Price Formats as a Source of Price Dispersion: A Study of Online and Offline Prices in the Domestic U.S. Airline Markets

Ramnath K. Chellappa
Goizueta Business School, Emory University, Atlanta, Georgia 30322, ram@bus.emory.edu

Raymond G. Sin
School of Business and Management, Hong Kong University of Science and Technology,
Clear Water Bay, Hong Kong, rsin@ust.hk

S. Siddarth
Marshall School of Business, University of Southern California, Los Angeles, California 90089,
siddarth@usc.edu

A large body of research in economics, information systems, and marketing has sought to understand sources of price dispersion. Previous empirical work has mainly offered consumer- and/or product-based explanations for this phenomenon. In contrast, our research explores the key role played by vendors' price-format adoption in explaining price dispersion. We empirically analyze over a half-million online and offline prices offered by major U.S. airlines in the top 500 domestic markets. Our study shows that a vendor's price format remains an important source of price dispersion in both channels *even* after accounting for other factors known to impact dispersion in airline ticket prices. Importantly, this finding is true for both transacted and posted tickets. We document several other interesting empirical findings. First, the lower variance in the prices of "everyday low price" (EDLP) firms serves to reduce the market-level dispersion in prices when such firms are present. Moreover, the price variance of non-EDLP firms in these markets is also lower than in those markets in which EDLP competitors are absent. Second, we also find that dispersion in offered prices increases closer to the departure date, which is consistent with theoretical assertion that price dispersion increases with reservation prices. Finally, we continue to observe dispersion of online prices *even* after accounting for vendor strategy and other known sources of dispersion, suggesting that the prices are unlikely to converge even in the presence of sophisticated online search mechanisms.

Key words: online markets; price dispersion; airline industry; EDLP; hierarchical linear modeling
History: Vallabh Sambamurthy, Senior Editor; Ram Gopal, Associate Editor. This paper was received on June 19, 2007, and was with the authors 15 months for 4 revisions. Published online in *Articles in Advance* March 1, 2010.

1. Introduction

Price dispersion studies can provide rich insights into the prevailing market conditions and are of particular relevance to vendors following market segmentation and price discrimination strategies. Although the price dispersion phenomenon is fundamental to any competitive market, the advent of the Internet and the emergence of online markets for many goods and services have led to renewed interest in this topic. The popular view is that the Internet lowers consumer search costs and alleviates a host of information asymmetry issues, leading to near convergence in prices (Bakos 1997). However, empirical studies show that markets, electronic or otherwise, demonstrate significant dispersion in prices even after taking into account many other known sources of price dispersion (see Pan et al. 2004 for a compressive discussion). An important goal of our current research is to add to the understanding of price dispersion in both online and offline markets. Our work goes beyond the consumer- and/or product-related explanation of price dispersion offered by past empirical research, and examines the key role played by the vendor's price-format strategies.

Literature in marketing suggests that the proactive choice of a particular "price format" can position a firm to attract certain types of consumers in the marketplace (Bell and Lattin 1998). Price formats are not necessarily conveyed by the prices of *individual* goods, but rather reflect the overall pricing strategy

of a firm. Thus, firms may convey a certain image about themselves through price structures alone. For example, it is believed that consumers attach a certain price image to prices ending in 9 (e.g., 44^{49}, 99^{99}) and several retail outlets position themselves as "value priced" by only stocking products priced lower than $0.99. Other classic price formats that firms can adopt are "everyday low price" (EDLP) and "promotional pricing," or HILO/PROMO, strategies. Whereas the former strategy promises lower average prices, the HILO/PROMO strategy uses deep price cuts on certain featured products to attract consumers. Note that simply practicing a certain price format may not always successfully lead to the creation of a corresponding price image in the minds of the consumer—that depends on the consumers' ability to see and compare other prices as well. However, such vendor strategies are quite likely to influence the dispersion of prices *posted* in a market. Moreover, if successful, these pricing strategies are also likely to affect the dispersion of *transacted* prices that eventually clear in the market. However, extant research has not examined the impact of such firm pricing strategies on the dispersion of either posted or transacted prices.

Hence, an important objective of our current research is to empirically examine whether and how vendors' choice of price formats contributes to price dispersion in a market. We are particularly interested in investigating what happens to price dispersion when two or more firms with different price-format strategies compete in the same market. We also seek to empirically test two other questions: First, is the extent of dispersion in transacted and posted prices the same, and if not, do price formats explain these differences? Second, how do consumers' reservation prices affect the extent to which prices are dispersed? Answers to both these questions can shed light on prior theoretical assertions on price dispersion.

To answer these questions, we study the U.S. domestic airline market. The airline market provides an ideal setting for this research not only because the same firms participate in both online and offline channels, but because at least two of these firms are self-declared practitioners of the EDLP strategy, whereas most of the other firms adopt a non-EDLP strategy. The U.S. airline industry has been extensively studied by researchers in economics, and to some extent by those in IS and marketing as well, and therefore, our empirical analysis carefully controls for the well-known determinants of price dispersion (e.g., market factors and ticket characteristics) identified by previous research.

We have two data sets that cover the same three-month period corresponding to the third quarter of 2004. The first data set has a total of 1,137,500 individual tickets posted by 14 network carriers and three regional airlines that together account for over 86% of all domestic passenger enplanements in the United States and is a near-comprehensive list of all *online* airfares *posted* by *both* online travel agents as well as individual airline websites. We put together this data set by developing intelligent Web agents scripted in open source languages that routinely pulled data for all origin-destination pairs in the United States The second data set was obtained from the Origin and Destination Survey that is managed by the U.S. Bureau of Transportation Statistics. This data set is a 10% sample of all airline tickets *transacted* in the United States, which includes *both online* and *offline* modes of purchase. Thus, whereas the former data contain *posted* prices, the latter contain *transacted* prices and is the standard data set for empirical airline research in economics. After narrowing down the specific markets and ticket types for analyses, we have between 200,000 and 250,000 individual tickets for *each* data set.

We develop and estimate a hierarchical linear model that relates price dispersion to its underlying determinants while accounting for the partial dependence among prices due to market- or airline-specific factors. The remainder of the paper is organized as follows. Section 2 reviews relevant literatures on price dispersion and pricing in the airline industry and develops a set of testable hypotheses. Section 3 describes the data and model, and presents the results of our empirical analyses. The paper concludes with a discussion of the results and the limitations of the study, as well as avenues for future research.

2. Sources of Price Dispersion in the U.S. Airline Industry

Three streams of research inform this work. First, a rich history of both theoretical and empirical research in economics has shown how various product, consumer, and market factors affect equilibrium prices and price dispersion in a market. Second, a significant body of research has identified the important factors influencing price dispersion in the airline industry. Finally, a relatively new and emerging stream of information systems (IS) research has studied the nature of online price dispersion. Indeed, recent IS research suggests that product, retailer, and market characteristics should be considered jointly to understand market-level dispersion of prices (Venkatesan et al. 2006). Our empirical models of price dispersion not only weave together findings from each stream, but contribute to the extant literature by examining the impact of vendors' price-format strategies.

The U.S. airline industry comprises 14 major domestic network airlines that operate from 79 cities and compete in over 2,000 routes or markets.[1] The existing literature has studied various elements of this industry, including pricing at the customer end and operational aspects, such as hub-and-spoke management at the back end (Berry et al. 1997, Borenstein 1989, Brueckner et al. 1992). Furthermore, there has also been a steady stream of work that has specifically examined the various sources of heterogeneity in airline pricing (e.g., Borenstein and Rose 1994). These streams of work are briefly reviewed below and are also incorporated into our empirical analysis.

