase in market power may result in a decrease in price discrimination. Hence, the theoretical direction of relationship between market power and price discrimination is ambiguous.

The empirical studies mostly concentrate on the relationship between market concentration and price dispersion.[1] However, again, there is no consensus on the direction of this relationship. Borenstein and Rose (1994) and Stavins (2001) find a negative relationship between market concentration and price dispersion while Gerardi and Shapiro (2009) find a positive relationship.[2] The structure-conduct-performance (SCP) paradigm argues that the structure of an industry (e.g., market concentration) affect the conduct of firm, and thus, the performance of the industry. In particular, this paradigm argues that there is an inverse

[1] Note that price discrimination is only one of the reasons for price dispersion. We say that price discrimination exists when the variation of prices cannot be entirely explained by variations in marginal costs. Price dispersion may happen when the consumers are not fully informed about prices and/or have limited memories. See, Varian (1980), Chen, Iyer, and Pazgal (2010), and Kutlu (2015a) for models where price dispersion may be present without price discrimination.

[2] See also Chakrabarty and Kutlu (2014) for a study that finds a non-monotonic relationship between market concentration and price dispersion.

relationship between the market concentration and the extent of competition. Hence, if we accept the SCP paradigm and consider market concentration as a proxy for market power, we may interpret the findings of these studies as evidence for relationship between market power and price dispersion.

There are reasons to believe that market concentration measures such as Herfindahl-Hirschman Index (HHI) may not be sufficient to capture market power. For instance, in an oligopoly setting, the actions of competitors, elasticity of demand, and marginal cost efficiency affect the market outcome.[3] The HHI does not provide relevant information about the elasticity of demand and the extent of cost efficiency. Moreover, HHI is a market-specific index, which cannot explain price dispersion among firms. It may make more sense to use a firm-specific market power measure when examining price dispersion. In order to address these concerns, we model price dispersion by conducts and marginal cost efficiencies, which are firm-specific.

In our study cost inefficiency refers to the increases in marginal costs due to the firm's suboptimal actions regarding production. Among others, one reason for inefficiency is the principle agent problem that the objective of manager is not fully aligned with profit maximization. Another reason would be potential optimization mistakes such as misallocation of inputs. If an increase in marginal cost efficiency leads to the same amount of price changes for both high-end and low-end segments, then marginal cost efficiency would be unrelated with price dispersion, measured by differences between high-end and low-end prices. In general, the relationship may also be positive or negative. On the one hand, the high-end customers would have less elastic demand compared with the low-end customers. An increase in marginal cost efficiency might result in a smaller decrease in the high-end prices compared with the low-end prices. Therefore, we can have a positive relationship between price dispersion and efficiency. On the other hand, even the same firm can implement different pricing strategies for different consumer segments and differences in these strategies

[3] A more extensive criticism for HHI as a market power index is provided by Borenstein et al. (1996). They only concentrate on the neoclassical settings where there are no inefficiencies. See Koetter, Kolari, and Spierdijk (2012) and Kutlu and Sickles (2012) for studies that consider the efficiency when measuring market power.

can lead to a negative relationship.[4] Moreover, if the share of the high-valuation customers is relatively large, the brand loyalty of business travelers becomes more important for the airlines. Hence, the airlines may be more reluctant to risk scaring off their loyal customers and losing them to other airlines on such routes. In response to a decrease in marginal cost efficiency, the airlines may be relatively more protective for the brand loyal customers considering potential future profits. Therefore, the effect of marginal cost efficiency on price discrimination can be ambiguous and depend on the particularities of the segment specific strategies. Finally, marginal cost efficiency may be related to the general "ability" of an airline to handle the operations optimally. Hence, an airline that is good at minimizing its costs may also be good at practicing price discrimination strategies. One of the purposes in the present study is identifying the direction of relationship between marginal cost efficiency and price dispersion empirically.

This chapter analyzes the determinants of price dispersion for the U.S. domestic airline market from 1999I to 2009IV. In particular, we focus on routes that originate from Chicago city. Our main findings are given as follows. First, we observe that omitting the marginal cost efficiency from the model may lead to over-estimation of the effect of conduct on price dispersion. However, we did not find strong evidence for over-estimation. Second, the airline conduct has opposite effects on big city and leisure routes. In particular, the airline conduct has a positive effect on price dispersion on big city routes while the effect is negative on leisure routes. Finally, marginal cost efficiency has a significantly negative influence on price dispersion. It turns out that this negative effect is higher on leisure routes. This is consistent with the lower price dispersion levels that we observe on leisure routes.[5] A 10 percentage points increase in marginal cost efficiency would lead to 0.0193 and 0.0328 decrease in Gini coefficient for big city and leisure routes, respectively.[6] This implies that a 10 percentage points increase in marginal cost efficiency would decrease the expected fare

[4] Chi, Koo, and Lim (2013) mention that different firms can have distinct pricing strategies. In our scenario, we suggest the possibility of different pricing strategies for different consumer segments.

[5] The average efficiency on leisure and big city routes are 84.2% and 79.7%, respectively. The difference is statistically significant at any conventional significance level. Hence, the airlines seem to be more careful about their cost minimizations on leisure routes.

[6] The Gini coefficient is equal to twice the expected absolute difference between two ticket prices that are drawn randomly from the population.

difference by 3.9% and 6.6% of the mean fares for two randomly selected tickets.[7]

The remaining parts of Chapter III are organized as followed. We describe the determinants of price dispersion in Section 3.2. In Section 3.3, we describe our airline data set. The empirical models and estimation results are presented in Section 3.4. In Section 3.5, we summarize our findings and make our concluding remarks. In Section 3.6, we present additional details in the Appendices.

3.2 Sources for Price Dispersion

The extent of price dispersion that the air traveler faces is considerable and it is our interest to examine the sources of this dispersion. There are a variety of pricing strategies that can lead to price dispersion such as peak-load pricing or stochastic demand pricing. Peak-load pricing is affected by the shadow cost of capacity. At the peak times, the shadow cost of additional seat is higher compared with off peak times. The stochastic demand pricing is related to the price dispersion due to the demand uncertainty that cannot be estimated by historical data. Also, price discrimination is one important source of price dispersion. However, it may be difficult to isolate these factors when studying price dispersion. In this study the determinants of price dispersion that we consider are: conduct, marginal cost efficiency, other cost based factors, population attributes, and product attributes. In particular, we are interested in how conduct and marginal cost efficiency affect price dispersion.

The studies of Borenstein and Rose (1994) and Gerardi and Shapiro (2009) provide empirical evidence that market concentration and price dispersion are related to each other. Even though these two papers hold different opinions about the direction of the relationship between market concentration and price dispersion, it is clear that market concentration has an effect on price dispersion. The structure-conduct-performance (SCP) paradigm argues that the structure of an industry (e.g., market concentration) affect the conduct of firm, and thus, the performance of the industry. In particular, this paradigm argues that there is an inverse relationship between the market concentration and the extent of competition.

[7] While Hazledine (2006), Kutlu (2009), and Kutlu (2012) do not consider marginal cost effciency in their theoretical Cournot and Stackelberg type price discrimination models, the findings are similar in the sense that a decrease in the marginal cost would increase price discrimination measured by the differences in prices.

In the light of SCP paradigm one may argue that, among others, these two papers present evidence for a relationship between market power and price dispersion. In our empirical part of this chapter, we aim to provide a closer look in to this relationship.

We consider two types of cost related factors: marginal cost of efficiency and other cost related factors. The marginal cost efficiency has both direct and indirect effects on price dispersion. The indirect effect follows from the efficient structure hypothesis (Demsetz, 1973) that more efficient firms gain more market power, which in turn affects the price dispersion. The direct effect is related to the cost reduction due to efficiency increase. In an oligopoly competition environment, a decrease in the marginal cost efficiency of the airline would affect the pricing strategies for high-end and low-end prices in a non-trivial way. The net effect of an increase in marginal cost efficiency on the price dispersion is ambiguous. On the one hand, due to the differences in demand elasticities, an increase in marginal cost efficiency may result in a smaller decrease in the high-end prices compared with the low-end prices. This supports a positive relationship between price dispersion and efficiency. On the other hand, the airline can implement different pricing strategies for different consumer segments. For example, if the share of the high-valuation customers is relatively large, the brand loyalty of business travelers and long term concerns would be relatively more important for the airlines. Therefore, the airline may be more reluctant to risk scaring off these loyal customers, and thus would not increase the high-end prices as a response to decrease in marginal cost efficiency. Moreover, as mentioned earlier, marginal cost efficiency may be one of the indicators of how successful an airline is in terms of implementing price discrimination, which is a source of price dispersion.[8] Also, we control for other cost related factors such as labor price, fuel price, and capital price. Finally, we control for quality-related characteristics that affect the costs. For instance, larger segment number is a relevant indicator for itinerary quality and costs.

Product attributes not only affect the costs of airlines but also affect the valuations of the consumers in a potentially heterogenous way. One of the important product characteristics

[8] The marginal cost efficiency is not a direct measure of price discrimination abilities. However, it may be considered as a proxy for the airline' intrinsic succcess in terms of optimizing its actions for a given objective.

is the number of segments. A higher number of segments implies a higher number of change of planes and more inconvenience for the traveler. Business travelers are less likely to sacrifice convenience for lower price compared to the leisure travelers. Stage length is the average distance of each segment. Longer stage length means more flight time, which is not convenient for the consumers. However, air travelers can accumulate miles and segments for future rewards through frequent flyer programs employed by airlines. This might balance off the direct negative effects of the stage length and number of segments on flight quality. Both number of segments and stage length are also cost related. Higher number of segments or a longer stage length increases the cost. Another related variable is the online rate which gives the percentage of tickets that do not involve change of operating carriers. Besides the extra costs, a change of carrier decreases the quality of an itinerary as this would cause some inconvenience to the consumers. The size of plane is both quality and cost related. On the one hand, larger aircraft are likely to be more stable and comfortable. On the other hand, on a flight involving a larger aircraft the waiting time for baggage would be longer and probability of mishandled baggage would potentially be higher. Hence, the overall net quality from size effect is ambiguous.

Population attributes are also related to price dispersion. A larger population is more likely to have a more heterogenous customer profile, which gives the airlines more opportunity to employ price discrimination strategies. Hence, we expect to observe a positive relationship between population size and price dispersion. Another route specific characteristic that can influence price dispersion is the per capita income. The income distribution is another relevant factor for price discrimination. A higher diversity in income would allow the airlines to benefit more from price discrimination. Also, the per capita income can control for the differences in pricing levels among different routes. The big city routes tend to have both leisure travelers, with high demand elasticity, and business travelers, with low demand elasticity; while leisure routes mostly have leisure travelers.[9] Hence, on the big city routes the airlines can distinguish business travelers from leisure travelers relatively more easily, and thus implement price discrimination more successfully.

[9]See Gerardi and Shapiro (2009) for more details about big city and leisure routes.

3.3 Data

This chapter focuses on domestic coach class airline tickets during the time period from 1999I to 2009IV. In particular, we consider the directional city-pairs (routes) that originate from Chicago. Our data set contains 108 routes and 14 carriers. The carriers are Frontier Airlines, JetBlue Airways, Southwest Airlines, and Spirit Airlines, Alaska Airlines, American Airlines, Continental Airlines, Delta Airlines, Northwest Airlines, United Airlines, US Airways, America West Airlines, ATA Airlines, and Trans World Airways. During the time period that we consider the U.S. airlines faced serious financial problems. Borenstein (2011) and Duygun, Kutlu, and Sickles (2014) point out that the financial losses for the domestic passenger airline operations in this time period is substantially higher than the losses between 1979 and 1999. Increases in taxes and jet fuel prices, relatively low fares, and a sharp decrease in demand are some of the challenges of this period for the U.S. airlines. Also, we observe a dramatic increase in load factors. More specifically, the average load factor increased from 71% to 81% between 2000 and 2009. Borenstein (2011) argues that such an increase might be attributed to improved yield management techniques. Thus, we would expect to observe some cost efficiency changes due to financial loss in this time period.

The product characteristics, quantity, and price data are collected from Passenger Origin-Destination Survey of the U.S. Department of Transportation (DB1B data set). This data set is a 10% random sample of all tickets that originate from the U.S. domestic cities. In our analysis, market is defined as a directional city-pair. Hence, we consider each city-pair as a separate market. Calculation of prices and quantities are based on the tickets that have no more than three segments in each direction. About 1% of tickets are eliminated during the elimination of tickets with more than 3 segments. Due to the differences in demand elasticities and other unobservable characteristics between coach class and higher-end classes (first class and business class), our study only focus on coach class tickets.

The product characteristics are captured by the average number of segments (SEG) for tickets sold by the carrier, average stage length (SL) for each segment, average size of the fleet ($SIZE$), and online rate ($ONLINE$). The size is defined as the average capacity of operating airplanes for each carrier in each quarter. The $ONLINE$ variable is the percentage of tickets where there is no change of operating carrier.

We use two different approaches to measure the price dispersion. First, we use the Gini coefficient ($GINI$) as a proxy for price dispersion.[10] Second, we generate another proxy, denoted by $PD_{90,10}$, using the 90^{th} and 10^{th} percentiles of the price distribution. Let P_x denote the price at x^{th} percentile. Then, we define $PD_{90,10} = (P_{90} - P_{10})/P_{10}$. The Gini coefficient concentrates more on the middle part of the price distribution. A price dispersion proxy that concentrates on the top and bottom tails of the price distribution provides additional information regarding the source of price dispersion. Hence, we use $PD_{90,10}$ for the sake of checking the robustness of our results.

The $GINI$ and $PD_{90,10}$ variables are generated after further elimination of incredible or outlier tickets. To be more specific, all tickets with incredible prices are dropped from our data set. Following Borenstein (1989) and Ito and Lee (2007), we dropped the open-jaw tickets since it would be difficult to distribute the ticket price into outbound and inbound segment for open jaw tickets. Moreover, the tickets that have a price less than 25 dollars or higher than 99 percentile or more than 2.5 times standard deviation from the mean for each airline within a route are eliminated from the data. The tickets that have prices less than 25 dollars are considered as frequent flyer program tickets and the tickets that have prices higher than 99 percentile or more than 2.5 times standard deviation are considered to be input (key punch) errors for the data set. For the round trip tickets, we divided the total price by two to get the one-way price.

Cost related characteristics are collected from BTS's T100 Domestic Database. As noted by Borenstein and Rose (1994), price dispersion can be caused by both cost-based price dispersion and discrimination-based dispersion. The cost data set is constructed from the firm level data of DOT's airline production data set (based on Form 41 and $T100$). Three types of important cost factors are controlled in our studies: labor cost, energy cost, and capital cost. This is done by constructing a price index for each of these categories. The

[10] See the Appendix A for more details.

salaries and benefits for five main types of personnel are provided in Form 41/P6. Annual employee number is given in Form 41, P10. We interpolated the annual employee data to get the quarterly values. Labor price is generated by combining five types of personnel. For energy price, we only capture the cost based on aircraft fuel. The energy input was generated by combining information on aircraft fuel gallons used with expense data per quarter. Flight capital is described by the average size (measured in number of seats) of the fleet. Also, the number of aircraft that a carrier operated from each different model of aircraft in airline's fleet is collected from DOT Form 41. For each quarter, the average number of aircraft in service is calculated dividing the total number of aircraft days for all aircraft types by the number of days in the quarter. This serves as a measure of the size of fleet. Capital price is calculated using these flight capital factors. All these cost factors are normalized by Tornqvist-Theil index.