2.1. Ticket Category

In his seminal paper, Varian (1980, p. 651) observes that " ...the 'law of one price' is no law at all." This observation implies that even the prices of relatively homogeneous products are unlikely to converge in a market with heterogeneous consumers. Moreover, product homogeneity is itself a theoretical concept that may be impossible to observe in real-world markets, such as in airline travel. Indeed, products are differentiated in many ways based upon differences in consumers' tastes and preferences (Hotelling 1929, Salop 1979), differences in quality (Mussa and Rosen 1978), or differences in the services that encompass the core product (Chellappa and Kumar 2005), such as shopping experience, convenience, delivery, and return policies (Smith et al. 2000). Thus, the features and prices of two airline tickets may be very different even if they are both issued for the same route and on the same day. One important source of differentiation between two airline tickets for the same route is whether or not there are Saturday-night stay requirements (Clemons et al. 2002, Dana 1998). Dana (1998) and Gale and Holmes (1993) have suggested that air carriers impose Saturday-night stay-over ticket restrictions to discriminate between business and leisure consumers, where business consumers are less elastic with respect to days of travel. Therefore, before we examine the impact of price-format strategies on dispersion, it is important to account for dispersion due to ticket types. Although one way to achieve this is to simply include a fixed effect for business/leisure tickets over all tickets, another approach is to conduct

our empirical analysis separately on each subsample. An advantage of the latter approach is that it directly identifies the impact of different factors on price dispersion in each ticket category without relying on complicated interaction effects. For example, if we account for the same set of variables for both ticket categories, and if the intercept is different in the two cases, then we know that the baseline dispersion itself is different. Because the two ticket categories are targeted towards two different segments, and because airlines price discriminate between the two categories, we should expect that the resulting distribution of ticket prices available in each category is different.

HYPOTHESIS 1 (H1). *Business (tickets with no weekend-stay restriction) and leisure (tickets with weekend-stay restrictions) tickets exhibit different levels of price dispersion.*

However, Stigler (1961, p. 214) observes, " ...a portion of the observed dispersion is presumably attributable to such differences. But it would be metaphysical, and fruitless, to assert that all dispersion is due to heterogeneity." Hence, for each ticket category, we further account for other known and unknown factors that may influence price dispersion.

2.2. Market Characteristics

In the airline industry, a market refers to the route or origin-destination pair of airports whose intrinsic characteristics will likely impact an individual firm's pricing, and hence overall price dispersion in the market. For example, firms might vary in how they price based on the distance between the origin and destination, thus leading to price dispersion. Similarly, prior research suggests that other route-related factors such as the number of slots available at an airport (a form of capacity constraint) and the number of airlines competing on a route also impact prices and, therefore, price dispersion (Berry 1990, Borenstein 1989, Borenstein and Rose 1994, Fournier and Zuehlke 2004, Morrison and Winston 1990, Neels 2000). We include three important variables suggested by extant research to account for price dispersion due to differences in market concentration, slot capacity, and distance.

Market concentration is a measure of the degree of competition on a route, and has been operationalized in two different ways. Borenstein and Rose (1994) use the concept of "density," measured by the total number of flights on the observed route, and flight Herfindahl to capture market concentration. Hayes and Ross (1998) and Stavins (2001), on the other hand, use the number of carriers offering services in each

---

[1] Note that in the airline industry, each route or unique origin-destination pair is referred to as a *market*; i.e., the route San Francisco–Los Angeles would be considered a market. However, following Borenstein (1989), Brueckner (1992), Borenstein and Rose (1994), and Berry et al. (1997), we adopt an even more rigorous definition of market by considering the directional component of the route as well (e.g., LAX-SFO is not the same as SFO-LAX), where the directional distinction is related to where the passenger originates in the outbound itinerary.

route as an alternative measure of market competition. We adopt this latter measure because it is appropriate to the models[2] that we consider. Stavins (2001, p. 202) finds that "price discrimination is higher on routes with more competition," implying that higher market concentration leads to a greater variety of prices. On the other hand, it is also possible that due to multimarket effects and tacit understanding, firms may tend not to undercut each other in very competitive conditions, which suggests that prices will converge when market concentration is high. Although there is evidence that mean prices are higher when market concentration is high (Fournier and Zuehlke 2004), the latter's impact on price dispersion can be ambiguous (Borenstein and Rose 1994). Because market concentration varies across routes, it is important to control for the influence of this variable on price dispersion.

The variable *slot* measures the extent to which one or both of the endpoint airports are congested. Currently there are four slot-constrained airports in the domestic air transportation market in the United States: Chicago O'Hare (ORD), Kennedy (JFK), and La Guardia (LGA) in New York City, and Ronald Reagan Washington National (DCA). The opportunity cost of operating in slot-constrained airports is higher, and hence higher prices are expected if one or both endpoints of a route are slot constrained. Prior studies find mixed results on the relationship between slot and price; whereas Berry et al. (1997) and Fournier and Zuehlke (2004) find a positive relationship, Stavins (2001) finds that the effect of slot on price is negative after controlling for Saturday-night stay-over and advance-purchase requirements. Although the impact of slot constraints on price dispersion has not been directly examined in previous work, Escobari and Gan (2007) suggest that costly capacity constraints will reduce price dispersion, which implies that slot-constrained routes should have lower price dispersion.

Finally, the variable costs of operating a flight are directly associated with flight distance between the two endpoint airports; hence, higher prices are expected for flights with longer distance (Borenstein 1989). On the other hand, the total distance of a flight varies by intermediate point (Hayes and Ross 1998). As the distance covered by a flight increases, the carrier can choose from a greater number of intermediate

airports and potentially generate greater economies of scale. Empirical findings from prior research suggest that cost is the dominant factor that explains the relationship between distance and price (Berry et al. 1997, Borenstein 1989, Stavins 2001). Thus, if the dominant effect of distance is in the form of cost or capacity, then once again we should observe low price dispersion (Escobari and Gan 2007); on the other hand, if longer distances provide greater possibilities for the firms, then we should observe greater price discrimination and hence greater dispersion.

HYPOTHESIS 2 (H2). *Characteristics of a market (route), such as its market concentration, slot capacity, and distance, are important sources of price dispersion.*

2.3. Firm Differences and Market Power
Price differences are often attributed to the relative market power of sellers. In many physical goods industries this is manifested in brand image, where consumers are known to pay higher prices at branded retailers (Brynjolfsson and Smith 2000). In the airline industry, differences in market power can arise from differences in route structures. For example, most airlines are based on the hub-and-spoke system in which the hubs are central to how planes are maintained and flight crews organized (Berry et al. 1997, Borenstein 1989, Brueckner et al. 1992). Airlines often prefer to route their flights through hubs because of the increased flexibility that comes from centralized operations at particular airports. An airline can exercise greater market power over flights that originate/end at a hub because their control over airport resources helps them offer superior services, such as more convenient gates and better departure times (Borenstein 1989), which may be valuable to business and other less price-sensitive travelers. In addition, hubs have the potential to increase entry barriers and drive up prices for hub-originating passengers, and have been used as a price mark-up proxy in previous research (e.g., Hayes and Ross 1998). Although cost efficiencies of hubs can potentially lead to lower prices, prior research suggests that firms typically exercise their market power and charge higher prices at hub airports (Borenstein 1989). A follow-up to the Borenstein and Rose's (1994) work extends this impact to study price dispersion and finds that "an increase in a firm's market power is expected to increase its ability to segment the market," hence, we could expect to see greater price dispersion in such markets (Gerardi and Shapiro 2007).

Furthermore, greater flight frequency has also been shown to lower frequency delay or the difference between a consumer's preferred time and actual time of a flight (Douglas et al. 1974), and is therefore considered to increase the value of services of an airline to a consumer. This is presumably because consumers

---

[2] Total number of flights suffers from collinearity problems with flight frequency; thus, it is not adopted in our analysis. Furthermore, the number of carriers serving a route is preferable because it highly (and negatively) correlates with three Herfindahl measures—flight Herfindahl (−0.69), origination Herfindahl (−0.73), and route Herfindahl (−0.62)—while providing better fit to the models.

who value their time more highly (and hence who are willing to pay higher fares) are likely to find the greater number of options provided by an airline with a higher frequency on a given route more attractive than airlines with lower flight frequencies (Borenstein 1989). Thus, frequency is considered a measure of market power and should be accounted for in any firm-level analysis of price dispersion.

HYPOTHESIS 3 (H3). *Greater market power provides increased ability to price discriminate, resulting in greater price dispersion.*

2.4. Advance Purchase Periods

Reservation price is simply the maximum price a consumer is willing to pay for a good. It is believed that consumers who purchase tickets closer to departure date are those with higher reservation prices for the trip. Although these consumer types may not be able to commit to a trip too far in advance, their opportunity costs (of not being able to make the trip) increase as the departure date approaches. Airlines utilize this "valuation" of time to discriminate between consumer types, and such segmentation serves to weed out consumers with high valuation of time (Stavins 2001). Hence, ticket prices are higher closer to the departure date, because most yield management algorithms (which also take into account seat availability) are programmed to take advantage of this difference in consumers' reservation prices;

> The common mechanism used to segment customers in yield-management situations is the time of purchase; that is, the less price-sensitive customer generally waits until the last minute to make reservations. On the other hand, people who make their reservations early are generally more price sensitive; they are willing to trade away some flexibility for a reduced price.
> (Weatherford and Bodily 1992, p. 832).

Although it is intuitive to see that prices will be higher when consumers' reservation values are higher, we are interested in the nature of price *dispersion* within each ticket segment, i.e., are 7-day advance purchase ticket prices more or less dispersed than 21-day ticket prices? In other words, although 7-day advance-purchase tickets may always be more expensive than 21-day advance-purchase tickets, it is not necessary that all prices for 7-day advance-purchase tickets be the same. For example, if all 7-day purchase tickets in a market were priced at $700 and all 21-day purchases were $300, then there would no variance in prices *within* an advance-purchase period. On the other hand, it is also possible that prices for 7-day tickets are more dispersed than 21-day ticket prices. Furthermore, the nature of the dispersion within a

particular advance-purchase period may be different for business and leisure tickets. Our goal is to account for such within-category price dispersion as well as to study the total dispersion in the market.

Varian's (1980) model of sales suggests that price dispersion will be greater in markets with higher reservation prices. The intuition behind this prediction is that firms will randomize their pricing strategy as long as there are consumers in the market who are not fully knowledgeable about the lowest prices available. Such a randomized strategy will therefore be bounded by the maximum willingness-to-pay of the consumers in the market. Because this upper bound is likely to be higher when consumers have high reservation prices, the range of prices offered in such markets is likely to be high as well. Recently, Baye et al. (2006) empirically examined this proposition in online markets for electronic products. Our goal is to examine whether this theoretical observation finds any support in online airline markets, where the presence of online travel agents that allow for easier search is a countervailing force that can potentially reduce the number of uninformed consumers.

HYPOTHESIS 4 (H4). *Price dispersion increases closer to flight departure.*

2.5. Online Price Dispersion

Information asymmetry, i.e., consumers' inability to fully know and compare prices, is a well-known source of price dispersion. For example, Varian (1980) suggests that price dispersion will persist when sellers intentionally vary prices over time so that consumers cannot learn about them (as long as there is some cost of acquiring price information). Salop and Stiglitz (1977) also note that the presence of at least some uninformed consumers ensures the existence price dispersion. Although maintaining information asymmetry was perhaps easier in pre-Internet days, reduced search costs associated with online markets may make it difficult, perhaps even impossible, for price dispersion to exist. Bakos (1997) argues that lowered search costs on the Internet would decrease the ability of sellers to charge monopolistic prices. However, subsequent empirical work consistently finds prices on the Internet to be dispersed, whether in the market for books, CDs (Brynjolfsson and Smith 2000), tickets from online travel agents (Clemons et al. 2002), or online brokerages (Chen and Hitt 2002). These empirical findings suggest that online markets are not necessarily frictionless, and that all consumers are not perfectly aware of product and pricing characteristics. Hence, even after accounting for extant sources of dispersion, we should find the persistence of price dispersion in online markets as well.

HYPOTHESIS 5 (H5). *We shall continue to observe dispersion of online prices even after accounting for extant sources of dispersion.*

2.6. Price Formats and Price Dispersion

Most of the previously described theoretical models discuss pricing strategies for single-product firms that set prices based purely on differences in consumers' price knowledge and on competitors' actions. However, in reality, most firms sell a variety of products, and do not price them with a myopic view of achieving a one-time sale, or necessarily react to their competitor's prices in a Bertrand-competition fashion. For example, retailers sometimes use products as loss leaders to attract consumers to buy their other products (Hess and Gerstner 1987). One pricing strategy that encompasses a firms' need to compete in multiple product markets on a long-term basis and represent a firm's overall approach to pricing is the creation and maintenance of a "price image" through price formats.

Previous research in marketing suggests that firms follow two basic price formats, namely, "everyday low price" (EDLP) and "promotional pricing" (PROMO or HILO). Although these formats are more appropriately regarded as a continuum rather than a dichotomy (Bell and Lattin 1998, Hoch et al. 1994, Shankar and Bolton 2004), it is commonly agreed that EDLP sellers tend to charge relatively stable, below-average prices with little or no temporary price discounts (Bell and Lattin 1998). These sellers aim to credibly convey to consumers that they consistently offer low prices; hence, sustaining this price image relies on both the magnitude and consistency of prices. On the other hand, HILO/PROMO sellers are promotion oriented. Their prices are normally higher than the market average, but are frequently accompanied by promotions that permit prices on some products to be temporarily lower than EDLP prices (Lal and Rao 1997). Typically, HILO/PROMO firms maintain this price image by offering promotional prices on a small subset of featured items for a limited period of time.

Because these complex price formats can make it difficult for consumers to become informed about the true nature of prices, firms can target consumers based on their preferences and price knowledge (Blattberg et al. 1981). Although segmenting consumers based on their opportunity cost of time (Lal and Rao 1997, Ortmeyer et al. 1991) or their basket size (Bell and Lattin 1998) are well-known short-term pricing tactics, price-knowledge based segmentation is perhaps most important in repeat-purchase product categories. Indeed, these strategies are commonly observed amongst retailers and grocery stores, in which price format significantly influences consumer decisions of where to shop and what to buy. In

the physical retail market, Wal-Mart is the most well-known practitioner of EDLP; similarly, in the U.S. domestic airline market, Southwest and JetBlue are known to have everyday low prices, whereas other airlines offer various forms of price-discounts and promotions consistent with the HILO strategy.

Vendors' conscious choice of particular pricing/segmentation strategies is likely to impact the distribution of prices in the market over and above other relevant factors. Research on Internet car retailing finds that prices obtained through Autobytel, a firm that targets consumers with high costs of information gathering and bargaining, are both lower and exhibit smaller variance compared to those obtained through their competitors (Morton et al. 2001, Zettelmeyer et al. 2001). Furthermore, prior studies based on offline retail prices have documented significant differences in the range and variability in prices posted by sellers adopting different price formats (Ho et al. 1998, Shankar and Bolton 2004). Therefore, the distribution of ticket prices should display three distinct characteristics. First, the dispersion of prices for EDLP vendors should differ from that of non-EDLP vendors. Second, price dispersion in markets where all competitors have a single price format will be different from those markets in which both strategies coexist. Third, the relatively stable prices posted by EDLP competitors increases the price knowledge of consumers, and restricts the extent to which non-EDLP firms in the same markets can price discriminate among consumers.

HYPOTHESIS 6A (H6A). *The adoption of a particular price format by an airline is an important source of price dispersion. In particular, markets with the presence of EDLP carriers will exhibit lower price dispersion compared to markets with no EDLP carriers.*

HYPOTHESIS 6B (H6B). *The prices of non-EDLP carriers in markets with EDLP competitors will be less dispersed relative to their prices in markets without EDLP carriers.*

3. Data and Methodology

Our research is based on two data sets that contain price and detailed ticket information. The first contains prices and descriptions of airline tickets obtained from online travel agents as well as individual airline websites during the third quarter of 2004. This raw data was gathered by a Web-based spider that we developed using Curl (a tool for scouring and extracting information from the Web) and later processed using a parser that we wrote using Perl and other database scripting languages.