Although the market power and efficiency are related concepts, the literatures on measuring market power and efficiency are developed independently.[11] Only recently there has been some studies that try to relate these two literatures.[12] One of the difficulties in the stochastic frontier literature is that estimation of cost efficiency requires the total cost data which is not available for the U.S. airlines at the route level. Chapter II in this thesis deals with this issue by combining the conduct parameter approach with the stochastic frontier analysis. A complication that they face is that the stochastic frontier models generally assume exogeneity of the regressors. Therefore, they use a control function approach in line with Kutlu (2010), Karakaplan and Kutlu (2015), and Kutlu (2015b).[13] Following the conduct parameter approach introduced by Chapter II, we obtain the marginal cost efficiencies (EFF) and conducts (CON) of airlines. Our conduct measure is firm-route-time specific, which contrasts with the route-time specific concentration measures of Borenstein and Rose (1994), Stavins (2001), and Gerardi and Shapiro (2009). Similarly, our efficiency variable

[11] See Perloff, Karp, and Golan (2007) and Kumbhakar and Lovell (2000) for book-length surveys on market power measurement and stochastic frontier analysis, respectively. See also Sickles (2005) for an extensive Monte Carlo study examining the performances of efficiency estimators.

[12] See, for example, Delis and Tsionas (2009), Koetter, Kolari, and Spierdijk (2012), Kutlu and Sickles (2012), and Kutlu and Wang (2015).

[13] See also Guan, Kumbhakar, and Myers (2009) and Tran and Tsionas (2013) for GMM based solutions for endogeneity.

is firm-route-time specific.

Market conditions can have a significant influence on price dispersion. So, we include the city-specific demographic variables in our analysis: per capita income and population. We get the city level per capita income and population data from Bureau of Labor Statistics. We interpolate the annual data to get the quarterly per capita income and population variables for each city. For each origin-destination city-pair, we define route-specific per capita income (PCI) as the population weighted per capita income for end-point cities. In order to get the real prices, we deflated the nominal prices by consumer price index (CPI) and use the year $1999I$ as the base time. We define the route-specific population as the average population (POP) of two end-point cities. Because not all airports are located in cities with demographic information, the number of the routes is reduced further in our final data set. In our final data set, we further eliminate the small firms and small routes. On a given route, firms with market shares less than 0.01 are eliminated. For a given quarter, any route with enplanements less than 1800, i.e., 20 passengers per day, is dropped. We also eliminate the routes with less than 30 observations.

We consider market concentration and marginal cost efficiency as potentially endogenous variables. One of the instrumental variables that we use is the geometric market share (GEO).[14] This is an instrumental variable that is used by Gerardi and Shapiro (2009) in their analysis of price dispersion. The formula for this variable is provided in Appendix A. Average marginal cost efficiency for all other routes that share the same origination city and other carriers on the same route ($EFFIV1$) and average marginal cost efficiency for all other carriers on other routes that share the same origination city ($EFFIV2$) are two other instrumental variables that we use in our estimations. We use $EFFIV2$ in our benchmark estimations and $EFFIV1$ for robustness check.

Since the demand structures of big city and leisure routes are different in terms of consumer heterogeneity, we also analyze price dispersion for big city and leisure routes separately. We classify a route as "big city" route if the destination is located within the

[14]The GEO variable is the same as $GEOSP$ variable that was used by Gerardi and Shapiro (2009). $GEOSP$ is similar to the $GEOSHARE$ in Borenstein and Rose (1994). The difference is that Borenstein and Rose (1994) use average daily enplanements while we use average quarterly enplanements.

30 largest MAs in the United States.[15] The leisure routes are defined by the median value of ratio of accommodation earnings to total nonfarm earnings through the time period of 2001-2009. Then, we label a route as leisure route if the median value of this ratio for the destination city belongs to the 85^{th} percentile.[16] Gerardi and Shapiro (2009) assume that the leisure route tickets are mainly purchased by leisure travelers, who have high demand elasticity and low reservation price. In contrast, big city routes contain a high proportion of business travelers, who have low demand elasticity and high reservation prices, along with tourism travelers. Hence, the big city routes are likely to show more consumer heterogeneity compared with the leisure routes, which makes it easier for airlines to employ price discrimination. As a consequence, if price discrimination effect dominates, the big city routes would have larger Gini coefficients compared with that of leisure routes. The median and mean of Gini coefficients for the big city routes, leisure routes, and whole sample are given in Table 3.1. We can see that the big city routes have larger price dispersion compared with the leisure routes. Based on this fact, it would be interesting to decompose the impacts of conduct and marginal cost efficiency on price dispersion by route types. For the sake of illustration, in Figure 3.1, we show the price distributions of two routes. Here, Orlando is a leisure city and Philadelphia is a big city. We can clearly see the differences in price dispersion on these two routes: Chicago-Orlando vs Chicago-Philadelphia. To be more specific, the distribution of price dispersion for the leisure route looks like a normal distribution while the distribution of price dispersion for the big city route looks like a bimodal distribution. This indicates that the airlines were more successful in identifying the business travelers in the big city routes compared with the leisure routes.

Table 3.1. Gini Coefficient Descriptive Statistics by Subsample

Route Types	Big City Routes	Leisure Routes	All Routes
Mean	0.2249	0.1929	0.2121
Median	0.2143	0.1854	0.2038

[15] For the destination cities that are both tourism city and big city, we just treat them as big city. The cities include Miami, Orlando, Phoenix, Portland, San Antonio, San Diego, and Tampa.

[16] The final lists of big city routes and leisure routes are given in Appendix A.

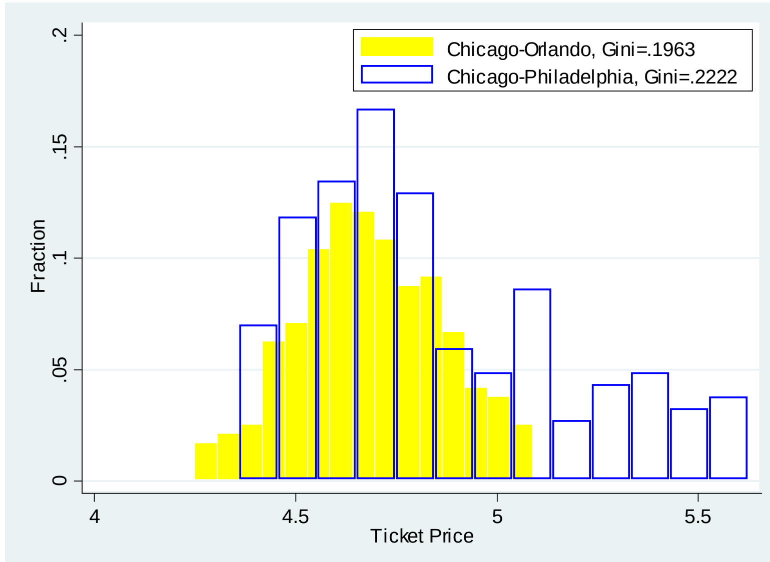

Figure 3.1 Price Dsipersion Orlando v.s. Philandelphia

Table 3.2. Descriptive Statistics

Variable Name	Mean	SD	Variable Description
Price Dispersion Measures			
GINI	0.2121	0.0735	Gini Coefficient
PD9010	1.9703	1.5742	(P90-P10)/P10
Competition Factors			
CON	0.2899	0.3772	Conduct Parameter
HHI	0.4106	0.1543	Herfindahl-Hirschman Index
Population Attributes			
ln(POP)	15.5092	0.1839	Logarithm of average population of origination and destination cities
ln(PCI)	10.4226	0.0383	Logarithm of population weighted per capita income of origination and destination cities
Leisure	0.2368	0.4251	Dummy for destination city, 1 if tourism city, 0 if non-tourism city
Big city	0.1857	0.3889	Dummy for destination city, 1 if big city, 0 if non-big city
Product Attributes			
ln(SL)	6.4084	0.5966	Logarithm of average stage length for an airline
ln(SEG)	0.3698	0.3217	Logarithm of average segment number for an airline
ONLINE	0.6674	0.3684	Percentage of tickets that do not involve change of operating carriers
ln(SIZE)	4.9726	0.0847	Logarithm of average number of available seats for an airline
Cost Factors			
ln(LP)	9.2320	0.7346	Logarithm of labor price by combining five categories of personnel normalized by Tornqvist-Theil index
ln(FP)	3.8727	0.5868	Logarithm of fuel price using only aircraft fuel parts normalized by Tornqvist-Theil index
ln(KP)	7.2798	0.7611	Logarithm of capital price normalized by Tornqvist-Theil index
EFF	0.7804	0.1361	Marginal cost efficiency
Instrumental Variables			
GEO	0.1942	0.1281	Geometric market share
EFFIV1	0.7804	0.0359	Average marginal cost efficiency for all other carriers or routes that share the same origination city
EFFIV2	0.7820	0.0359	Average marginal cost efficiency for all other carriers on other routes that share the same origination city
Observations	18209		

Table 3.2 presents the descriptive statistics for the Chicago (as origin city) based routes for three segments tickets. We measure the price dispersion using the Gini coefficient. The Gini coefficient reflects price inequality across the entire range of different prices paid by the consumers. As mentioned earlier, it is equal to the half of expected absolute differences in prices as a proportion of the mean price for two consumers drawn randomly from the sample. For example, our whole sample has a Gini coefficient of 0.2121 which implies that the expected absolute price difference is 42.42% of the mean fare. The conduct parameter, CON, has a mean value of 0.2899. This implies that on average the airline routes are neither competitive nor monopolistic. Finally, the average marginal cost efficiency in our sample is 0.7804. Hence, there is a room for efficiency improvement.

3.4 Empirical Specification and Estimation Results

In this section, we present our empirical analysis of price dispersion. As described earlier, we model the price dispersion by the conduct, marginal cost efficiency, other cost related variables, product attributes, and population attributes. The empirical specification estimated in our analysis is given by:

$$
\begin{aligned}
Y_{irt} &= \beta_0 + \beta_1 MP_{irt} + \beta_2 EFF_{irt} + f\left(COST_{it}; \delta\right) \\
&\quad + g\left(PRODATT_{irt}; \gamma\right) + h\left(POPATT_{rt}; \eta\right) + F_i + R_r + T_t + \varepsilon_{irt}
\end{aligned} \tag{16}
$$

where Y is the price dispersion measure which is $GINI$ or $PD_{90,10}$; MP is a variable which is either conduct (CON) or HHI; EFF is the marginal cost efficiency; f is a function of cost related variables, $COST$; g is a function of product attributes, $PRODATT$; h is a function of population attributes, $POPATT$; and F_i, R_r, and T_t are firm, route, and time dummies; and ε_{irt} is the error term. The indices i, r, and t refer to the firm, route, and time period, respectively. The variables that are representing product attributes are the average number of segments (SEG), average stage length (SL), online rate ($ONLINE$), and average size of aircraft ($SIZE$). The cost related variables are labor price (LP), fuel price (FP), and capital price (KP). Finally, the variables that are representing the population attributes are the average population (POP) and population weighted per capita income (PCI). The Gini coefficient is our benchmark measure for price dispersion. All of CON, HHI, and EFF variables lie in the unit interval. The conduct parameter values of 0, market share, and 1 correspond to perfect competition, Cournot competition, and monopoly (joint profit maximization), respectively.

As argued in the introduction, the HHI is an imperfect indicator of market power. Hence, it seems that using conduct parameter and marginal cost efficiency may provide a better understanding of the relationship between market power and price dispersion. The earlier studies that are examining the relationship between concentration measures and price dispersion do not consider the marginal cost efficiencies in their analysis. Hence, the effect of the marginal cost efficiency is embedded in the residual term. The quiet life and efficient structure hypotheses suggest that the market power and efficiency are related to each other. Hence, the parameter estimates in the earlier studies may be inconsistent if the instrumental variables are correlated with the efficiency. For example, Borenstein and Rose (1994) and Gerardi and Shapiro (2009) use the GEO variable as an instrument which turns out to be correlated with the efficiency. In our sample, the correlation of EFF and GEO is equal to -0.3754. While this correlation does not seem to be very high, it is still worthwhile to see the effect of ignoring inefficiency in the analysis of price dispersion.

In all estimations we assume that CON and HHI are endogenous; and we use GEO as an instrument. The marginal cost efficiency can be endogenous if there is a feedback from price dispersion to marginal cost efficiency. For example, reaching higher levels of price dispersion might complicate the manager's general tasks and negatively affect the firm performance in terms of reaching optimal level of cost. Based on this argument, it is reasonable to treat marginal cost efficiency as an endogenous variable. For each estimation table, we provide four sets of estimates: the "No Efficiency" model excludes the efficiency from the estimations; the "Exogenous Efficiency" model assumes that EFF is exogenous; and "Endogenous Efficiency I" and "Endogenous Efficiency II" models assume that EFF is endogenous but use different instrumental variables for marginal cost efficiency. To be more specific, "Endogenous Efficiency I" and "Endogenous Efficiency II" models use $EFFIV1$

and $EFFIV2$ as instruments (along with GEO), respectively. In what follows, the Endogenous Efficiency II would be our benchmark setting. Hence, if not specified, we are referring to estimations from "Endogenous Efficiency II" models. In Table 3.3-3.5 we assume that $Y = GINI$ and $MP = CON$.