We consider homogeneous categories of tickets commonly used by prior research on airline pricing in economics, namely, coach-class, nonrefundable, and

round-trip tickets. The prices cover one-to four-week advance-purchase tickets for both leisure tickets, in which a Saturday-night stay is required, and for business tickets, with no weekend restriction. The data pertain to the 500 busiest routes in the United States, resulting in a total of 1,137,500 individual tickets written by 14 network carriers and three regional airlines that together account for over 86% of all domestic passenger enplanements in the United States. To control for any possible price difference that may be attributed to differences in flight duration, number of connections for any given route, or departure and return schedules, only direct flights are considered in our analyses. Restricting our analysis to comparable tickets avoids the possibility of spurious effects being associated with the pricing strategy variable. Our final data set consists of 251,547 unique observations.[3] The second data set is based on the Origin and Destination Survey (DB1B) provided by the U.S. Bureau of Transportation Statistics (BTS). The data are a sample of 10% of all airline tickets that originated in the United States on domestic carriers for the same quarter, allowing us to compare online and offline prices of the exact same carriers during the same time period as our online data collection. The total number of observations for a nonstop, round-trip ticket in the second data set is 209,273. Appendix A summarizes how the variables included in the models were operationalized.

In the following section, we shall first discuss the three models that are used in this study; we then discuss the various measures of the dependent variables, and subsequently the details of our estimation method and results.

3.1. Models

Airline pricing data exhibit a three-level structure, consisting of tickets (level 1) nested within airlines (level 2) nested within markets (level 3). Due to the hierarchical nature of the data, two critical assumptions of OLS—independence and homoscedasticity of random errors—may not hold. Hence, least-squares regression is not an appropriate analysis technique. The unique cost structure and pricing strategy of each airline may result in its ticket prices in different routes to be correlated. This violation of the OLS assumption of independent errors, also known as Intra-Class Correlation (ICC), can cause the standard errors of the coefficients to be underestimated

---

[3] Tickets written by the three regional airlines (Sun Country, USA 3000, and Midwest Express) in our original data set have been excluded from the tables and subsequent analyses due to insufficient number of observations.

(Kreft and De Leeuw 1998), thus raising the risk of type-I error (Pedhazur 1997). Furthermore, because unit-level random error varies across airlines, the assumption of homoscedasticity is also likely to be violated.

To overcome these problems, we analyze the data using the hierarchical linear modeling (HLM) approach. Hierarchical linear models extend traditional regression models by accounting for the partial dependence of individual tickets within the same airline, and also for tickets in one route being more similar than those belonging to another route. This approach has been recommended for the analysis of airline data by Borenstein and Rose (1994), as well as in a recent study of price dispersion by Venkatesan et al. (2006), to simultaneously account for the correlation induced by retailer characteristics and product similarities. Other researchers in IS (Ang et al. 2002) have also used HLM when examining nested data such as when separating the effects of individual- and institutional-level predictors of compensation. In describing our model levels, we follow the approaches suggested by these extant works.

In the subsequent discussions, we use subscript m to denote a market and subscript k to denote a carrier. Model 1 addresses market-level price dispersion (dependent variable: $Disp_m$). We apply Model 1 separately on the full set of data (all carriers) and a subset of data that includes only non-EDLP carriers to test Hypotheses 6A and 6B, respectively. Model 2 is a firm-level analysis of the variance in prices of individual carriers within each market (dependent variable: $Pvar_{km}$). The goal of Model 2 is to specifically tease out the differences in the price variation of EDLP and non-EDLP carriers. Note that the data in Model 1 are aggregated across all tickets and carriers within each route, reducing the measures to level-3 units. Because the hierarchical structure no longer exists in the market-level model, HLM is applied only for Model 2.

Model 1—Market-Level Price Dispersion

$$Disp_m = \alpha + \beta_1 DD7 + \beta_2 DD14 + \beta_3 DD21$$
$$+ \beta_4 shorthaul_m + \beta_5 slot_m + \beta_6 mktcon_m$$
$$+ \beta_7 EDLPmkt_m + \varepsilon_m, \qquad (1)$$

where

$$\varepsilon_m \sim N(0, \sigma^2). \qquad (2)$$

The dispersion measures in Model 1 are constructed for each market, m. Estimation is carried out separately for both measures of dispersion, i.e., range and coefficient of variation. The independent variable *EDLPmkt* is introduced to identify markets where

at least one of the EDLP carriers operates; other independent variables are defined in Appendix A. Because the unit of analysis is the route (i.e., one observation per route per week), level-2 control/explanatory variables (e.g., airline fixed effect) are excluded from this model.

Model 2—Firm-Level Variance in Prices

Level 1 (Firm-Level) Model:

$$Pvar_{km} = \beta_{0m} + \beta_{1m} DD7 + \beta_{2m} DD14 + \beta_{3m} DD21$$
$$+ \beta_{4m} freq_{km} + \beta_{5m} hub_{km} + u_{k0} + \varepsilon_{km} \qquad (3)$$
$$\varepsilon_{km} \sim N(0, \sigma^2).$$

Level 2 (Market-Level) Model:

$$\beta_{0m} = \gamma_{00} + \gamma_{01} shorthaul_m + \gamma_{02} slot_m$$
$$+ \gamma_{03} mktcon_m + u_{0m}. \qquad (4)$$
$$\beta_{1-5,m} = \gamma_{1-5,m}.$$

Model in the combined form:

$$Pvar_{km} = \alpha + \gamma_{01} shorthaul_m + \gamma_{02} slot_m + \gamma_{03} mktcon_m$$
$$+ \gamma_{10} DD7 + \gamma_{20} DD14 + \gamma_{30} DD21 + \gamma_{40} freq_{km}$$
$$+ \gamma_{50} hub_{km} + \varepsilon_{km}, \qquad (5)$$

where

$$\alpha = \gamma_{00} + u_{k0} + u_{0m},$$
$$u_{0m} \sim N(0, \varphi). \qquad (6)$$

Consistent with Borenstein and Rose (1994), route effects (u_{0m}) are treated as random, whereas airline effects (u_{k0}) are considered fixed. Similar to Clay et al. (2002), we infer the impact of the particular strategy followed by the EDLP carriers, Southwest and JetBlue, using the fixed effects for each airline. The dispersion measures in Model 2 are constructed for each unique firm-market pair, km. Note that because level-1 units (ticket-level) are aggregated in computing the dependent measures, the model exhibits a two-level (rather than three-level) hierarchy. Again, estimation is carried out separately for both measures of dispersion, i.e., range and coefficient of variation (calculated at the carrier-route level). Frequency in this model is defined as the weekly average number of actual flight departures *per airline per route*.

3.2. Measuring Price Dispersion

Our dependent variable for Model 1 is $DISP_m$, i.e., dispersion of prices in a market. There are a number of acceptable ways to measure price dispersion,

although the most commonly used ones are the price range and the coefficient of variation (CV).[4] These two measures have been widely used in prior research on price dispersion (Baye and Morgan 2004; Baye et al. 2004, 2006; Carlson and Pescatrice 1980; Sorensen 2000). Although raw measures of variability such as variance and standard deviation have also been analyzed (Dahlby and West 1986, Pratt et al. 1979), the scaled measure (CV) is superior because it can distinguish between two markets with the same variance in prices. In this situation, the CV will accurately reflect the higher price dispersion in the market with a lower average price relative to one in which the average price is higher. For the same reasons, we use CV to measure firm-level price variance as well. Table 1 shows how the different firm- and market-level dispersion measures were operationalized in this study.

3.3. Estimation and Results

Two maximum-likelihood methods are commonly used in estimating hierarchical linear models: the full maximum likelihood (ML) and the restricted maximum likelihood (REML). In ML, both fixed effects and variance components are included in the likelihood function. Variance-covariance parameters and second-level fixed coefficients are estimated by maximizing the joint likelihood. In REML, variance-covariance components are first estimated with maximum likelihood that integrates over all possible values of the fixed effects, which are then recovered using generalized least square (GLS) given the variance-covariance estimates obtained from the first step (Goldstein 1995, Raudenbush and Bryk 2002, Raudenbush et al. 2001). REML minimizes the deviance of the least-squares residuals as opposed to minimizing deviance of the data.

We adopt REML to estimate our models because ML, although consistent and asymptotically efficient, does not adjust for the number of fixed effects that are being estimated. As a result, the variance components will tend to be underestimated with small sample sizes or when the number of groups is small (Jones and Steenbergen 1997). Although the source data sets used in this research do not fall into the "small sample size" category, our dependent variables are aggregated measures (i.e., range and coefficient of variation) that significantly reduce the number of observations in some of our analyses. Furthermore, the use of REML is also consistent with a majority of previous research that has used the HLM approach.