Table 3.3. Price Dispersion Estimation using GINI for All Routes

GINI	No Efficiency	Exogenous Efficiency	Endogenous Efficiency I	Endogenous Efficiency II
CON	0.0176**	0.0077	0.0076	0.0101
	(0.0071)	(0.0071)	(0.0071)	(0.0071)
EFF		-0.2433***	-0.2432***	-0.1837***
		(0.0113)	(0.0113)	(0.0322)
ln(POP)	0.4685***	0.4434***	0.4396***	0.4486***
	(0.1696)	(0.1393)	(0.1394)	(0.1442)
ln(PCI)	-0.0182	0.1295	0.1271	0.0935
	(0.1622)	(0.1635)	(0.1635)	(0.1609)
ln(SL)	-0.2070***	-0.2366***	-0.2367***	-0.2294***
	(0.0538)	(0.0514)	(0.0514)	(0.0516)
ln(SL) Square	0.0148***	0.0183***	0.0183***	0.0175***
	(0.0041)	(0.0039)	(0.0039)	(0.0039)
ln(SEG)	-0.0105	0.0169	0.0169	0.0102
	(0.0105)	(0.0109)	(0.0109)	(0.0111)
ln(SIZE)	-0.0360	-0.0225	-0.0225	-0.0258
	(0.0265)	(0.0242)	(0.0242)	(0.0242)
ONLINE	0.0037	-0.0048	-0.0048	-0.0027
	(0.0042)	(0.0039)	(0.0039)	(0.0041)
ln(LP)	-0.0731**	-0.0805***	-0.0805***	-0.0787***
	(0.0289)	(0.0260)	(0.0260)	(0.0261)
ln(KP)	0.0266	0.0149	0.0149	0.0179
	(0.0326)	(0.0315)	(0.0315)	(0.0314)
ln(FP)	0.1276***	0.1323***	0.1323***	0.1311***
	(0.0436)	(0.0429)	(0.0429)	(0.0425)
ln(FP)*ln(LP)	0.0052	0.0051	0.0051	0.0051
	(0.0061)	(0.0057)	(0.0057)	(0.0057)
ln(KP)*ln(LP)	0.0028	0.0034	0.0034	0.0032
	(0.0028)	(0.0026)	(0.0026)	(0.0026)
ln(FP)*ln(KP)	-0.0031	-0.0009	-0.0009	-0.0014
	(0.0057)	(0.0052)	(0.0052)	(0.0053)
Time Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	18209	18209	18209	18209
Centered R Square	0.4044	0.4802	0.4802	0.4747

Note: Robust standard errors in parentheses, clustered by each ticketing carrier on each route. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.4. Price Dispersion Estimation using GINI for Big City Routes

GINI	No Efficiency	Exogenous Efficiency	Endogenous Efficiency I	Endogenous Efficiency II
CON	0.0522***	0.0336***	0.0335***	0.0317**
	(0.0121)	(0.0123)	(0.0124)	(0.0145)
EFF		-0.1746***	-0.1760***	-0.1927**
		(0.0261)	(0.0264)	(0.0867)
ln(POP)	0.7175***	0.6744***	0.6736***	0.6704***
	(0.2759)	(0.2187)	(0.2183)	(0.2084)
ln(PCI)	-0.0197	0.1040	0.1046	0.1170
	(0.1637)	(0.1646)	(0.1651)	(0.1774)
ln(SL)	0.1018	-0.0040	-0.0050	-0.0149
	(0.1848)	(0.1749)	(0.1750)	(0.1825)
ln(SL) Square	-0.0074	0.0027	0.0027	0.0037
	(0.0137)	(0.0131)	(0.0131)	(0.0140)
ln(SEG)	-0.0308	0.0009	0.0011	0.0041
	(0.0370)	(0.0341)	(0.0341)	(0.0384)
ln(SIZE)	0.0605	0.0778*	0.0780*	0.0797*
	(0.0490)	(0.0441)	(0.0441)	(0.0466)
ONLINE	-0.0058	-0.0054	-0.0054	-0.0053
	(0.0141)	(0.0125)	(0.0125)	(0.0124)
ln(LP)	0.0224	0.0176	0.0176	0.0171
	(0.0615)	(0.0524)	(0.0524)	(0.0519)
ln(KP)	-0.0946	-0.1453**	-0.1457**	-0.1506**
	(0.0691)	(0.0639)	(0.0638)	(0.0738)
ln(FP)	0.2926***	0.3240***	0.3243***	0.3273***
	(0.0895)	(0.0809)	(0.0809)	(0.0801)
ln(FP)*ln(LP)	-0.0246**	-0.0283***	-0.0284***	-0.0287**
	(0.0121)	(0.0108)	(0.0107)	(0.0112)
ln(KP)*ln(LP)	0.0069	0.0103*	0.0104*	0.0107*
	(0.0061)	(0.0055)	(0.0055)	(0.0061)
ln(FP)*ln(KP)	0.0213*	0.0245**	0.0245**	0.0248**
	(0.0119)	(0.0102)	(0.0102)	(0.0104)
Time Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	3382	3382	3382	3382
Centered R Square	0.5988	0.6388	0.6389	0.6396

Note: Robust standard errors in parentheses, clustered by each ticketing carrier on each route. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.5. Price Dispersion Estimation using GINI for Leisure Routes

GINI	No Efficiency	Exogenous Efficiency	Endogenous Efficiency I	Endogenous Efficiency II
CON	-0.0036	-0.0171	-0.0170	-0.0172
	(0.0132)	(0.0136)	(0.0136)	(0.0151)
EFF		-0.3254***	-0.3227***	-0.3282***
		(0.0305)	(0.0309)	(0.1159)
ln(POP)	0.0319	-0.0326	-0.0285	-0.0316
	(0.2145)	(0.1726)	(0.1727)	(0.1723)
ln(PCI)	-0.5491	-0.5040*	-0.4977	-0.5013*
	(0.3467)	(0.3034)	(0.3035)	(0.3037)
ln(SL)	-0.1395	-0.1894**	-0.1890**	-0.1899**
	(0.0884)	(0.0817)	(0.0817)	(0.0867)
ln(SL) Square	0.0064	0.0123**	0.0122**	0.0123*
	(0.0062)	(0.0059)	(0.0059)	(0.0066)
ln(SEG)	-0.0326*	0.0053	0.0050	0.0057
	(0.0183)	(0.0176)	(0.0176)	(0.0219)
ln(SIZE)	-0.0539	-0.0459	-0.0460	-0.0459
	(0.0414)	(0.0360)	(0.0360)	(0.0357)
ONLINE	0.0050	-0.0039	-0.0038	-0.0040
	(0.0074)	(0.0069)	(0.0069)	(0.0078)
ln(LP)	-0.1064**	-0.1071***	-0.1072***	-0.1072***
	(0.0478)	(0.0400)	(0.0401)	(0.0400)
ln(KP)	0.0320	0.0180	0.0182	0.0179
	(0.0465)	(0.0449)	(0.0449)	(0.0459)
ln(FP)	0.0012	0.0745	0.0736	0.0750
	(0.0676)	(0.0718)	(0.0717)	(0.0760)
ln(FP)*ln(LP)	0.0122	0.0102	0.0103	0.0102
	(0.0096)	(0.0088)	(0.0088)	(0.0088)
ln(KP)*ln(LP)	0.0054	0.0058*	0.0058*	0.0058*
	(0.0039)	(0.0035)	(0.0035)	(0.0035)
ln(FP)*ln(KP)	-0.0110	-0.0081	-0.0082	-0.0081
	(0.0092)	(0.0084)	(0.0084)	(0.0085)
Time Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	4311	4311	4311	4311
Centered R Square	0.4475	0.5255	0.5255	0.5255

Note: Robust standard errors in parentheses, clustered by each ticketing carrier on each route. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The EFF variable has a negative sign, for all estimates in Table 3.3-3.5, indicating that as marginal cost efficiency increases the price dispersion is tempted to decrease. For big city routes, a 10 percentage points increase in marginal cost efficiency would lead to 0.0193 decrease in Gini coefficient. Moreover, 10 percentage points increase in marginal cost efficiency would lead to 0.0328 decrease in the Gini coefficient for the leisure routes. The implication is that the expected fare difference would fall by 3.9% and 6.6% as the ratio of the mean fares for two randomly selected tickets, respectively. The CON variable is not significant in Table 3.3 and Table 3.5. However, it is significant and positive in Table 3.4. Hence, in big city routes the conduct has a positive effect on price dispersion. In particular, 0.1 point increase in conduct[17] would lead to 0.0032 increase in the Gini coefficient. The implication is that the expected fare difference would increase by 0.64% as the ratio of the mean fares for two randomly selected tickets. This is in line with the idea that in the heterogeneous routes it is easier for the airlines to distinguish the consumers, and thus the extent of price discrimination in these routes is expected to be higher compared with the homogenous routes. This can be explained by differences in demand elasticities. Business travelers have relatively lower demand elasticities for air transportation compared with the leisure travelers, so airlines can charge a higher price to the business travelers. However, the extent of price dispersion due to conduct is not high. Hence, the marginal cost efficiency is a much more effective determinant of price dispersion than the conduct. This result is robust in all our estimations. Comparing the conduct estimates from the first and last columns for Table 3.3-3.5, we see that the conduct parameter is over-estimated when the marginal cost efficiency is not included. Although this pattern that omitting efficiency leads to higher coefficients for the conduct is robust to different subsamples, the differences in these coefficient estimates are not statistically significant. So, there is some evidence for over-estimation (due to robustness of the over-estimation) but this evidence is weak at best.

Gini coefficient puts more weight on median values of prices, and thus it may not be capturing some of the essential aspects of price dispersion. Therefore, we also use another measure of price dispersion, $PD_{90,10}$, which concentrates on the tail distributions of prices. This measure would also be useful in testing the robustness of our results. In Table 3.6-3.8 we assume that $Y = PD_{90,10}$ and $MP = CON$. The EFF variable has a negative sign, for all estimates in Table 3.6-3.8, which agrees with our earlier estimates. The CON variable is not significant in Table 3.7. However, it is significant and negative in Table 3.6 and Table 3.8. Hence, based on these estimates, in leisure routes the conduct has a negative effect on

[17] For comparison, the change in conduct when the number of symmetric Cournot players change from 3 to 4 would roughly be equal to 0.08. More precisely, it is $\frac{1}{3} - \frac{1}{4} = \frac{1}{12}$.

price dispersion. This is in line with the idea that in the homogenous routes it is harder for the airlines to distinguish the consumers. Again, the extent of price dispersion due to the conduct is not high. In particular, 0.1 point increase in conduct would lead to a 1.07% decrease in $PD_{90,10}$ for leisure routes. Omitting the EFF variable from the estimations lead to a positive bias on the coefficient of CON. The direction of this bias is robust to different model specifications and subsamples that we used.

Table 3.6. Price Dispersion Estimation using PD9010 for All Routes

(P90-P10)/P10	No Efficiency	Exogenous Efficiency	Endogenous Efficiency I	Endogenous Efficiency II
CON	-0.3112**	-0.5556***	-0.5590***	-0.5366***
	(0.1441)	(0.1431)	(0.1433)	(0.1456)
EFF		-5.9524***	-6.0350***	-5.4902***
		(0.2364)	(0.2404)	(0.6624)
ln(POP)	9.4251***	8.8194***	8.7953***	8.8726***
	(3.5219)	(2.6106)	(2.6028)	(2.6583)
ln(PCI)	-0.1188	3.4838	3.5438	3.2135
	(2.9949)	(2.6057)	(2.6090)	(2.5763)
ln(SL)	-3.6859***	-4.4109***	-4.4206***	-4.3545***
	(1.3456)	(1.3228)	(1.3236)	(1.3264)
ln(SL) Square	0.2946***	0.3809***	0.3821***	0.3742***
	(0.0996)	(0.0976)	(0.0977)	(0.0986)
ln(SEG)	-0.0582	0.6130***	0.6223***	0.5610**
	(0.2101)	(0.2125)	(0.2132)	(0.2278)
ln(SIZE)	-0.8438	-0.5147	-0.5100	-0.5402
	(0.5160)	(0.4463)	(0.4466)	(0.4447)
ONLINE	-0.0183	-0.2245***	-0.2274***	-0.2085**
	(0.0897)	(0.0826)	(0.0826)	(0.0840)
ln(LP)	-1.1453**	-1.3291***	-1.3312***	-1.3148***
	(0.5587)	(0.4573)	(0.4570)	(0.4568)
ln(KP)	0.6047	0.3199	0.3161	0.3420
	(0.6764)	(0.6240)	(0.6248)	(0.6252)
ln(FP)	1.9086**	2.0206***	2.0217***	2.0124***
	(0.7637)	(0.7051)	(0.7060)	(0.7006)
ln(FP)*ln(LP)	0.1194	0.1177	0.1176	0.1179
	(0.1243)	(0.1099)	(0.1100)	(0.1098)
ln(KP)*ln(LP)	0.0506	0.0650	0.0652	0.0639
	(0.0583)	(0.0525)	(0.0526)	(0.0525)
ln(FP)*ln(KP)	-0.0982	-0.0443	-0.0435	-0.0485
	(0.1154)	(0.0976)	(0.0976)	(0.0987)
Time Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	18209	18209	18209	18209
Centered R Square	0.2839	0.3760	0.3759	0.3755

Note: Robust standard errors in parentheses, clustered by each ticketing carrier on each route.
* p < 0.1, ** p < 0.05, *** p < 0.01.

Table 3.7. Price Dispersion Estimation using PD9010 for Big City Routes

(P90-P10)/P10	No Efficiency	Exogenous Efficiency	Endogenous Efficiency I	Endogenous Efficiency II
CON	0.6027***	0.0286	0.0062	0.1315
	(0.2074)	(0.1962)	(0.1961)	(0.2763)
EFF		-5.3868***	-5.5967***	-4.4212**
		(0.4753)	(0.4786)	(1.7466)
ln(POP)	16.5544***	15.2415***	15.1840***	15.4744***
	(6.1889)	(4.2658)	(4.2049)	(4.5195)
ln(PCI)	1.6205	5.4444**	5.5908**	4.7579*
	(3.2073)	(2.6112)	(2.6320)	(2.7381)
ln(SL)	4.2874	1.0250	0.8977	1.6097
	(2.7399)	(2.3533)	(2.3530)	(2.6115)
ln(SL) Square	-0.3253*	-0.0155	-0.0034	-0.0710
	(0.1942)	(0.1684)	(0.1684)	(0.1989)
ln(SEG)	-0.6830	0.2932	0.3312	0.1182
	(0.5977)	(0.5076)	(0.5067)	(0.6100)
ln(SIZE)	0.5995	1.1353	1.1564	1.0393
	(1.1044)	(0.8854)	(0.8833)	(0.9498)
ONLINE	-0.1912	-0.1787	-0.1782	-0.1809
	(0.3128)	(0.2499)	(0.2485)	(0.2569)
ln(LP)	-0.7585	-0.9062	-0.9117	-0.8798
	(1.6528)	(1.2441)	(1.2341)	(1.3039)
ln(KP)	-0.6692	-2.2327**	-2.2940**	-1.9524
	(1.4611)	(1.1269)	(1.1210)	(1.3286)
ln(FP)	4.2823***	5.2537***	5.2919***	5.0795***
	(1.6211)	(1.3093)	(1.3074)	(1.3209)
ln(FP)*ln(LP)	-0.1250	-0.2420	-0.2467	-0.2210
	(0.3305)	(0.2531)	(0.2512)	(0.2641)
ln(KP)*ln(LP)	0.1568	0.2631***	0.2673***	0.2441**
	(0.1243)	(0.0966)	(0.0961)	(0.1133)
ln(FP)*ln(KP)	0.0389	0.1368	0.1407	0.1193
	(0.3314)	(0.2525)	(0.2504)	(0.2633)
Time Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	3382	3382	3382	3382
Centered R Square	0.3283	0.3854	0.3853	0.3651

Note: Robust standard errors in parentheses, clustered by each ticketing carrier on each route. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.8. Price Dispersion Estimation using PD9010 for Leisure Routes