---

[4] In an empirical analysis of price dispersion in the airline industry, Borenstein and Rose (1994) report highly similar results based on dispersion measured with the Gini coefficient and the coefficient of variation.

Table 1 **Dispersion Measures**

Unit of analysis[a]	Measure	Explanation
Firm-level variance in prices ($Pvar_{km}$)	$range_{km} = \dfrac{\max price_{km} - \min price_{km}}{\max price_m - \min price_m}$	Standardized difference between the maximum and minimum prices available from carrier k in market (route) m at period t.
	$CV_{km} = \dfrac{\sqrt{1/(I_{km} - 1)\sum_i^{I_{km}} (price_{ikm} - \overline{price_{km}})^2}}{\overline{price_{km}}}$, where $\overline{price_{km}} = \dfrac{1}{I_{km}}\sum_i^{I_{km}} price_{ikm}$	Coefficient of variation of prices of a set of tickets I_{km} offered by carrier k in market m at period t.
Market-level price dispersion ($Disp_m$)	$range_m = \max price_m - \min price_m$	Raw range of prices of all tickets in a given in market m at period t.
	$CV_m = \dfrac{\sqrt{1/(I_m - 1)\sum_i^{I_m} (price_{im} - \overline{price_m})^2}}{\overline{price_m}}$, where $\overline{price_m} = \dfrac{1}{I_m}\sum_i^{I_m} price_{im}$	Coefficient of variation of prices of a set of tickets I_m offered by *all* carriers in market m at period t.

[a] All dependent variables are measured at each time period t. For simplicity in exposition, the subscript has been suppressed from all tables and models.

In Tables 2 and 3, we present the summary statistics for ticket prices at the individual airline level for online (posted) prices and transacted (DB1B) prices, respectively.

First note that both the range (Max–Min) and mean of posted online prices (pooled across advance-purchase periods) are lower than their transacted counterpart. The transacted prices (DB1B data) are always pooled because there is no information available on advance-purchase characteristics. Table 3 reveals that correlations among the variables are low, ruling out concerns about multicollinearity.

Table 4 presents the main results from our analysis that examines market-level price dispersion for online prices as given by Model 1 (Table 7 does the same with DB1B data). We can see a significant relationship between advance-purchase requirements and price dispersion, with the exception of the 21-day advance-purchase period for leisure tickets. Observe that the magnitudes of coefficients decrease with the number of days of advance purchase (i.e., coefficient of *DD*7 is greater than that of *DD*14, whereas the coefficient of *DD*14 is greater than that of *DD*21[5]). This finding suggests that compared to the base advance-purchase period of 28 days, ticket prices become more and more dispersed as the departure date approaches. Hence, Hypothesis 4 is supported. Consistent with the fact that consumers with higher opportunity costs (those who cannot always plan earlier) have higher reservation prices, we know that reservation prices are increasing towards flight departure as preplanning becomes less constrained. Therefore, whereas *DD*7 is representative of the market with the highest reservation prices, *DD*28 (base

dummy) is representative of the market with the lowest reservation prices. Furthermore, we conducted pairwise *t*-tests and confirmed that the intercepts and coefficients are statistically different for business and leisure tickets, thus providing support for Hypotheses 1.

Note that both *slot* (a measure of resource scarcity) and *mktcon* (a measure of intensity of competition) are positively correlated with price dispersion for both types of tickets. On the other hand, ticket prices for routes shorter than 500 miles (*shorthaul*) are less dispersed than the market average in *range*, even if the same cannot be conclusively said for the CV measure. Overall, these results provide general support for Hypothesis 2, which suggests that market characteristics are a source of price dispersion. Of particular interest to this research is the variable *EDLPmkt* that defines whether a market is served by an EDLP carrier. The significant and negative sign indicates that whenever a market is served by one or more carriers that practice the EDLP price format, the resulting prices in this market are less dispersed than others in which only non-EDLP carriers operate. This provides strong support for Hypotheses 6A, which suggests that markets with EDLP carriers should exhibit lower price dispersion compared to other markets where EDLP carriers are absent.

In addition to examining whether vendor pricing strategies impact the overall dispersion of prices in the market, results from Model 2 provide insights into how these strategies impact firm-level price dispersion. We interpret the results by considering a non-EDLP airline as the base and examine the coefficients of the fixed effects for the two EDLP airlines. Table 5 shows that even after controlling for all known market and ticket characteristics, both EDLP

[5] *T*-tests confirm that the coefficients for *DD*7, *DD*14, and *DD*21 are statistically different within each ticket category.

Table 2   Summary Statistics (Online Posted Prices)

	Mean	STD	Min	Max	Correlation matrix								
Price	328.16	265.88	89	2,959	1.00								
DD7	0.25	0.43	0	1	0.17	1.00							
DD14	0.25	0.43	0	1	−0.02	−0.33	1.00						
DD21	0.25	0.44	0	1	−0.08	−0.33	−0.34	1.00					
Freq	43.34	39.54	0	192	−0.15	0.00	−0.01	0.00	1.00				
Hub	0.74	0.44	0	1	0.05	0.00	0.00	0.00	0.3	1.00			
Shorthaul	0.37	0.48	0	1	−0.27	0.01	−0.01	0.01	0.13	−0.07	1.00		
Slot	0.45	0.5	0	1	−0.02	0.00	0.01	−0.01	−0.17	−0.24	0.00	1.00	
Mktcon	7.44	2.07	1	14	0.23	−0.01	0.01	0.00	−0.34	−0.09	−0.39	−0.1	1.00
EDLPmkt	0.17	0.38	0	1	−0.27	0.01	−0.01	0.00	0.34	−0.12	0.26	−0.27	−0.29	1.00

Table 3   Summary Statistics (Online Plus Offline Transacted Prices—DB1B Data)

	Mean	STD	Min	Max	Correlation matrix						
Price	353.66	245.18	61	3,275	1.00						
Freq	43.34	39.54	0	192	0.05	1.00					
Hub	0.74	0.44	0	1	0.16	0.30	1.00				
Shorthaul	0.37	0.48	0	1	−0.12	0.13	−0.07	1.00			
Slot	0.45	0.5	0	1	−0.07	−0.17	−0.24	0.00	1.00		
Mktcon	7.44	2.07	1	14	0.08	−0.34	−0.09	−0.39	−0.1	1.00	
EDLPmkt	0.17	0.38	0	1	−0.15	0.34	−0.12	0.26	−0.27	−0.29	1.00

Table 4   Market-Level Online Price Dispersion

	Model 1				
	Business			Leisure	
Variables	Range	CV		Range	CV
Intercept	0.1917	0.0299***		29.4895	0.0860***
DD7	124.2662***	0.0592***		147.2428***	0.0626***
DD14	49.5949***	0.0391***		118.8963***	0.0675***
DD21	28.4760*	0.0222**		−0.6808	0.0033
shorthaul	−48.6108***	−0.0105		−70.1794***	−0.0106
slot	43.3150***	0.0355***		33.3705**	0.0366***
mktcon	68.9733***	0.0377***		89.6429***	0.0383***
EDLPmkt	−74.7711***	−0.0194**		−97.1278***	−0.0429***
−2LL	25,569.5	−1,820.4		26,069.4	−1,752.3
BIC	25,577	−1,812.9		26,077	−1,744.8
N	1,864	1,864		1,867	1,867

airlines exhibit lower variance in their price portfolio (with the exception of one measure for JetBlue). Note that the Department of Transportation defines seven airlines to be low-cost carriers based on their operational infrastructure, including the two EDLP practitioners Southwest and JetBlue, as well as others such as America West and AirTran. However, it is important to note that relative to other airlines, only the two EDLP airlines have negative coefficients for price dispersion (whereas the coefficients for other airlines are either insignificant or positive if significant). This is perhaps sufficient evidence that market-level dispersion of prices is a function of front-end pricing strategy rather than operational infrastructure

alone. Although it may indeed be true that the two (back-end operations and front-end pricing) need to be aligned in order for a strategy to be successful, it is beyond the scope of the current paper to empirically examine this relationship.