(P90-P10)/P10	No Efficiency	Exogenous Efficiency	Endogenous Efficiency I	Endogenous Efficiency II
CON	-0.6597***	-0.9270***	-0.9310***	-1.0676***
	(0.2551)	(0.2904)	(0.2910)	(0.3810)
EFF		-6.4590***	-6.5564***	-9.8573***
		(0.9274)	(0.9795)	(2.7810)
ln(POP)	3.7111	2.4526	2.4436	1.7676
	(4.8276)	(4.0482)	(4.0390)	(4.2265)
ln(PCI)	-2.6760	-1.7466	-1.7084	-1.2822
	(7.4863)	(6.7686)	(6.7684)	(6.6986)
ln(SL)	-0.3396	-1.3322	-1.3452	-1.8537
	(1.5822)	(1.6879)	(1.6891)	(1.9994)
ln(SL) Square	0.0393	0.1569	0.1585	0.2187
	(0.1055)	(0.1199)	(0.1201)	(0.1560)
ln(SEG)	0.1452	0.8976*	0.9091*	1.2934*
	(0.4184)	(0.5321)	(0.5370)	(0.7431)
ln(SIZE)	-0.3243	-0.1679	-0.1654	-0.0852
	(0.8576)	(0.7577)	(0.7577)	(0.7460)
ONLINE	0.0444	-0.1322	-0.1349	-0.2252
	(0.1943)	(0.1865)	(0.1864)	(0.2196)
ln(LP)	-0.8354	-0.8511	-0.8516	-0.8588
	(0.7764)	(0.6323)	(0.6322)	(0.6667)
ln(KP)	-1.5694	-1.8453	-1.8494	-1.9912
	(1.3345)	(1.3147)	(1.3160)	(1.3662)
ln(FP)	-0.5363	0.9134	0.9348	1.6773
	(1.1273)	(1.1643)	(1.1706)	(1.4071)
ln(FP)*ln(LP)	-0.1211	-0.1596	-0.1601	-0.1800
	(0.2288)	(0.2176)	(0.2177)	(0.2258)
ln(KP)*ln(LP)	0.2014*	0.2095*	0.2096*	0.2138*
	(0.1111)	(0.1072)	(0.1072)	(0.1105)
ln(FP)*ln(KP)	0.1533	0.2101	0.2109	0.2401
	(0.1926)	(0.1806)	(0.1807)	(0.1929)
Time Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	4311	4311	4311	4311
Centered R Square	0.5248	0.6095	0.6096	0.6058

Note: Robust standard errors in parentheses, clustered by each ticketing carrier on each route. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Now, we combine our findings from Table 3.3-3.8. We start with Table 3.3 and Table 3.6. The coefficient of CON in Table 3.3 is statistically insignificant. Hence, we rely on the coefficient from Table 3.6, which is negative and significant. The sign difference seems to be incidental due to insignificance of the coefficient of CON in Table 3.3. Thus, we argue that as the conduct increases, the overall price dispersion falls. This finding coincides with the findings of Borenstein and Rose (1994). Based on Table 3.4 and Table 3.7, we argue that the conduct has a positive effect on price dispersion on big city routes. Finally, based on Table 3.5 and Table 3.8, we conclude that the conduct has a negative effect on price dispersion. Hence, if the whole sample is used in estimations, the positive and negative effects in the big city and leisure routes may cancel out each other. This maybe one of the potential reasons for the mixed results in the literature.

Finally, in Table 3.9, we consider HHI as a determinant of price dispersion and assume that $Y = GINI$ and $MP = HHI$. This table helps us see the differences between CON and HHI when modeling price dispersion. In order to save space, we announce only the coefficient estimates for HHI and EFF. Again, the EFF variable has a negative sign, for all estimations, which agrees with our earlier results.

When we do not include the cost efficiency, HHI have a positive effect on price dispersion for all routes and big city routes. The effect is not statistically significant for leisure routes. Hence, the negative sign seems to be incidental. The estimation results are in line with the findings of Gerardi and Shapiro (2009) that market concentration has a positive effect on price dispersion. In big city routes, compared with the leisure routes, the market concentration have a larger effect on price dispersion. As argued by Gerardi and Shapiro (2009), this may be an evidence of larger price discrimination on big city routes. However, once we include the marginal cost efficiency, the statistical significance of these results disappears. As we mentioned earlier, HHI might not be a good proxy for market power since it does not capture the demand elasticities and marginal cost efficiencies. Moreover, HHI is a route specific measure, which may not control for the firm specific variations. Finally, our CON estimates are jointly estimated with EFF, which makes CON a more compatible measure to use along with EFF.

Table 3.9. Price Dispersion Estimation using HHI

GINI	A. All Routes		B. Big-city Routes		C. Leisure Routes	
	No EFF	With EFF	No EFF	With EFF	No EFF	With EFF
HHI	0.1969**	0.1186	0.8143**	1.7546	-0.0595	-0.3594
	(0.0879)	(0.0873)	(0.3993)	(1.1820)	(0.2276)	(0.3906)
EFF		-0.1708***		-0.1575		-0.4253*
		(0.0363)		(0.1647)		(0.2323)
Time Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18209	18209	3382	3382	4311	4311
Centered R Square	0.3806	0.4617	0.2274	-0.1280	0.4338	0.3537

Note: Robust standard errors in parentheses, clustered by each ticketing carrier on each route. The columns of "With EFF" exhibit the estimation results including marginal cost efficiency using EFFIV2. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.5 Summary and Concluding Remarks

In this study, we examined the determinants of airline price dispersion. We particularly concentrated on the conduct and marginal cost efficiency determinants of price dispersion. Marginal cost efficiency is an important yet unobservable determinant of price dispersion, which is ignored by earlier studies. Due to its strong relationship with market power, ignoring the airline efficiencies can lead to inconsistent parameter estimates. Hence, we compared the estimation results with and without marginal cost efficiency to understand the extent of potential biases in the earlier studies. The estimation results show that the marginal effect of conduct may be over-estimated when the marginal cost efficiency is omitted. We considered the firm-specific conduct, CON, as a determinant of price dispersion in contrast to the route specific HHI that are used the literature. If we consider CON and HHI as proxies for market power, the CON variable has the advantage that it may enable us to capture the relationship between market power and price dispersion more precisely.

We argued that the conduct has opposing effects on big city and leisure routes. In particular, for the big city routes the relationship is positive and for the leisure routes it is negative. This contrasting results may be a potential reason for mixed empirical results in the literature. The marginal cost efficiency has a negative effect on price dispersion and this finding was robust in our all specifications. We found that the marginal cost efficiency is a relatively more important component of price dispersion. The marginal cost efficiency not only has a direct effect on social welfare loss but also has an indirect effect through its significant influence on the extent of price dispersion. While it would be difficult to measure this indirect effect on the social welfare, it is revealed that marginal cost efficiency has a shrinking role on the effect of price dispersion for social welfare by decreasing the price dispersion.[18] This shrinkage effect dominates the corresponding effect of conduct on social welfare. Finally, we found that the marginal cost efficiency is more effective in leisure routes compared with the big city routes. This may be partially because the price discrimination would be harder to implement for the airlines (due to lack of consumer heterogeneity).

[18] See Kumar and Kutlu (2015) for a study that considers the effect of price discrimination on social welfare.

CHAPTER IV

REVENUE EFFICIENCY AND DYNAMIC PRICING IN AIRLINE INDUSTRY

4.1 Introduction

The pricing of the flight tickets is always mystery for the consumers. It is common to have one roundtrip ticket cheaper than an one-way ticket even for the same city pair. Also, it is popular to find that the prices for the same flight change as frequently as daily, or even hourly. Borenstein and Rose (1994) show that the expected difference in fares between two randomly chosen travelers is 36% as the ratio of mean fare. Gerardi and Shapiro (2009) and Gaggero and Piga (2011) find even higher values for price dispersion, 44% and 66%, respectively. In Chapter III, we also find evidences for price dispersion. For Chicago based routes, the Gini coefficient is 0.21. Based on these observations, it would be interesting to examine how the airlines dynamically make the pricing decisions in order to maximize their revenues. For the airline industry, once the flight schedule is determined, the fixed cost is large enough to allow us to ignore the variable cost. Another reason is that in the short run, the purchases of new aircrafts are not common, which contribute to the highest percentage of fixed cost in airline industry. So, it is reasonable to concentrate on revenue optimization when modeling airline behavior.[1]

Revenue management is also called yield management or seat inventory management or dynamic pricing. According to Kimes (2002), yield management is defined as a method that can help a firm sell the right inventory to the right customer at the right time for the right price. In the airline industry, airlines sell identical seats at different prices to maximize revenues. Since American Airlines introduced the yield management technique in early 1980s, as a response to People's Express's deeply discounted fares, the yield management

[1] In this Chapter, the time period covered is from Apr 6th, 2015 to May 10th, 2015, which is a very short time period.

has become a more and more important strategy to gain competing advantages. Yield management is realized through the computerized reservation system, which keeps records of airline seats booking and fare information. Nowadays, this computerized reservation system has become available not only for the airlines, but also for the travel agents. By using this computerized system, it is easier for the airlines to track the demand states and adjust the prices based on the demand forecast and competition dynamics.

As pointed out by Pinder (1995), Sridharan (1998) and Barut and Sridharan (2004), the made-to-order (MTO) manufacturing industry is suitable for yield management. Yield management is a general practice for perishable inventory control, such as hotel, rental car, cruises and flight tickets. There are two common characteristics of these products or service. First, the product/service expires at a certain point of time. Second, the capacity is fixed in advance and the capacity constraint can only be extended at a very high marginal cost. These two characteristics make the dynamic pricing strategy highly important for these perishable goods/service. Davis (1994) and Smith et al. (1992) show that American Airlines benefit from Yield Management (YM) and Smith et al. (1992) state that American Airlines made an extra $1.4 billion between 1989 and 1991 because of its advanced yield management techniques.

However, even though the importance of yield management is highly acknowledged, there is not much research concerning yield management. The main reason might be that the dynamic pricing data and load factor associated with dynamic price data are not easy to get. This study overcomes this difficulty by scripting the online data from *priceline.com*, which is a popular website to book flight tickets in the U.S.. This website provides us the information about ticket price, airline name, flight number, departure time, arrival time, departure airport, arrival airport and seat map. We can count total number of available seats, total number of unavailable seats and total capacity from seat maps using *Perl* program. By scripting the daily data from top 10 Chicago based metropolitan city routes and combining this with codeshare and hub information, we create a unique dynamic pricing database which allows us to make an analysis about revenue efficiency and dynamic pricing.

In this chapter, we first analyze the revenue management efficiency of the U.S. domestic airlines for each flight. Then we divide the whole sample into high-efficient flights and low-efficiency flights based on the revenue efficiency levels. By comparing the high-efficiency flights' dynamic pricing pattern with the low-efficiency flights' pattern, we summarize the potential points that airline could pay attention to in the future in order to improve revenue. In this chapter, we mainly focus on the effects from load factor and advanced days purchased (ADP) on dynamic prices, which are two main drivers for dynamic price changes. By doing this, we also testify the validity of capacity-based theories and time-based theories in the airline industry.

The main findings in this chapter are illustrated as follows. First, we find large differences in revenue efficiency among different flights, firms and routes. Second, the dynamic pricing patterns for the high-efficiency flights are different from those of low-efficiency flights. To be more specific, the high-efficiency flights tend to be less responsive to timing changes (advanced days purchased, ADP). Especially for the time period between two weeks and three weeks before departure, the low-efficiency flights responds more sharply compared with high-efficiency flights. So, the implication of this is that airlines could further improve the revenue by adjusting their reservation systems so that the prices do not change so frequently based on timing, because such frequent changes in prices might worsen the revenue optimization. Third, we only find weak evidences for capacity-based theories, but find stronger evidences for time-based theories.

The rest of Chapter IV is structured as follows: Section 4.2 contains a detailed literature review about revenue management and dynamic pricing strategies. Section 4.3 introduces the data scripted from *priceline.com* and data from *flightstats.com*. Section 4.4 exhibits the empirical specifications and the relevant results. In Section 4.5 we present our conclusions and make policy suggestions.

4.2 Related Literature

This part summarizes the main theoretical and empirical studies about revenue/yield management. Generally speaking, there are three sets of theories. One is classical price discrimination theory, another is capacity-based theories and the third is time-based theories. In classical price discrimination models, the airlines try to distinguish and segment the consumers by using different ticket restrictions, such as Saturday night stay and non-refundability restrictions. By adding different ticket restrictions to different tickets, the travelers self select themselves into different groups based on their different budgets and demand elasticities.

Capacity-based theories argue that capacity is limited in the airline industry and the cost associated with augmented capacity is large, that is airline capacity is costly but perishable. Airline market is characterized by perishable goods, costly capacity and uncertain demand. Prescott (1975), Eden (1990) and Dana (1999b) explain the relationship between fares and seat availability under the assumption of uncertain demand. Also Dana(1999a) shows that price dispersion increases demand shifting and thus increases the social welfare by allocating the consumers into available seats.

Gale and Holmes (1992, 1993) employ a mechanism design approach to model advance purchase discounts in a monopoly market. The monopoly firm utilizes fare discounts to divert the consumers from "peak" to "off peak" flights. In their studies, it is assumed that the consumers can only learn their time preferences right before the departure and different customers have different opportunity costs of waiting. Based on their studies, the peak flights should have higher average fare compared with off peak flights and the price dispersion from off peak flight should be higher, as stated in Puller et al. (2009). Dana (1998) states that in a perfectly competitive market, there might be advance purchase discounts. Deneckere and Peck (2012) build a model to allow the current price to depend on the evolution of the aggregate quantity sold in last time period on the supply side. On the demand side, they allow inter-temporal substitution so that the consumers can decide whether or not to delay their purchases based on their knowledge of demand states and their expectations of future prices. By doing this, they relax the assumption in Dana (1998) that the airlines cannot adjust the prices as they learn new information about demand states and other information.

Kimes (1989) explains the general requirements to implement revenue management. Also he points out possible implementation approaches based on whether the price is variable and whether the duration can be predictable. Kimes (1989) shows that different industries should implement different revenue strategies based on their industry specific characteristics. He states that, similar with hotel industry, airline industry lies in the Quadrant-2 industries, so airline firms are trying their best to control for both duration and dynamic pricing, which makes the revenue management more sophisticated. Brooks and Buttons (1994) use shipping industry as an example to illustrate the rise of yield management during 1970s and 1980s. Talluri and van Ryzin (2006) talk about the details in yield management in their textbook.

Due to the difficulty of the empirical data on the dynamic occupancy ratio corresponding to dynamic fares, only a limited number of empirical studies try to testify and analyze these theories. Puller et al. (2009) use a census of ticket transactions from one computer reservation system to study the relationship between fares, ticket characteristics and flight load factors. They find mixed supports for the scarcity pricing theories. Escobari and Gan (2007) employ a panel data analysis and find empirical support for the capacity-based theories. By developing an effective cost of capacity (ECC) model, they show that higher ECC would lead to higher prices. Also, they find that the effect from ECC on price is higher in competitive markets. Also, Piga and Bachis (2007) find empirical support for Deneckere and Peck (2012) that price does not necessarily increase over time. Escobari (2012) further confirms that fares decrease until about two weeks before departure and then increase, holding inventories constant.