Also note that in Hypothesis 3, consistent with prior research, we had suggested that market power (freq and hub[6]) is a source of price dispersion. The

---

[6] At the time of our data collection, JetBlue Airways had two hub airports (JFK and LGB), whereas Southwest Airlines had six (ABQ, BNA, DAL, HOU, PHX, STL) in our sample. Of the 500 routes in our sample, 65 routes include an airport that is a hub for either/both of the airlines.

Table 5 Firm-Level Variance in Online Prices of All (Both EDLP and Non-EDLP) Carriers[a]

| | Model 2 | | | |
| | Business | | Leisure | |
Variables	Range	CV	Range	CV
Intercept	0.9598***	0.0408***	0.9915***	0.1344***
*DD*7	0.0602***	0.0648***	0.0039	0.0584***
*DD*14	0.0514***	0.0431***	0.0000	0.0577***
*DD*21	0.0278*	0.0259***	0.0027	−0.0021
Freq	0.0020***	0.0002*	0.0021***	0.0000
Hub	0.0539***	0.0213***	0.0584***	0.0135*
Shorthaul	0.0100	−0.0009	0.0422**	0.0059
Slot	−0.0175	0.0256***	0.0222	0.0330***
Mktcon	−0.1639***	0.0071*	−0.1527***	0.0028
*EDLP*1 (Southwest)	−0.2356***	−0.0724***	−0.2093***	−0.1163***
*EDLP*2 (JetBlue)	−0.1302***	−0.0285	−0.2219***	−0.0867***
−2 *LL*	2,501.4	−4,562.7	1,893.3	−3,613.7
BIC	2,513.7	−4,550.4	1,905.6	−3,601.4
N	3,592	3,592	3,600	3,600

Note. For parsimony, coefficients of other airline fixed effects and variance of market random effects are not reported.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.01$ (same notation used for all tables).

firm-level analysis given by Model 2 supports this result across both ticket types and segments. Note that the positive coefficients imply that the greater the market power of firm, the greater will be the variance in prices it offers.

Another hypothesis (H6B) was that the presence of EDLP airlines may alter the pricing behavior of non-EDLP airlines, thus affecting the price dispersion in the market as a whole. To examine this question, we apply Model 1 to a subset of data that includes only non-EDLP carriers, and analyze the extent to which the dispersion in prices of these airlines varies with the presence or absence of EDLP competitors. The results are presented in Table 6. We can observe that the coefficient for *EDLPmkt* is both negative and sig-

nificant, which suggests that non-EDLP carriers lower the range of their prices whenever they face EDLP competitors (though CV is not affected). This finding provides direct support for Hypothesis 6B, and once again offers support to the intuition that practice of particular vendor pricing strategies is an important source of price dispersion in the marketplace (Hypothesis 6A).

It is also interesting to note that although the presence of EDLP airlines impacts the price range, it does not have the same effect on the CV. Perhaps one explanation for this observation is that price ranges are more subject to search technologies than price variances. For example, through simple sorting a consumer might be able to find the lowest/highest prices in the market, and because EDLP prices are often used as reference prices (Blattberg et al. 1995), non-EDLP carriers may want to appear in search results by providing greater discounts. On the other hand, variance is reflective of overall prices in the market and is less easily processed by the consumer.

We then conducted two similar market-level analyses with the DB1B data obtained from the U.S. Bureau of Transportation Statistics. These data contain both online and offline prices, and results of the corresponding analyses are reported in Tables 7 and 8. Note that in contrast to the online prices *offered* by the airlines, which we analyzed previously, the current analysis relates to prices that were actually paid by consumers. Results in Table 7 parallel those obtained from our previous analysis of posted prices: Transacted prices in markets with EDLP carriers exhibit lower dispersion than in non-EDLP markets. Note that although the *shorthaul* variable continues to correlate negatively with price range, as in Table 4, it does not have a significant impact on CV. This suggests that the range of market-clearing prices is smaller for shorthaul markets than for long-haul ones. Also, similar to the results for online posted prices, competi-

Table 6 Market-Level Dispersion in Online Prices Offered by Non-EDLP Carriers

| | Business | | Leisure | |
Variables	Range	CV	Range	CV
Intercept	1.9915	0.037***	30.5176	0.0955***
*DD*7	141.2426***	0.063***	166.5061***	0.0677***
*DD*14	55.9343***	0.041***	137.5549***	0.0750***
*DD*21	31.4467*	0.023**	−4.3803	−0.0013
Shorthaul	−55.7507***	−0.003	−81.2292***	−0.0085
Slot	46.1849***	0.039***	33.8186**	0.0365***
Mktcon	67.0064***	0.032***	86.1184***	0.0328***
EDLPmkt	−88.9240***	−0.010	−80.7096***	0.0029
−2 *LL*	22,081.4	−1,428	22,561.9	−1,372.8
BIC	22,088.7	−1,420.6	22,569.3	−1,365.5
N	1,593	1,593	1,601	1,601

Table 7 Market-Level Online and Offline (Transacted)[a] Price Dispersion

Variables	Range	CV
Intercept	472.34***	0.3817***
Shorthaul	−175.67***	0.0143
Slot	−34.57	−0.006
Mktcon	79.11***	0.0170***
EDLPmkt	−520.34***	−0.1904***
−2 *LL*	7,329.8	−528.3
BIC	7,335.9	−522.1
N	486	486

[a]Since there is no information on the timing of ticket purchase (number of days prior to departure) in the DB1B data, all advance purchase dummies (*DD*7, *DD*14, *DD*21), and Saturday-night stay-over restrictions (business versus leisure) are excluded from analyses with these data.

Table 8 Market-Level Dispersion in Online and Offline (Transacted) Prices, Non-EDLP Carriers

Variables	Range	CV
Intercept	857.89***	0.5121***
Shorthaul	−441.58***	−0.0050
Slot	−96.82*	−0.0094
Mktcon	152.29***	−0.0104
EDLPmkt	−284.51***	−0.0990***
−2 LL	6,240.8	−389.7
BIC	6,246.8	−383.7
N	486	486

tion (*mktcon*) impacts both dispersion measures. Interestingly, slot has no significant effect on dispersion in the transacted prices, implying that resource constraints have no significant impact on the dispersion of prices that clear the market. Note that although prior research has established the importance of this resource constraint in price dispersion, it has not offered any conclusive evidence of its impact on price dispersion.

The results presented in Table 8 are similar to those reported in Table 6 earlier, except that non-EDLP transacted prices are analyzed instead of posted prices. Consistent with our earlier findings for online posted prices, the negative coefficient of the *EDLPmkt* variable implies that even non-EDLP carriers tighten their price range and lower their CV in the presence of EDLP competitors.

3.4. Comparing Posted and Transacted Prices

Note that the government-maintained DB1B data do not have all the variables that are in our self-acquired data set (posted online prices). The DB1B does not contain advance-purchase period information nor does it explicitly identify the channel (online versus offline) in which the ticket was purchased. Therefore, we remove these identifiers from the posted price data set to create a new data set that is reasonably comparable to the DB1B counterpart. We carried out *t*-tests to examine whether the range and CV of prices for each carrier-route observation in the two data sets were the same. We find that the range of prices of the posted (online only) tickets is smaller than the range of transacted prices (i.e., a mix of online and offline tickets) 80% of the time. Similarly, we find that the CV of the former is smaller than in the latter data set 83% of the time. All findings are statistically significant at the 99.99% confidence level. We also repeated this analysis using random samples consisting of about 25% of the observations in each data set and got similar results.

Although we cannot make conclusive statements on online versus offline dispersion afrom this finding

(because the data sets are not perfectly comparable), the results do provide some indication that offline prices may be less dispersed than online prices. The logic underlying this assertion is as follows. First, note that transacted prices must be a subset of posted prices because only a portion of the posted prices will clear. Because we find transacted prices to have a higher range, variance, and coefficient of variance than the posted prices, and because posted prices come purely from online sources, whereas the transacted prices are a mix of online and offline, this suggests that the additional dispersion in prices is likely to have been introduced by the offline component.