4.3 Data Sources and variable constructions

The main data source in this chapter comes from data scripted from *priceline.com*, which is among the top five biggest online travel agents. The other four online agents are *Expedia*, *Kayak*, *Orbitz* and *Travelocity*. In 2014, 52% of the tickets are booked through online travel websites. *Priceline* offers almost lowest prices by providing lowest price guarantee. Even though the importance of revenue management is widely acknowledged by the economists, airlines and consumers, there are not many literatures that work on the dynamic pricing based on the dynamic inventory in the airline industry. One of the difficulties is that the dynamic capacity occupancy at the time of a fare is not easy to get for the researchers. In this chapter, we write a *Perl* program to script the daily dynamic inventory change information associated with daily dynamic pricing data from *priceline.com*. Stavins (2001) points out that since there are different booking classes, each with different sets of ticket restrictions, it is necessary to control for ticket restrictions. By scripting ticket specific data from *priceline.com*, we do not need to worry about the different ticket restrictions because tickets offered by *priceline.com* have the same ticket restrictions.[2] Considering that high-end (first class and business class) tickets and low-end (coach class) tickets are not so comparable because they have different market segments with different consumers, we only focus on coach class tickets. What is more, we only focus on the non-stop flights for two main reasons. One is that we do not want to introduce more heterogeneity between non-stop flights and multi-stops flights, making our analysis complicated. The other reason is that it takes too much time to download the multi-stop flights due to the large volume of multi-stop flights.[3]

In this study, we track the dynamic pricing behavior through *priceline.com* each day at the same time. The main variables scripted from *priceline.com* are price, airline name, flight number, origin city, origin airport, destination city, destination airport, quote date, departure date, departure time, arrival time and seat map. Our *Perl* program counts the total number of seats, total number of available seats and their IDs, total number of unavailable seats and their IDs from the seat map for each flight. In order to illustrate this process better, Figure 4.1 is provided here. As shown in Figure 4.1, after entering the city pairs and departure date, our *Perl* program will automatically list the full list of available flights. Then this program automatically open the seat map associated with each ticket.

[2]There are some refundable full price tickets offered by *priceline.com* on departure date, but we delete them from the dataset to make sure all tickets in this study have same ticket restrictions.

[3]The volume of multi-stop flights is about ten times of the volume of non-stop flights.

And the seat map provides different colors and symbols for the "open seats" and "taken seats", so our *Perl* program automatically remembers each color and symbol and counts the total number of unavailable seats and total number of available seats, including their corresponding IDs. Dividing the total number of unavailable seats by total number of seats, we obtain the daily dynamic load factor for each flight on each scheduled departure date. Using the changes in total number of unavailable seats, we calculate the quantity sold each day, from where we obtain the total revenue for each flight. This is one of the important highlights in the data construction part in this study.

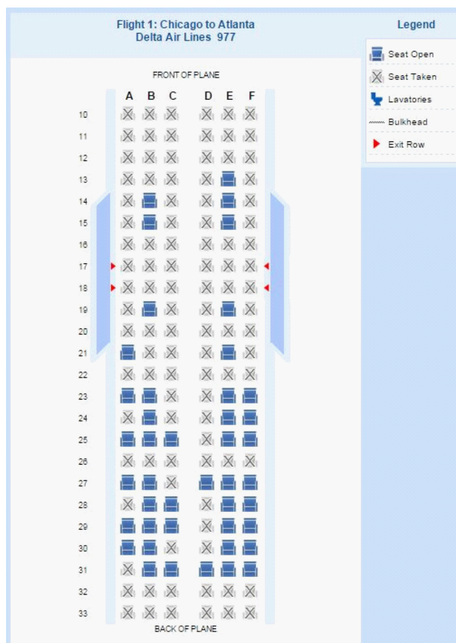

Figure 4.1 Example for data scripting

Due to computation time constraints, the final dataset only contains the flights that originate from Chicago to top ten metropolitan cities, including Atlanta, New York, Houston, Dallas, Boston, Los Angles, San Francisco, Miami, Washington, D.C. and Philadelphia.[4] Since Chicago is a big city with large airports and heavy passenger traffic, it is popular to use Chicago-based routes. We follow Brander and Zhang (1990), using Chicago based routes. And by doing this, we keep consistent throughout the thesis. We focus on the flights that have scheduled departure dates between Apr 6th, 2015 and May 10th, 2015 to

[4] Chicago is one of the top 10 Metropolitan cities, so we extend the city list to top 11 cities.

avoid the summer period. The summer period is known to have special demand and pricing patterns compared with non-summer period. So, in this chapter, we only focus on non-summer period to make things simple. We follow each flight from 60 days before departure date till the departure date. The scription of the dynamic pricing and inventory data is conducted each day at the same time. From *priceline.com*, we obtain the daily price, load factor, advanced days purchased (ADP), total flight time. However, since the load factor and quantity (revenue) are calculated from the online seat maps but the low cost carriers (LCCs) do not have the seat maps posted on *priceline.com*,[5] we lose the information about low cost carriers for the revenue and dynamic load factor. So, the LCCs are not included in the revenue efficiency estimation and dynamic pricing estimations. However, we consider the competition effects from the LCCs by taking them into consideration when calculating the available flights on market. Also, we calculate the dynamic available number of flights for each day, including airline-route-quoted date level, airline-quoted date level and route-quoted date level. These dynamic number of flights are based on dynamic data of available flights left on market. Since not all airlines post their tickets online (such as Southwest[6]), the calculation of number of available flights might have some negative bias. However, this might be the best way to capture the dynamic changes in number of available flights based on our data. The idea is that if there are a wide range of flights available from other carriers, the price should be lower due to the competition effect. If some flights become unavailable and exit the market, the prices might increase for flights left on market due to the adjustments in market structure. It may be considered as a dynamic measure of competition level. If one carrier has a lot of flights available on the same routes, some consumers will be willing to pay higher price since it would be easier to reschedule the flight if their flights are delayed or they miss their flights. It is also possible that high delay rate due to high utilization rate of aircrafts might lead to lower demand for such flights. Daily revenue is calculated based on changes in the total number of unavailable seats and daily price. Then, we sum up the total revenue for each flight. Due to the data limitations, we assume that

[5] Most LCCs donot allow consumers to choose their seats before departure.

[6] Southwest does not advertise its prices on the travel agents' website to keep its cost low.

prices stay constant beyond 60 days before departure.

In addition to the data collected from *priceline.com*, we also collect the codeshare information from *flightstats.com*. From this website, we collect the number of domestic codeshare carriers and the number of foreign codeshare carriers for each flight. Codeshare agreement is that two or more carriers share the same flight. So, codeshare means that the carriers under the agreement are using the same operating carrier. It is a nowadays a popular strategy for the airlines to cooperate with both domestic airlines and foreign airlines. However, codeshare partners still publish and market the flight under their own airline designators and flight numbers as part of their published timetable or schedule. The hub information is collected from each airline's official website. We create two dummy variables for hub, Hub_Origin (equal to 1 when the origination city is a hub for the carrier) and Hub_Dest (equal to 1 when the destination is a hub for the carrier). The number of airports in origin city and destination city is counted based on the number of airports that have available flights between the city pairs. For instance, for Chicago-Atlanta route, there are flights from ORD airport to ATL airport and from MDW airport to ATL airport, so on route Chicago-Atlanta, we count the number of airports in the origination city, $N\_Airport\_Origin$, as 2. The data used in Chapter IV is summarized in Table 4.1.[7]

[7]This summary data is the daily flight level data used in dynamic pricing estimations. However, in revenue efficiency estimation, we only keep one record for each flight to avoid duplicate revenue for the same flight.

Table 4.1. Descriptive Statistics

Variable Name	Mean	St.D.	Variable Description
ln(Revenue)	10.01	0.64	Logarithm of revenue for each flight
Price	219.33	125.7	Dynamic price for each ticket
ln(Flight Time)	4.92	0.16	Logarithm of total flight time for given flight
Peak	0.57	0.5	Peak dummy, equal to 1 if during peak hour (9am to 6pm)
Domestic_Codeshare	0.52	0.61	The number of domestic codeshare partners on flight level for given carrier
Foreign_Codeshare	1.75	1.82	The number of foreign codeshare partners on flight level for given carrier
Hub_Origin	0.69	0.46	Hub dummy for origin city for given carrier
Hub_Dest	0.56	0.5	Hub dummy for destination city for given carrier
N_Airport_Origin	1.76	0.43	Number of airports in origination city that have flights for given route
N_Airport_Dest	1.89	0.31	Number of airports in destination city that have flights for given route
N_Flight_irt	16.5	7.6	Number of flights actively provided on date *t* by given carrier on given departure date on given route
N_Flight_rt	64.44	31.45	Number of flights actively provided on date *t* on given departure date on given route
N_Flight_it	68.74	29.64	Number of flights actively provided on date *t* by given carrier on given departure date
Load Factor	0.62	0.24	Daily load factor, defined as total number of unavailable seats divided by total number of seats
ADP	23.92	15.67	Advanced Days Purchased, defined as the difference between quoted date and departure date
Observations	196753		

Longer total flight time corresponds to longer distance, higher supply cost, higher supply price and thus higher revenue. Also, longer distance means that there are fewer substitutes such as automobiles and buses. This leads to higher demand and thus higher revenue. So, we expect total flight time to have positive influence on revenue and price. Codeshare agreement is a more and more popular strategy among the airline firms. By sharing the same flight, the airlines tend to improve their capacity utilization rates and reduce their operating costs. However, since codeshare partners could also be competitors, the effect from the codeshare agreement could be ambiguous. Flights from the same airline could serve as substitutes before consumers make purchase decisions and serve as complementary after they make purchase decisions and miss the purchased flight. So, the final effect is ambiguous here, depending on the tradeoff between these two conflicting effects. Flights from other carriers serve as substitutes, especially the flight during nearby hours. By examining the effect from total available flights on the same route, we can have some insight about how competition affects the revenue and dynamic pricing behavior.

Number of airports serves similarly as number of flights. However, if the number of airports increase, it would highly reduce the passenger traffic and the waiting cost for the consumers and airlines. Based on this argument, it improves the service quality. Hub is a very important determinant for the airlines. Airline hubs are airports that an airline uses as a transfer point to get passengers to their intended destinations. It is part of a hub and spoke model. The hub can highly reduce the cost, however, the effect from hub on the revenue is unclear. If an airline has a hub in an airport, it is supposed to have a large volume of traffic and thus cause some delays in the baggage waiting. However, the hub may also be good because travelers can easily transfer to other cities through hub airport. The overall effect from hub on revenue is not so clear.

In addition to above variables, load factor and ADP are two very important variables when we model dynamic pricing behaviors. By including dynamic load factor and ADP into the model, we would testify the validity of the two important theories, capacity-based yield management theories and time-based yield management theories. Load factor is calculated as the ratio of number of total unavailable seats to number of total seats, which is occupation ratio. Since this variable could be endogenous in the dynamic pricing estimation, we include the lag of load factor as the instrument to control for the endogeneity. ADP is another determinant when making pricing decisions. In Figure 4.2, we show the price changes for two flights that departed on May 10th, 2015, Flight 1648 and Flight 2207. As shown in Figure 4.2, the fares stay constant for a while and then increase sharply after 2 weeks before departure date. However, there is a price decrease around 20 days before departure date.

This finding coincides with Piga and Bachis (2007) and Deneckere and Peck (2012).

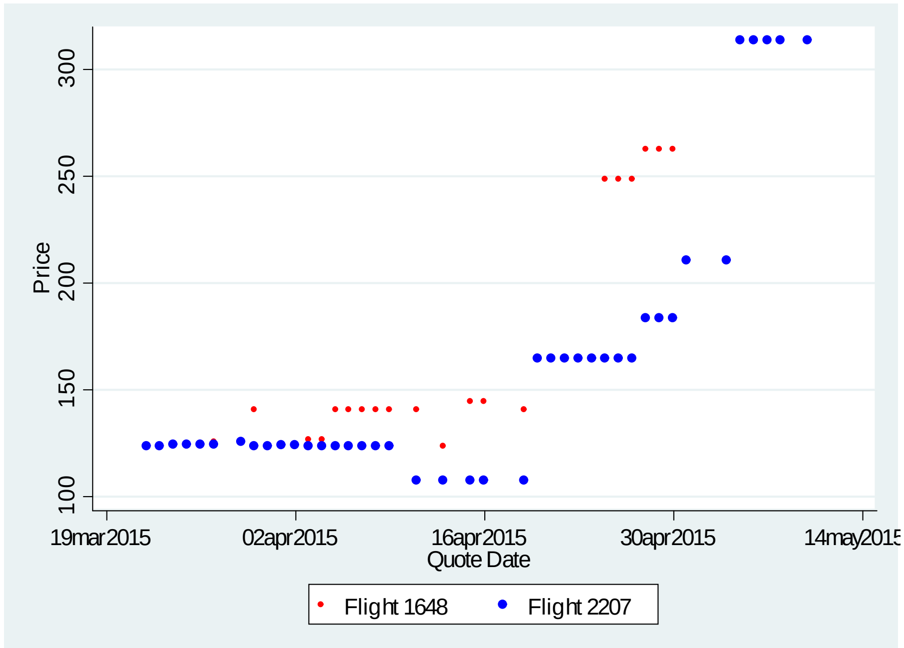

Figure 4.2 Dynamic Prices

4.4 Empirical Specifications and Results

In order to distinguish the high-efficiency firms from the low-efficiency firms, we first estimate the revenue efficiency on flight level. The observation unit in this this section is flight, that is, for each departure date, we calculate the total revenue for each flight f. Rather than airline level study, flight level data could exhibit the detailed dynamic pricing strategy for each flight. Some airlines might be more successful in dynamically making prices for peak flights, compared with off peak flights. It is possible for an airline to be better at pricing weekday departures than weekend departures. Moreover, there are many other factors that would lead to heterogeneity among the flights, leading to differences in revenue efficiencies even for the same airline, such as different operation styles in different city-pairs. Taking these concerns into consideration, we treat each flight on each departure date separately. Also, we treat same flight on different departure dates as different flights. Additionally, we treat one same flight marketed by different airlines as separate flights.[8] After aggregating

[8]Same flight marketed by different airlines are treated as different flights, because even for the same flights, different prices are offered by different airlines.

the total revenue for each flight, the empirical specification for revenue efficiency is given by

$$\begin{aligned} R\left(X_f\right) &= R^*\left(X_f\right)\exp\left(-u_f+v_f\right) \Leftrightarrow \\ \ln R\left(X_f\right) &= \ln R^*\left(X_f\right) - u_f + v_f \end{aligned} \tag{17}$$

The deterministic component of best practicing firm's revenue is given by

$$\begin{aligned} \ln R^*\left(X_f\right) &= \beta_0 + \beta_1 \ln(Flight\ Time_f) + \beta_2 Peak_f + \mathbf{N\_Airport}\boldsymbol{\beta}_3 \\ &\quad + \mathbf{Codeshare}\boldsymbol{\beta}_4 + \mathbf{N\_Flights}\boldsymbol{\beta}_5 + \mathbf{Hub}\boldsymbol{\beta}_6 + \mathbf{DOW}_f\boldsymbol{\beta}_7 \end{aligned} \tag{18}$$

Here f refers to flight; $R\left(X_f\right)$ is the total revenue for flight f; $R^*\left(X_f\right)$ denotes the best practice airline's revenue, which serves as benchmark revenue. The benchmark revenue is a function of revenue related factors, X_f^d, including total flight time ($Flight\ Time_f$), a vector of number of codeshare carriers, **Codeshare** (number of domestic codeshare carriers, $Domestic\_Codeshare$ and number of foreign codeshare carriers, $Foreign\_Codeshare$), a vector of number of airports (number of airports in origin city, $N\_Airport\_Origin$ and number of airports in destination city, $N\_Airport\_Dest$), a vector of number of flights (number of flights from the same airline on same route on the same departure date, $N\_Flight\_ir = max_t(N\_Flight\_irt)$, number of flights on the same route on given departure day, $N\_Flight\_r = max_t(N\_Flight\_rt)$ and maximum number of flights from the same carrier on given departure day $N\_Flight\_i = max_t(N\_Flight\_it)$), a vector of day of the week dummies for the departure date, **DOW** (six dummies from Monday to Saturday), peak hour dummy ($Peak$, 9am to 6pm departure time) and a vector of hub dummies, **Hub** (one for origination city and the other for destination city, denoted by Hub_Origin and Hub_Dest respectively). u_f is a one-sided error term, which captures the inefficiency part of the revenue management. Here we assume that that $u_f = h_f\tilde{u}_f$ and $\tilde{u}_f \sim \mathbf{N}^+\left(0, \sigma_u^2\right)$. Here $\sigma_u^2 = \exp\left(\mathbf{X}_u\boldsymbol{\beta}_u\right)$ and $\mathbf{X}_u$ are the factors that would affect the heterogeneity in revenue efficiency, including $Flight\ Time_f$ in this study. And $v_f \sim \mathbf{N}\left(0, \sigma_v^2\right)$ where $\sigma_v^2 = \exp\left(\beta_\varepsilon\right)$. In Stochastic Frontier Analysis, the inefficiency can be identified from the asymmetry of u_f term. v_f term is the conventional two sided error term, which follows normal distribution in this study. When measuring revenue efficiency, it is reasonable to suspect that the number of flights and hub dummies could be endogeneous variables. Higher revenue attracts entries of new airlines, which induces the endogeneity of number of flights. Moreover, it is reasonable to believe that higher revenue encourages the airlines to build a hub in the city, indicating the endogeneity of hub dummies. However since our sample only covers a very short time period (from Apr 6th to May 10th), during which the airlines could not adjust these factors, we consider these variables to be exogenous in this short time period estimation.

Based on the estimations of revenue efficiency, we show the estimation results in Table 4.2 and summarize the revenue efficiency in Table 4.3. As shown in Table 4.3, the mean value of revenue efficiency is 0.7463 and the median value is 0.7560. So, there is still large room for the airlines to improve their revenue efficiencies. As shown in Table 4.2, longer flight time is correlated with higher revenue. Longer flight time for the non-stop flights corresponds to longer distance, which indicates lower possibility to be substituted by automobile transportation or bus. This leads to higher revenue for the airlines.[9] Peak dummy, as expected, has a positive effect on revenue due to lower demand for off peak flights. The number of domestic codeshare carriers has a negative effect on revenue. On the one hand, if one carrier has more domestic codeshare carriers, the possibility of colluding with other carriers is higher, leading to higher revenue. On the other hand, domestic codeshare partners could be potential competitors. For the same flight, sometimes different ticketing carriers post different prices to compete with each other. What is more, different consumers belong to different loyalty programs, even for the same prices, different consumers might purchase from different airlines in order to accumulate the mileage for their future flights. These two conflicting effects cancel out with each other, leading to ambiguous results. Based on our estimations, the latter effect dominate the former effect, leading to negative effect from number of domestic codeshare carriers on revenue. Most of the time, foreign codeshare

[9]Of course, longer distance means higher operating cost. Since we only focus on the revenue part, not profit, we decide to ignore this effect.

partners cooperate with the domestic carriers on the domestic segment, not involving the domestic segment's market. So, the competition effect from the foreign codeshare partners is pretty low in the domestic market. Additionally, the foreign codeshare partners help to reduce the demand uncertainty for the domestic carriers. We find positive effect from number of foreign codeshare partners in this study.

Hub in origin city and destination city both improve the revenue. Hub could potentially improve operation efficiency, leading to higher revenue. Higher number of airports, both in origination city and in destination city, contributes to reduce the volume of traveler traffic and thus reduce the flight-level revenue. Number of flights from the same carrier on the same route on the same departure date as analyzed above might have ambiguous effect from two conflicting effect. Based on Table 4.2, we find positive effect and we argue that people prefer the carriers with higher flexibility. Number of flights on the same route measures the competition among different flights. Higher competition level leads to shrink of the revenue, coinciding with traditional textbook theory. If one carrier has more flights on the same day, it is more likely that there is higher possibility of delays since delays from earlier flights most likely affect the timing of later flights. Since we have $\ln(\sigma_u^2) = X_u\beta$, negative sign from flight time indicates that routes with longer distance have higher revenue efficiency. In Table 4.3, we also find large diversity in revenue efficiencies, which encourage us to explore the dynamic pricing patterns that lead to such big differences in revenue efficiency.

Table 4.2. Estimations of Revenue Efficiency

ln(Revenue)	Estimates
ln(Flight Time)	0.8532***
	(0.0311)
Peak	0.0449***
	(0.0088)
Domestic_Codeshare	-0.1200***
	(0.0084)
Foreign_Codeshare	0.0169***
	(0.0030)
Hub_Origin	0.2188***
	(0.0212)
Hub_Dest	0.0726***
	(0.0123)
N_Airport_Origin	-0.5663***
	(0.0156)
N_Airport_Dest	-0.0775***
	(0.0158)
N_Flight_ir	0.0292***
	(0.0011)
N_Flight_r	-0.0005**
	(0.0002)
N_Flight_i	-0.0043***
	(0.0003)
Day of Week Dummies	Yes
ln(sigma_u^2)	
ln(Flight Time)	-0.6802***
	(0.1630)
Constant	1.3764*
	(0.8109)
ln(sigma_v^2)	
Constant	-1.3875***
	(0.0294)
Observations	16207
Log Likelihood	-13100.914

Note: Standard errors in parentheses, clustered by each flight on each departure date. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.3. Summary of Revenue Efficiency

	Mean	Median	Min	Max
Revenue Inefficiency	0.7463	0.7560	0.2753	0.9175

In revenue efficiency estimations, we only keep one record for each flight in order to avoid duplicate revenues for the same flight. In the dynamic pricing estimations, we keep the full records for the flights that are examined in revenue efficiency part, and analyze the dynamics of pricing behavior. We categorize one flight as high-efficiency flight if its efficiency is above the median value (0.7560) and categorize it as low-efficiency flight if its efficiency is below the median value. We obtain dynamic pricing estimations separately for these two groups. The empirical specification of dynamic pricing is given by

$$\begin{aligned} P_{ft} &= \beta_0 + \beta_1 ADP_{ft} + \beta_2 LoadFactor_{ft} + \beta_3 Peak_f + \mathbf{N\_Airport}_{ft}\boldsymbol{\beta}_4 \\ &\quad + \mathbf{Codeshare}\boldsymbol{\beta}_5 + \mathbf{N\_Flight}_{ft}\boldsymbol{\beta}_6 + \mathbf{Hub}\boldsymbol{\beta}_7 + \mathbf{DOW}_f\boldsymbol{\beta}_8 + \mathbf{F}_f\boldsymbol{\beta}_f + \mathbf{R}_r\boldsymbol{\beta}_R + \varepsilon_{ft} \end{aligned} \tag{19}$$

The dependent variable here is the price for flight f on quoted date t ($ADP = t$), P_{ft}. So, for each flight, we have a series of prices. In this estimation, we control for ADP, which captures the importance of timing in dynamic pricing, corresponding to the time-based theories. Also, load factor ($LoadFactor$) is another important factor that affects the changes in prices. We would expect that higher load factor means the scarcity of the seats, leading to a higher price, indicated by capacity-based theories. Peak time flights are widely known to have much higher prices than off peak flights. So, we also control for peak dummy, $Peak$, in our estimations. Weekend flights tend to have higher price due to higher demand during weekend, so we include the weekday dummies from Monday to Saturday, **DOW**. A vector of hub dummies, **Hub** (Hub_origin and Hub_dest) is also included in the estimations. A vector of number of codeshare partners are controlled here (**Codeshare**), including number of domestic codeshare and foreign codeshare partners ($Domestic\_codeshare$ & $Foreign\_codeshare$). A vector of number of flights, **N_Flight**, is also controlled in this estimation, including N_flight_irt, N_flight_it, N_flight_rt.

These number of flights capture the dynamic number of available flights.[10] A vector of number of airports, **N_Airport** ($N\_Airport\_Origin$ and $N\_Airport\_Dest$) is also included in the estimations. Firm dummies ($\mathbf{F}_f$) and route dummies ($\mathbf{R}_r$) are controlled in this estimation.

The estimation results are given in Table 4.4, which has four columns. The first column is the high-efficiency flights' estimations with the assumption of exogenous load factor, the second column is the high-efficiency flights' estimation with the assumption of endogenous load factor. The third column and fourth column correspond to the estimations for low-efficiency group when the load factor is assumed to be exogenous and endogenous, respectively. Since the endogeneity tests indicate that there is endogeneity problem, we mainly focus on the estimations that treat load factor as endogenous variable. As mentioned above, the lag value of load factor is the instrumental variable for load factor. The exogenous version estimates are provided here as robustness check. From Table 4.4, we find a positive relationship between load factor and price for both low-efficiency flights and high-efficiency flights. However, the coefficient is not significant at 10% significance level except for the low-efficiency group with exogenous load factor. Based on this, we conclude that we only find weak evidence of capacity-based theories. An airline with higher load factor tends to price higher due to the scarcity of the remaining seats for the low-efficiency flights. For the low-efficiency flights, in column three, we find that 10 percent increase in load factor would lead to about \$4 price change for one ticket. This is the average effect since 60 days before departure. We would suspect that this effect could be larger when departure date nears.[11] Compared with the high-efficiency flights, it looks like that the low-efficiency flights might be too responsive to the capacity change, leading to smaller market share, compared with high-efficiency flights. It is possible that the low-efficiency flights want to optimize their revenues by frequently change their prices. However, this could lead to suboptimization of the revenue. We only find weak evidences of capacity-based theory from Table 4.4.[12]

[10] These three variables are different from what we used in revenue efficiency.
[11] We leave this topic to future studies.
[12] Table 4.5 also shows similar results.

As departure date nears, the price level goes up for both high-efficiency flights and low-efficiency flights. We do not find statistically significant differences between high-efficiency flights and low-efficiency flights on time-based pricing patterns since the difference between the estimates for ADP is not significant at 10% significance level. Most of the travelers do not learn their exact preferences about flight time until a couple of days before departure. The business travelers do not want to sacrifice their time flexibility for the price differences. So, the airlines charge much higher to the last minute buyers, who have lower demand elasticities and higher budgets. Flights during peak hours tend to have much higher price. Generally speaking, the price different between peak flights and off peak flights is around \$9. As mentioned above, the domestic codeshare partners mostly act as competitors while foreign codeshare partners help to reduce the demand uncertainty. Hub in origination city plays little role in price, but the hub in destination city has negative effect on price. One the one hand, hub provides high convenience for the consumers to transfer to other cities. On the other hand, hub means heavy traffic and longer waiting time for the baggage if the hub is in destination city. From Table 4.4, we find that the consumers would like to pay less if they expect to wait longer for their luggage. The three types of number of flights affect dynamic pricing in the same way as they affect revenue efficiency.

Table 4.4. Comparison of Dynamic Pricing between High-efficiency with Low-efficiency Flights

Price	High Efficiency (1)	High Efficiency (2)	Low Efficiency (1)	Low Efficiency (2)
Load Factor	17.8474	14.8241	39.8553**	11.2583
	(18.5738)	(17.1821)	(15.9798)	(16.1958)
ADP	-2.4563***	-2.4541***	-2.4515***	-2.6186***
	(0.4982)	(0.5394)	(0.2124)	(0.2322)
Peak	8.3341***	8.3171***	8.7528***	9.8216***
	(2.4754)	(2.6074)	(1.5450)	(1.4435)
Domestic_Codeshare	-13.0996**	-12.8836**	-14.1333*	-11.4425*
	(6.1415)	(5.6328)	(7.7737)	(6.5159)
Foreign_Codeshare	2.5927**	2.5952**	0.7564	0.6900
	(1.0583)	(1.0320)	(1.8116)	(1.6110)
Hub_Origin	0.0000	39.4922	0.0000	-25.3397
	(0.0000)	(36.7955)	(0.0000)	(22.1735)
Hub_Dest	-28.7831***	-28.2012***	-24.4487**	-22.1744**
	(6.0983)	(6.5010)	(11.8082)	(11.3060)
N_Flight_irt	2.0894***	2.1586***	0.8131	1.1549
	(0.7038)	(0.6715)	(0.9591)	(0.9031)
N_Flight_rt	-1.3238***	-1.3885***	-0.3592*	-0.4348**
	(0.2642)	(0.2776)	(0.2068)	(0.2074)
N_Flight_it	-0.8639*	-0.8638*	-0.1183	-0.1252
	(0.4530)	(0.4512)	(0.1811)	(0.1815)
Day of Week Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	98411	92800	98342	93838
R Square	0.5128	0.5352	0.3476	0.4677

Note: Standard errors in parentheses, clustered by each flight on each departure date. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Since timing plays a very important role in dynamic pricing as shown in Table 4.4, we further categorize the ADP into 0-6 day, 7-13 days, 14-21 days and 22 days plus to explore the detailed dynamic pricing patterns based on timing. In Table 4.5, we put these dummy variables into the regression. We find that there are large price differences if one purchase the ticket at different time. To be specific, the travelers who buy tickets in the last week would pay \$140 more if they buy from high-efficiency flights, compared with buying the same ticket three weeks before the departure. This price difference further increases to \$159 for the low-efficiency flights. Again, we did not find statistically significant differences for the periods of ADP0-6 and ADP7-13. However, we find some differences in time period of

ADP14-21, that is, the low-efficiency flights seem to be more sensitive to timing during this time period compared with the high-efficiency flights. The estimates for ADP14-21 is not significant for high-efficiency estimations at 10% significance level. The higher sensitivity to timing for the low-efficiency flights would make the travelers switch to other flights. Also, the estimation results from Table 4.5 validate the robustness of estimations in Table 4.4. Again, same as Table 4.4, we only find weak evidences for capacity-based theories from low-efficiency flights. Generally speaking, by comparing the high-efficient flights with low-efficiency ones, we provide some insights for the low-efficiency flights about how to change their pricing strategies so as to improve their revenue efficiencies and thus total profits. It is possible that the low efficiency might come from the excessive response to the changes in available seats and timing, leading to loss in market share. The low-efficiency flights might also underestimate the elasticity of demand of the travelers, or they underestimate the substitute possibilities from other flights. Another explanation is that low-efficiency flights try to change their prices more frequently in order to fix inefficiency problem. However, the frequent changes make the low-efficiency flights hard to fully optimize their changes. For instance, it gets harder to segment customers for the right prices when there are frequent price changes. According to *Customer Satisfaction Report* for online travel agency in 2014, 66% of the consumers put prices as the primary reason that customers reserve tickets from an online travel agency. So, too responsive changes in price would scare off the customers, leading to loss in market share and thus revenue. So, based on this finding, it would be beneficial for the low-efficiency flights/airlines to adjust their computer reservation systems so that the prices do not react so sharply to the available capacity changes and timing changes. Also, the low-efficiency flights should think about how the frequent changes of the price damage the revenue efficiencies.