Finally, Hypothesis 5 is supported by our analysis, despite the inclusion of other known explanatory factors, no model of online prices fully accounts for all the dispersion in online prices—i.e., the residuals remain significant. This can also be verified by examining the pseudo *R*-square. The persistence of price dispersion in online markets goes against the notion that online search costs are zero, but lends support to the view that online markets do have nonzero frictional costs (Hann and Terwiesch 2003). Also note that we are the first to simultaneously examine dispersion in both posted and transacted prices *across the same set of firms*. We find dispersion in posted as well as transacted prices. The dispersion in posted prices reflects vendors' price discrimination strategy, whereas the dispersion in transacted prices confirms that these posted prices do clear the market. This suggests that the price discrimination strategies are successful and that not all consumers are informed about all prices. The results from our analyses in this study are summarized in Table 9.

4. Discussion and Conclusions

Research in economics has strived to identify and explain the many sources of price dispersion even when some leading theorists have consistently warned that not all of price dispersion is necessarily explainable. More recently, this topic has been raised again by the expanding scope of electronic markets and the prevailing misconception that online markets will lead to convergence of prices. Our research is motivated by the theoretical interests that surround this problem of price dispersion, as well as the need to empirically explain and reconcile the findings in online marketplaces.

Most of the discussions on online markets generally portray a marketplace dominated by consumers, in which negligible search costs are purported to reduce vendors to compete in the Bertrand equilibrium

Table 9   Summary of Analyses and Results

Hypothesis	Variables	Online data	DB1B data
H1: Different ticket categories (business versus leisure) exhibit different levels of dispersion	Online data set (*posted prices*), segmented by weekend stay-over restriction	Yes (Table 4)	N/A
H2: Market characteristics of a route are a source of price dispersion	Market concentration, slot, distance (*mktcon, slot, shorthaul*)	Yes[+] (Table 4)	Yes[+*] (Table 7)
H3: Market power of the airline is a source of price dispersion	Hub, Frequency (*hub, freq*)	Yes (Table 5)	N/A
H4: Price dispersion increases closer to flight departure	Advance purchase (*DD*7, *DD*14, *DD*21)	Yes[++] (Tables 4 and 6)	N/A
H5: Online price dispersion cannot be fully explained even after accounting for extant sources of dispersion	Online data set (*posted prices*)	Yes (Table 4)	
H6A: Markets in which EDLP carriers are present exhibit lower price dispersion compared to those without EDLP carriers	EDLP markets (*EDLPmkt*)	Yes** (Tables 4 and 6)	Yes (Tables 7 and 8)
H6B: Prices of tickets offered by non-EDLP carrier are less dispersed in markets where EDLP carriers are present compared to those without EDLP carriers			

*Slot not significant, **Except for CV; [+] *shorthaul* not significant for CV; [++] *DD*21 for leisure tickets not significant.

fashion. However, the empirical reality does not match these observations. First, even for the same product, it is possible to observe a wide range of prices in both online and offline markets. Second, rarely do vendors adjust the prices of each and every one of their goods in reaction to competitor pricing and/or consumer search costs. Indeed, both theoretical and empirical research in marketing suggests that conveying and maintaining a certain image to consumers is as much a part of pricing as economic factors such as costs, market characteristics, and competitor strategies—price can be a vehicle to convey a message, and price format is one common way through which it is accomplished. Therefore, if indeed price format is an integral part of a firm's conscious branding strategy, then its influence should be evident in any distribution of prices at the market level. Hence, it is one of our goals in this paper to examine the role of vendors' price-format strategies on price dispersion.

The empirical setting that we analyze in this study is the domestic airline market in the United States We choose the airline industry for a number of reasons: First, it is a well-studied industry that has identified innumerable sources of dispersion in economics. Hence, our demonstration of price-format strategy as a source of dispersion will be sound if this variable remains a significant factor in explaining dispersion of prices in the market *even after* accounting for other sources from prior research. Second, the availability

of online *and* offline prices for the *same* set of firms represent an improvement over previous research in which *different* sets of firms provided online and offline prices. Third, this setting provides us with an opportunity to acquire both posted and transacted prices. Whereas research in economics has generally used transacted prices to study dispersion, IS research has generally used posted prices. Combining the two gives us the opportunity to see if pricing strategies employed (as gleaned from posted prices) by the firms do end up clearing when consumers eventually purchase (as can be understood from transacted prices).

Our results control for and verify the role of factors established by prior research on airline price dispersion. Because the roles of these variables have been elaborated upon earlier, we focus on our new findings here. First, we examine whether price dispersion is significantly different in markets where EDLP airlines operate compared to markets in which only non-EDLP carriers compete. We find that, indeed, dispersion of prices, in terms of both range and CV, are lower in the former. Second, we investigate why these markets have lower dispersions—which could potentially stem from a combination of two occurrences: EDLP firms directly reducing down the dispersion of the prices in the market, or their existence affects the competitive behaviors of other firms so that these firms lower their price variance as well. Hence, we ascertain whether there is a distinct difference in the variance of prices

set by EDLP and non-EDLP firms. Our results show that: (1) The range and CV of prices offered by EDLP firms are indeed lower than those of their non-EDLP counterparts; (2) the range and CV of prices offered by *non*-EDLP firms are lower in markets with EDLP competitors than in markets where these competitors are absent. Hence, our hypothesis that pricing strategy through price formats is a source of price dispersion is supported. An important managerial implication of our findings is that non-EDLP competitors will not shy away from imitating the EDLP firms when faced with this type of competition.

Although theories of online search have suggested price convergence, we know that price dispersion persists even in mature markets (Ratchford et al. 2003). Consistent with these early findings, and more recent work that reaffirms the existence of online price dispersion (Walter et al. 2006), our analysis also finds evidence of price dispersion online.

Prices closer to departure date are higher because they are specifically targeted towards consumers with greater opportunity cost of time. An important finding of our paper is that price dispersion is also increasing closer to departure date. There are two possible reasons: First, higher reservation price implies higher upper bound for prices, and thus the possibility of price offerings is higher. Second, price-insensitive consumers (often corporate travelers) are also perhaps less inclined to search for fares, and hence may be unless informed about the entire distribution of prices in the market. Both these reasons resonate with Varian's (1980) observation that price dispersion increases with reservation prices and the proportion of uninformed consumers.

Finally, we are also in a unique position to compare the dispersion of posted and transacted prices. Our own data and that of DB1B suggest that EDLPs only offer the lowest prices in the market about 50% of the time. In other words, these firms appear to maintain their price image not necessarily by focusing on the "low price" component *alone*, but they also strategically manipulate their "everyday" or consistency component. However, decreased search costs online could potentially lead EDLP airlines to be "found out," i.e., their low price image could take a beating if their prices were always compared to lower promotional prices from a non-EDLP airline. However, in contrast, our research shows that the EDLP prices *do* clear in the marketplace, suggesting that price-format strategies do work even in online markets. This observation has implications for other nonairline vendors such as Wal-Mart that practice EDLP strategy offline. Our findings would suggest that Wal-Mart need not significantly alter its online practices to be competitive with exclusively online retailers like Amazon.com and others. Although further examination of firm-specific category-level analysis is warranted, it may very well

be that firms selectively employ EDLP strategy or perhaps even vary in the degree to which they practice EDLP across product categories.

It is interesting to observe that EDLP airlines do not list themselves with any of the online travel agencies, thus avoiding making themselves vulnerable to easy price comparisons.[7] This means that although firms such as Target.com colist their products on Amazon.com's marketplace and allow purchase through Google's payment system, it may not be wise for Wal-Mart to pursue such a strategy simply to expand their potential market base. Rather, it may be better off by operating independently and minimizing price comparisons.

4.1. Limitations and Future Research

The main limitations of our work stem from the data collected by the government on transacted prices. Whereas we were able to acquire data on posted prices where we could categorize by weeks of advance purchase and weekend stay-over restrictions, this information is not available in the data put forth by the Bureau of Transportation Statistics. Hence, a one-to-one comparison of posted and transacted tickets is not possible. Furthermore, the BTS data do not identify the actual channel of ticket purchase if it was bought through a travel agent or purchased from an airline website. This once again limits some of our comparisons, because for the posted prices (which we collect ourselves) we do know if the ticket prices actually came from an airline's own website or from an online travel agent. We also do not possess any data on frequent fliers and loyalty card holders for the different airlines; clearly, there will be a predisposition of these consumer types towards particular airlines, which may in turn affect pricing strategies. Availability of more detailed data could lead to interesting extensions, including estimation of search costs.