Table 4.5. Comparison of Dynamic Pricing between High-efficiency with Low-efficiency Flights V2

Price	High Efficiency (1)	High Efficiency (2)	Low Efficiency (1)	Low Efficiency (2)
Load Factor	21.4311	16.2360	40.0366**	21.0103
	(17.6512)	(16.0385)	(15.8822)	(15.9211)
ADP0-6	140.2287***	143.5072***	158.9166***	163.2612***
	(24.5199)	(26.2268)	(11.1958)	(11.6413)
ADP7-13	43.8342***	45.0737***	46.2882***	50.0924***
	(13.1233)	(14.5414)	(6.5316)	(6.9091)
ADP14-21	11.2864	11.8438	6.7886**	9.1178**
	(7.7311)	(8.5181)	(3.3044)	(3.6538)
Peak	8.0297***	8.1272***	8.5657***	9.3403***
	(2.5416)	(2.6589)	(1.4163)	(1.3549)
Domestic_Codeshare	-13.7461**	-13.5266**	-15.2643**	-13.6047**
	(6.6176)	(6.0657)	(7.1470)	(6.2039)
Foreign_Codeshare	2.7719**	2.7600**	0.6908	0.6315
	(1.1375)	(1.1068)	(1.6946)	(1.5468)
Hub_Origin	0.0000	59.8544**	0.0000	-22.7134
	(0.0000)	(27.6042)	(0.0000)	(22.0840)
Hub_Dest	-29.7924***	-29.0121***	-25.1978**	-23.8226**
	(6.1721)	(6.4733)	(11.1660)	(10.8333)
N_Flight_irt	2.0589***	2.1610***	0.8051	1.0161
	(0.6855)	(0.6722)	(0.9468)	(0.8937)
N_Flight_rt	-1.2591***	-1.3195***	-0.3674*	-0.4067**
	(0.2616)	(0.2714)	(0.2121)	(0.2057)
N_Flight_it	-0.9280**	-0.9234**	-0.1567	-0.1738
	(0.4202)	(0.4144)	(0.1384)	(0.1339)
Day of Week Dummies	Yes	Yes	Yes	Yes
Firm Dummies	Yes	Yes	Yes	Yes
Route Dummies	Yes	Yes	Yes	Yes
Observations	98411	92800	98342	93838
Centered R Square	0.5493	0.4739	0.5352	0.4677

Note: Standard errors in parentheses, clustered by each flight on each departure date. * p < 0.1, ** p < 0.05, *** p < 0.01.

4.5 Summary and Concluding Remarks

Using a unique self-collected data from *priceline.com*, this chapter examines the revenue efficiency on the flight level. Employing the stochastic frontier analysis estimations, we obtain the revenue efficiencies of the flights. The revenue efficiencies between different flights are shown to be highly diverse. By comparing the dynamic pricing patterns between high-efficiency flights with the low-efficiency flights, we find that low-efficiency flights response more sensitively to the available capacity changes and timing change. To be specific, the prices of low-efficiency flights respond more sensitively than high-efficiency flights from 3 weeks to 2 weeks before departure. The frequently changes in price would lead to suboptimization of the revenue for these flights. Our suggestions for the low-efficiency flights is that they might try to adjust their computer reservation systems so that the prices are less responsive to the load factor changes and timing changes. Or, they should not provide so frequent changes in prices, which might cause loss in suboptimization of prices.

Also, in this chapter, we find evidences for both capacity-based theories and time-based theories. To be specific, higher load factors relates to higher price levels. And prices increase as departure nears, especially for the last minute buyers. However, we do not find strong evidences for capacity-based theory for the high-efficiency flights. We find some evidences for capacity-based theory from low-efficiency flights. The evidences for time-based theory are robust to both high-efficiency flights and low-efficiency flights.

CHAPTER V

CONCLUSIONS

This thesis studies market power, cost efficiency and pricing strategies in the airline industry. In Chapter II, this thesis builds up a theoretical framework that enables the estimation of marginal cost efficiency without total cost data. Chapter III analyzes the factors that affect the price dispersion based on the estimates obtained from Chapter II. Chapter IV provides an analysis of the revenue efficiency and dynamic pricing patterns for the U.S. domestic airlines. So, this thesis provides a comprehensive study of the market competition, cost efficiency and pricing strategies in the airline industry.

Chapter II develops a theoretical framework that allows simultaneous estimation of marginal cost efficiency and conduct parameter without total cost data. This is the first study that enables estimation of marginal cost efficiency without total cost data, to our best knowledge. Traditional SFA obtains the cost efficiency from the cost function, which requires total cost data. Lerner Index also has the same problem. Conduct parameter framework derives the "perceived marginal revenue" from the demand and supply system. In Chapter II, this thesis combines the conduct parameter framework with SFA literature and develops a conduct parameter based model to estimate marginal cost efficiency. By doing this, this thesis provides a theoretical framework that enables us to estimate marginal cost efficiency and conduct parameter simultaneously. The airline data is employed in this study to testify the theoretical framework in Chapter II. The airline data is collected from DB1B, Form 41 and T100. And the time period covered by Chapter II is 1999I-2009IV and we only focus on coach class tickets on Chicago based routes. Using control function approach to deal with the endogeneity problem, this thesis validates the theoretical framework provided in Chapter II. In Chapter II, we find the cost efficiency differences among the firm-route pairs. Moreover, Chapter II explores the relationship between conduct parameter and marginal cost efficiency and finds evidence for QLH from the airline data.

Chapter III studies the determinants of price dispersion. In particular, we are interested in how market power and marginal cost efficiency affect the price dispersion. Based on the estimates of conduct parameter and marginal cost efficiency from Chapter II, Chapter III examines the direction of conduct and marginal cost efficiency's effect on price dispersion. The finding of Chapter III supports Borenstein and Rose (1994) that generally the market power, proxied by conduct, has a negative effect on price dispersion for all routes sample. Moreover, we find that for big city routes, market power has a positive effect on price dispersion while the effect is negative for leisure routes. Additionally, we find that the effect of marginal cost efficiency is negative. We find negative relationship between conduct and marginal cost efficiency in Chapter II. The instrumental variable used in previous studies, geometric market share, is correlated with marginal cost efficiency (the omitted variable) in previous studies. Chapter III also examines whether this correlation leads to bias in previous studies. We only find weak evidences for overestimation. In Chapter III, we also compare conduct with HHI (widely used in previous studies) and we find that conduct is preferred to HHI since HHI does not contain information about elasticity of demand and marginal cost efficiency information.

Chapter IV analyzes the dynamic pricing patterns for the airlines. The daily dynamic pricing data and its associated load factor information are scripted from *priceline.com* using *Perl* program. This unique dataset allows us to explore how the airlines make dynamic pricing decisions, especially based on available seats left and advanced days purchased. Chapter II and Chapter III employ the post sale data from DB1B, while Chapter IV utilizes the data scripted from *priceline.com*. In Chapter IV, this study first analyzes the revenue efficiencies of the flights. From here, we find large differences in the revenue efficiency among different flights. The mean value of revenue efficiency is 0.75, which means that the airlines still have a very large room for improvement. By dividing the flights into high-efficiency and low-efficiency flights, we compare the different dynamic pricing patterns for these two groups, focusing on load factor and ADP. Also, we find weak evidences for capacity-based theories and stronger evidence for time-based theories.

APPENDIX A

DATA AND VARIABLES

A.1 Details in data construction

In this Appendix we provide the details in the data construction process. First, all multi-destination tickets are dropped as it is difficult to identify the ticket's origin and destination without knowing the exact purpose of the trip. Second, any itinerary involving international flights is eliminated. Third, we adjust the fare class for high-end carrier. That is, for some airlines, due to marketing strategies, only high-end (first class and business class) tickets are provided to consumers on all routes, especially some small airlines. However, the quality should be taken as coach class. Therefore, we consider all such tickets as coach class tickets. In different time periods, due to changes in the pricing strategy, sometimes high-end-only carrier switches to a regular carrier which sells both coach class tickets and high-end tickets. For instance, Sun Country Airlines does not provide coach class tickets in 2001 but provides coach tickets in 2005 and years after. Hence, we treat the tickets in each quarter separately when considering the adjustment. That is, we treat high-end tickets from Sun Country Airlines as coach class tickets in 2001, not in year 2005. Fourth, tickets that have high-end segments and unknown fare classes are dropped. We followed Borenstein (1989) and Brueckner, Dyer, and Spiller (1992) by using ticketing carrier as our airline as an observation unit. Also, firm specific average segment numbers, average stage length on a given route are calculated after further elimination of multi-ticketing-carrier tickets. Moreover, our data set includes a distance variable which is the shortest directional flight distance.

A.2 Calculation of Gini Coefficient

The $GINI$ variable is calculated as:

$$\begin{aligned} GINI &= 1 - 2(\sum_{i=1,N} ((fare_i * \frac{PAX_i}{REVENUE}) \\ &\quad *(\frac{1}{2}\frac{PAX_i}{PAX} + (1 - \sum_{j=1,i} \frac{PAX_j}{PAX}))) \end{aligned}$$

where N is the number of different fare level tickets by a certain carrier on certain route in certain quarter, $fare_i$ is the reported fare for the ith ticket, PAX_i is the number of passengers purchased at $fare_i$, $REVENUE$ is the total revenue by certain carrier on certain route in certain quarter and PAX is the total passenger number that traveled by a certain carrier on certain route in certain quarter.

A.3 Calculation of Geometric Market Share

The GEO variable is defined as:

$$GEO = \frac{\sqrt{ENP_{xo}ENP_{xd}}}{\sum_y \sqrt{ENP_{yo}ENP_{yd}}}$$

where x is the observed airline, y indexes all airlines and ENP_{xo} and ENP_{xd} are the quarterly enplanement at origin city and destination city by airline x.

REFERENCES

[1] Aigner, D.J., Lovell, C.A.K. and Schmidt, P.J. (1977), Formulation and Estimation of Stochastic Frontier Production Function Models, Journal of Econometrics, 6, 21-37.

[2] Alderighi, Marco & Nicolini, Marcella and Piga, Claudio A. (2012), Combined Effects of Load Factors and Booking Time on Fares: Insights from the Yield Management of a Low-Cost Airline, FEEM Working Paper No. 14.2012.

[3] Almanidis, P. and Kutlu, L. (2015), A Dynamic Stochastic Frontier Model with Threshold Effects, Unpublished manuscript.

[4] Bandyopadhyay, D. and Das, A. (2006), On Measures of Technical Inefficiency and Production Uncertainty in Stochastic Frontier Production Model with Correlated Error Components, Journal of Productivity Analysis, 26, 165-180.

[5] Barros, Pedro Pita. (1999), Multimarket competition in banking, with an example from the Portuguese market, International Journal of Industrial Organization, 17, 335-352.

[6] Barut, M., & Sridharan, V. (2004), Design and evaluation of a dynamic capacity apportionment procedure, European Journal of Operational Research, 155(1), 112-133.

[7] Battese, G.E. and Coelli, T.J. (1992), Frontier Production Functions, Technical Efficiency and Panel Data with Application to Paddy Farmers in India, Journal of Productivity Analysis, 3, 153-169.

[8] Berger, A. and Hannan, T.H (1998), The Efficiency Cost of Market Power in the Banking Industry: A Test of the "Quiet Life" and Related Hypothesis, Review of Economics and Statistics, 80, 454-465.

[9] Berry, S., & Jia, P. (2008), Tracing the woes: An empirical analysis of the airline industry. National Bureau of Economic Research Working Paper No. w14503.

[10] Borenstein, S. (1989), Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry, RAND Journal of Economics, 20, 344-365.

[11] Borenstein, Severin (2005), U.S. Domestic Airline Pricing, 1995-2004, Berkeley, CPC Working Paper No. CPC05-48.

[12] Borenstein, S. (2011), On the Persistent Financial Losses of U.S. Airlines: A Preliminary Exploration, Working paper.

[13] Borenstein, S., Bushnell, J., Kahn, E., and Stoft, S. (1996), Market Power in California Electricity Markets, Utilities Policy, 5, 219-236.

[14] Borenstein, S. and Rose, N.L. (1994), Competition and Price Dispersion in the U.S. Airline Industry, Journal of Political Economy, 102, 653-683.

[15] Borenstein, Severin & Rose, Nancy (1995), Bankruptcy and Pricing Behavior in U.S. Airline Markets, American Economic Review, 85(2), 397-402.

[16] Borenstein, S., & Rose, N. L. (2007). How airline markets work... or do they? Regulatory reform in the airline industry, National Bureau of Economic Research Working Paper No. w13452.

[17] Brander, J.A. and Zhang, A. (1990), Market Conduct in the Airline Industry: An Empirical Study, RAND Journal of Economics, 21, 567-583.

[18] Bresnahan, T.F. (1982), The Oligopoly Solution is Identified, Economics Letters, 10, 87-92.

[19] Bresnahan, T.F. (1989), Studies of Industries with Market Power, The Handbook of Industrial Organization. Amsterdam: North-Holland.

[20] Brooks, M., & Button, K. J. (1994), Yield Management: A Phenomenon of the 1980s and 1990s?, International Journal of Transport Economics, 177-196.

[21] Brueckner, J., Dyer, N., Spiller, P.T. (1992), Fare Determination in Airline Hub-and-Spoke Networks, RAND Journal of Economics, 23, 309-333.

[22] Busse, Meghan (2002), Firm Financial Condition and Airline Price Wars, RAND Journal of Economics, 33(2), 298-318.

[23] Chakrabarty, D. and Kutlu, L. (2014), Competition and Price Discrimination in the Airline Market, Applied Economics, 46, 3421-3436.

[24] Chen, Y., Iyer, G., and Pazgal, A. (2010), Limited Memory, Categorization, and Competition, Marketing Science, 29, 650-670.

[25] Chi, J., Koo, W.W., and Lim, S.H. (2013), Carriers Price Dispersion in the U.S. Airline Industry, Working paper.

[26] Ciliberto, Federico & Schenone, Carola (2008), Bankruptcy and Product-Market Competition: Evidence from the Airline Industry, University of Virginia Working Paper.

[27] Comanor, W.S. and Leibenstein, H. (1969), Allocative Efficiency, X-Efficiency and the Measurement of Welfare Losses, Economica, 36, 304-309.

[28] Cornwell, C., Schmidt, P., and Sickles, R.C. (1990), Production Frontiers with Time-Series Variation in Efficiency Levels, Journal of Econometrics, 46, 185-200.

[29] Corts, K.S. (1999), Conduct Parameters and the Measurement of Market Power, Journal of Econometrics, 88, 227-250.

[30] Dana, Jr, J. D. (1998), Advance-purchase discounts and price discrimination in competitive markets, Journal of Political Economy, 106(2), 395-422.

[31] Dana Jr, J. D. (1999a), Using yield management to shift demand when the peak time is unknown, The Rand Journal of Economics, 456-474.

[32] Dana Jr, J. D. (1999b), Equilibrium price dispersion under demand uncertainty: the roles of costly capacity and market structure, The RAND Journal of Economics, 632-660.

[33] Dana, J. D., & Orlov, E. (2009), Internet penetration and capacity utilization in the US airline industry, Northeastern U. College of Business Administration Research Paper, (08-001).

[34] Davis, Paul 1994, Airline ties profitability yield to management, SIAM News, 27(5).

[35] Debreu, G. (1951), The Coefficient of Resource Utilization, Econometrica, 19, 273-292.

[36] Delis, M.D., Tsionas, E.G. (2009), The Joint Estimation of Bank-Level Market Power and Efficiency, Journal of Banking and Finance, 33, 1842-1850.

[37] Demsetz, H., (1973), Industry Structure, Market Rivalry, and Public Policy, Journal of Law and Economics, 16, 1-9.

[38] Deneckere, R., & Peck, J. (2012), Dynamic competition with random demand and costless search: A theory of price posting, Econometrica, 80(3), 1185-1247.

[39] Duygun, M., Kutlu, L., and Sickles, R.C. (2014), Measuring Efficiency: A Kalman Filter Approach, Unpublished manuscript.

[40] Eden, B. (1990), Marginal cost pricing when spot markets are complete, Journal of Political Economy, 1293-1306.

[41] Escobari, D. (2009), Systematic peak-load pricing, congestion premia and demand diverting: Empirical evidence, Economics Letters, 103(1), 59-61.

[42] Escobari, D., & Gan, L. (2007), Price dispersion under costly capacity and demand uncertainty, National Bureau of Economic Research Working Paper No. w13075.

[43] Escobari, D. (2012). Dynamic Pricing, Advance Sales and Aggregate Demand Learning in Airlines. The Journal of Industrial Economics, 60(4), 697-724.

[44] Farrell, M.J. (1957), The Measurement of Productive Efficiency, Journal of the Royal Statistical Society, 120, 253-281.

[45] Forbes, Silke Januszewski (2008), The Effect of Air Traffic Delays on Airline Prices, International Journal of Industrial Organization, 26(5), 1218-1232.

[46] Gaggero, A. A., & Piga, C. A. (2011), Airline market power and intertemporal price dispersion. The Journal of Industrial Economics, 59(4), 552-577.

[47] Gale, I. L., & Holmes, T. J. (1992), The efficiency of advance-purchase discounts in the presence of aggregate demand uncertainty, International Journal of Industrial Organization, 10(3), 413-437.

[48] Gale, I. L., & Holmes, T. J. (1993), Advance-purchase discounts and monopoly allocation of capacity, The American Economic Review, 135-146.

[49] Gerardi, K. S., & Shapiro, A. H. (2009), Does competition reduce price dispersion? New evidence from the airline industry, Journal of Political Economy, 117(1), 1-37.

[50] Genesove, D. and Mullin, W. (1998), Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914, RAND Journal of Economics, 29, 355-77.

[51] Good, D., Sickles, R.C., and Weiher, J. (2008), A Hedonic Price Index for Airline Travel, Review of Income and Wealth, 54, 438-465.

[52] Goolsbee, A., & Syverson, C. (2008), How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines, Quarterly. In Journal of Economics.

[53] Greene, W. (1980a), Maximum Likelihood Estimation of Econometric Frontier Functions, Journal of Econometrics, 13, 27-56.

[54] Greene, W. (1980b), On the Estimation of a Flexible Frontier Production Model, Journal of Econometrics, 3, 101-115.

[55] Greene,W. (2003), Simulated Likelihood Estimation of the Normal-Gamma Stochastic Frontier Function, Journal of Productivity Analysis, 19, 179-190.

[56] Guan, Z., Kumbhakar, S.C., Myers, R.J., and Lansink, A.O. (2009), Measuring Excess Capital Capacity in Agricultural Production, American Journal of Agricultural Economics, 91, 765-776.

[57] Harvey, A.C. (1989), Forecasting Structural Time Series Models and the Kalman Filter, Cambridge: Cambridge University Press.

[58] Hazledine, T. (2006), Price Discrimination in Cournot-Nash Oligopoly, Economics Letters, 93, 413-20.

[59] Hicks, J.R. (1935), Annual Survey of Economic Theory: The Theory of Monopoly, Econometrica, 3, 1-20.

[60] Ito, H. and Lee, D. (2007), Domestic Code Sharing, Alliances, and Airfares in the U.S. Airline Industry, Journal of Law and Economics, 50, 355-380.

[61] Iwata, G., (1974), Measurement of Conjectural Variations in Oligopoly, Econometrica, 42, 947-966.

[62] Lerner, A.P. (1934), The Concept of Monopoly and the Measurement of Monopoly Power, Review of Economic Studies, 1, 157-175.

[63] Jondrow, J., Lovell, C.A.K., Materov, I.S., and Schmidt P. (1982), On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model, Journal of Econometrics, 19, 233-238.

[64] Karakaplan, M.U. and Kutlu, L. (2015), Handling Endogeneity in Stochastic Frontier Models, Unpublished manuscript.

[65] Kim, C.J. (2006), Time-Varying Parameter Models with Endogenous Regressors, Economics Letters, 91, 21-26.

[66] Kim, C.J. and Nelson , C.R. (2006), Estimation of a Forward-Looking Monetary Policy Rule: A Time-Varying Parameter Model Using Ex Post Data, Journal of Monetary Economics, 53, 1949-1966.

[67] Kim, D. (2005), Measuring Market Power in a Dynamic Oligopoly Model: An Empirical Analysis, Working paper.

[68] Kim, Y. and Kim C.J. (2011), Dealing with Endogeneity in a Time-Varying Parameter Model: Joint Estimation and Two-Step Estimation Procedures, Econometrics Journal, 14, 487-497.

[69] Kimes, S. E. (1989), Yield management: a tool for capacity-considered service firms, Journal of operations management, 8(4), 348-363.

[70] Kimes, S. E. (2002), Perceived fairness of yield management, The Cornell Hotel and Restaurant Administration Quarterly, 43(1), 21-30.

[71] Kimes, S. E., & Wirtz, J. (2003)., Has revenue management become acceptable? Findings from an international study on the perceived fairness of rate fences, Journal of Service Research, 6(2), 125-135.

[72] Koetter, M., Kolari, J.W., and Spierdijk, L. (2012), Enjoying the quiet life under deregulation? Evidence from adjusted Lerner indices for US banks, Review of Economics and Statistics, 94, 462-480.

[73] Koetter, M. and Poghosyan, T. (2009), The Identification of Technology Regimes in Banking: Implications for the Market Power-Fragility Nexus, Journal of Banking and Finance, 33, 1413-1422.

[74] Kumbhakar, S.C. and Lovell, C.A.K. (2000), Stochastic Frontier Analysis. Cambridge University Press.

[75] Kumar, R. and Kutlu, L. (2015), Price Discrimination in Quantity Setting Oligopoly, Forthcoming in Manchester School.

[76] Kutlu, L. (2009), Price discrimination in Stackelberg competition, Journal of Industrial Economics, 57, 364.

[77] Kutlu, L. (2010), Battese-Coelli Estimator with Endogenous Regressors, Economics Letters, 109, 79-81.

[78] Kutlu, L. (2012a), Price discrimination in Cournot competition, Economics Letters, 117, 540-543.

[79] Kutlu, L. (2012b), U.S. banking efficiency, 1984-1995, Economics Letters, 117, 53-56.

[80] Kutlu, L. (2013), Misspecification in Allocative Inefficiency: A Simulation Study, Economics Letters, 118, 151-154.

[81] Kutlu, L. (2015a), Limited Memory Consumers and Price Dispersion, Review of Industrial Organization, 46, 349-357.

[82] Kutlu, L. (2015b), Time-varying true individual effects with endogenous regressors, Working Paper.

[83] Kutlu, L., and Sickles C.R. (2012), Estimation of Market Power in the Presence of Firm Level Inefficiencies, Journal of Econometrics, 168, 141-155.

[84] Kutlu, L., and Sickles C.R. (2015), Measuring Market Power When Firms Price Discriminate, Unpublished manuscript.

[85] Kutlu, levent & Wang, Ran (2015a), Estimation of Cost Efficiency Without Cost Data. Working Paper.

[86] Kutlu, levent & Wang, Ran (2015b), Price Dispersion, Competition, and Effciency in the U.S. Airline Industry. Working Paper.

[87] Kumbhakar, S.C., Lovell, C.A.K. (2000), Stochastic Frontier Analysis. Cambridge University Press.

[88] Kumbhakar, S.C., and Wang, H.J. (2006), Pitfalls in the Estimation of a Cost Function That Ignores Allocative Inefficiency: A Monte Carlo Analysis, Journal of Econometrics, 134, 317-340.

[89] Lau, L.J. (1982), On Identifying the Degree of Competitiveness from Industry Price and Output Data, Economics Letters, 10, 93-99.

[90] Lee, C.Y. and Johnson, A.L. (2012), Rational Inefficiency: A Nash-Cournot Oligopolistic Market Equilibrium, Working paper.

[91] Lerner, A.P. (1934), The Concept of Monopoly and Measurement of Monopoly Power, Review of Economic Studies, 1, 157-175.

[92] Mantin, B., & Koo, B. (2009), Dynamic price dispersion in airline markets, Transportation Research Part E: Logistics and Transportation Review, 45(6), 1020-1029.

[93] Maudos, J. and Fernandez de Guevara, J. (2007), The Cost of Market Power in Banking: Social Welfare Loss vs. Cost Efficiency, Journal of Banking and Finance, 31, 2103-2125.

[94] McAfee, R. P., & Te Velde, V. (2006), Dynamic pricing in the airline industry. Handbook on Economics and Information Systems, Ed: TJ Hendershott, Elsevier.

[95] McGill, J. I., & Van Ryzin, G. J. (1999), Revenue management: Research overview and prospects, Transportation science, 33(2), 233-256.

[96] Meeusen, W., and van den Broeck, J. (1977), Efficiency Estimation from Cobb-Douglas Production Function with Composed Error, International Economic Review, 8, 435-444.

[97] Oun, T.H., Zhang A., and Zhang, Y. (1993), Inter-Firm Rivalry and Firm-Specific Price Elasticities in Deregulated Airline Markets, Journal of Transportation Economics and Policy, 27, 191-192.

[98] Perloff, J. M., Karp, L.S., and Golan, A. (2007), Estimating Market Power and Strategies, Cambridge University Press, Cambridge.

[99] Perloff, J.M. and Shen, E.Z. (2012), Collinearity in Linear Structural Models of Market Power, Review of Industrial Organization, 40, 131-138.

[100] Phillips, R. (2005), Pricing and revenue optimization, Stanford University Press.

[101] Piga, C., & Bachis, E. (2007), Hub premium, airport dominance and market power in the European airline industry. Working Paper.

[102] Pinder, J. P. (1995), A revenue management approach to demand management and order booking in assemble-to-order manufacturing, Journal of Operations Management, 13(4), 299-309.

[103] Prescott, E. C. (1975), Efficiency of the natural rate, The Journal of Political Economy, 1229-1236.

[104] Puller, S.L. (2007), Pricing and Firm Conduct in California s Deregulated Electricity Market, Review of Economics and Statistics, 89, 75-87.

[105] Puller, S.L. (2009), Estimation of Competitive Conduct when Firms are Efficiently Colluding: Addressing the Corts Critique, Applied Economics Letters, 1497-1500.

[106] Puller, S. L., Sengupta, A., & Wiggins, S. N. (2009). Testing theories of scarcity pricing in the airline industry. National Bureau of Economic Research Working Paper No. w15555.

[107] Rotemberg, J.J. and Saloner, G. (1986), A Supergame-Theoretic Model of Price Wars During Booms, American Economic Review, 76, 390-407.

[108] Schmidt, P. and Sickles, R.C. (1984), Production Frontiers and Panel Data, Journal of Business and Economic Statistics, 2, 367-374.

[109] Sengupta, A., & Wiggins, S. N. (2006), Airline pricing, price dispersion and ticket characteristics on and off the internet, NET Institute Working Paper, (06-07).

[110] Sickles, R.C. (2005), Panel Estimators and the Identification of Firm-Specific Efficiency Levels in Semi-Parametric and Non-Parametric Settings, Journal of Econometrics, 126, 305-324.

[111] Smith, B. C., Guther, D. P., Rao, B. V., & Ratlife, R. M. (2001), E-commerce and operations research in airline planning, marketing, and distribution, Interfaces, 31(2), 37-55.

[112] Smith, M.D. (2008), Stochastic Frontier Models with Dependent Error Components, Econometrics Journal, 11, 172-192.

[113] Sridharan, S. V. (1998), Managing capacity in tightly constrained systems, International Journal of Production Economics, 56, 601-610.

[114] Stahl, D. O. (1989), Oligopolistic pricing with sequential consumer search, The American Economic Review, 700-712.

[115] Stahl, D. O. (1996), Oligopolistic pricing with heterogeneous consumer search, International Journal of Industrial Organization, 14(2), 243-268.

[116] Stavins, J. (2001), Price Discrimination in the Airline Market: The Effect of Market Concentration, Review of Economics and Statistics, 83, 200-202.

[117] Stevenson, Rodney E. (1980), Likelihood Functions for Generalized Stochastic Frontier Estimation, Journal of Econometrics, 13, 57-66.

[118] Talluri, K., & Van Ryzin, G. (2004), Revenue management under a general discrete choice model of consumer behavior, Management Science, 50(1), 15-33.

[119] Talluri, K. T., & Van Ryzin, G. J. (2006), The theory and practice of revenue management, Springer Science & Business Media.

[120] Tran, K.C., and Tsionas, E.G. (2013), GMM Estimation of Stochastic Frontier Model with Endogenous Regressors, Economics Letters, 118, 233-236.

[121] Varian, H.R. (1980), A Model of Sales, American Economic Review, 70, 651-659.

[122] Weiher, J.C. (2002), Pricing Issues in the United States Airline Industry, Unpublished manuscript.