The current paper focuses on understanding market-level dispersion of prices. However, many interesting questions with regard to a firm's strategic practice of its price-format strategy remain unexplored, particularly in the online context. A possible extension to this work is to investigate how firms may strategically vary their practice of price formats in different markets or for different product categories. There is also a limited amount of empirical work that coexamines the front-end pricing strategy and the back-end operational aspects of airline operations. We may have an opportunity to empirically examine whether EDLP airlines possess a particular cost structure, and how this may play a role in an airline's decision on price-format adoption and pricing.

---

[7] We would like to thank one of the reviewers for encouraging us to explore this aspect further.

Appendix A. Description of Variables

Factor	Variable	Related Literature	Explanation
Ticket categories and consumer segments	Business/leisure tickets (weekend stay-over requirement) Number of days of advance purchase (*DD*7, *DD*14, *DD*21)	Clemons et al. (2002), Dana (1998), Gale and Holmes (1993), Stavins (2001)	*DD*7 = 1 if observed ticket is generated within 0–7 days of departure date; 0 otherwise. *DD*14 = 1 if observed ticket is generated within 7–14 days of departure date; 0 otherwise. *DD*21 = 1 if observed ticket is generated within 14–21 days of departure date; 0 otherwise. Baseline comparison is *DD*28 (tickets generated four weeks and beyond from the date of departure).
Market characteristics	Market concentration (*mktcon*)	Hayes and Ross (1998), Stavins (2001)	*Mktcon* is the total number of airlines offering services (nonstop or indirect) in the observed route.
	Slot-constrained airports (*slot*)	Berry et al. (1997), Stavins (2001), Verlinda and Lane (2004)	*Slot* = 1 if either or both origin and destination airports on the observed route are slot controlled; 0 otherwise.
	Distance (*shorthaul*)	Berry et al. (1997), Borenstein (1989), Hayes and Ross (1998), Stavins (2001)	*Shorthaul* = 1 if the nonstop distance between origin and destination airports on the observed route is equal to or less than 500 miles; 0 otherwise.
	EDLP market (*EDLPmkt*)	New variable	*EDLPmkt* = 1 if the market is served by EDLP carriers (Southwest and/or JetBlue).
Market power	Hub (*hub*)	Berry et al. (1997); Borenstein (1989, 1991); Hayes and Ross (1998)	*Hub* = 1 if the origin and/or destination airport(s) in a given route is (are) a hub(s) for an airline offering the ticket; 0 otherwise.
	Frequency (*freq*)	Borenstein and Rose (1994), Hayes and Ross (1998)	*Freq* is the weekly average number of flights scheduled for departure from the origin to destination on a given route by the carrier writing the ticket.
Price-format	*EDLP*	New variable	*EDLP* = 1 if the ticket is written by a carrier adopting the "everyday low price" strategy (Southwest or JetBlue); 0 otherwise.

References

Ang, S., S. Slaughter, K. Y. Ng. 2002. Human capital and institutional determinants of information technology compensation: Modeling multilevel and cross-level interactions. *Management Sci.* **48**(11) 1427–1445.

Bakos, J. Y. 1997. Reducing buyer search costs: Implications for electronic marketplaces. *Management Sci.* **43**(12) 1676–1692.

Baye, M. R., J. Morgan. 2004. Price dispersion in the lab and on the Internet: Theory and evidence. *RAND J. Econom.* **35**(3) 449–466.

Baye, M. R., J. Morgan, P. Scholten. 2004. Temporal price dispersion: Evidence from an online consumer electronics market. *J. Interactive Marketing* **18**(4) 101–115.

Baye, M. R., J. Morgan, P. Scholten. 2006. Persistent price dispersion in online markets. D. W. Jansen, ed. *The New Economy and Beyond: Past, Present and Future.* Edward Elgar Publishing, New York, 122–143.

Bell, D. R., J. M. Lattin. 1998. Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. *Marketing Sci.* **17**(1) 66–88.

Berry, S. 1990. Airport presence as product differentiation. *Amer. Econom. Rev.* **80**(2) 394–399.

Berry, S., M. Carnall, P. T. Spiller. 2006. Airline hubs: Costs, markups and the implications of customer heterogeneity. D. Lee, ed. *Competition Policy and Antitrust Advances in Airline Economics,* Vol. 1. Elsevier, Oxford, UK, 183–214.

Blattberg, R. C., R. Briesch, E. J. Fox. 1995. How promotions work. *Marketing Sci.* **14**(3) G122–G132.

Blattberg, R. C., G. Eppen, J. Liberman. 1981. A theoretical and empirical evaluation of price deals for consumer nondurables. *J. Marketing* **45** 116–129.

Borenstein, S. 1989. Hubs and high fares: Dominance and market power in the U.S. airline industry. *RAND J. Econom.* **20**(3) 344–365.

Borenstein, S., N. L. Rose. 1994. Competition and price dispersion in the U.S. airline industry. *J. Political Econom.* **102**(4) 653–683.

Brueckner, J. K., N. J. Dyer, P. T. Spiller. 1992. Fare determination in airline hub-and-spoke networks. *RAND J. Econom.* **23**(3) 309–333.

Brynjolfsson, E., M. Smith. 2000. Frictionless commerce? A comparison of Internet and conventional retailers. *Management Sci.* **46**(4) 563–585.

Carlson, J., D. Pescatrice. 1980. Persistent price distributions. *J. Econom. Bus.* **33** 21–27.

Chellappa, R. K., R. Kumar. 2005. Examining the role of "free" product—Augmenting online services in pricing and customer retention strategies. *J. Management Inform. Systems* **22**(1) 355–377.

Chen, P., L. M. Hitt. 2002. Measuring switching costs and the determinants of customer retention in Internet-enabled businesses: A study of the online brokerage industry. *Inform. Systems Res.* **13**(3) 255–274.

Clay, K., R. Krishnan, E. Wolff, D. Fernandes. 2002. Retail strategies on the web: Price and non-price competition in the online book industry. *J. Indust. Econom.* **50**(3) 351–367.

Clemons, E. K., I. Hann, L. M. Hitt. 2002. Price dispersion and differentiation in online travel: An empirical investigation. *Management Sci.* **48**(4) 534–549.

Dahlby, B., D. West. 1986. Price dispersion in an automobile insurance market. *J. Political Econom.* **94**(2) 418–438.

Dana, J. D. 1998. Advance-purchase discounts and price discrimination in competitive markets. *J. Political Econom.* **106**(2) 395–422.

Douglas, G. W., G. Warren, J. C. Miller. 1974. Economic regulation of domestic air transport: Theory and policy. Report, Brookings Institution, Washington, DC.

Escobari, D., L. I. Gan. 2007. Price dispersion under costly capacity and demand uncertainty. NBER Working Paper W13075, http://ssrn.com/abstract=986925.

Fournier, G. M., T. Zuehlke. 2004. Price effects of reciprocal multi-market contacts among airline carriers. Working paper, Emory University, Atlanta.

Gale, I. L., T. J. Holmes. 1993. Advance-purchase discounts and monopoly allocation of capacity. *Amer. Econom. Rev.* **83**(1) 135–146.

Gerardi, K., A. H. Shapiro. 2007. The effects of competition on price dispersion in the airline industry: A panel analysis. Working papers, Federal Reserve Bank of Boston, Boston.

Goldstein, H. 1995. *Multilevel Statistical Models*, 2nd ed. Halstead Press, New York.

Hann, I.-H., C. Terwiesch. 2003. Measuring the frictional costs of online transactions: The case of a name-your-own-price channel. *Management Sci.* **49**(11) 1563–1579.

Hayes, K. J., L. B. Ross. 1998. Is airline price dispersion the result of careful planning or competitive forces? *Rev. Indust. Organ.* **13**(5) 523–541.

Hess, J. D., E. Gerstner. 1987. Loss leader pricing and rain check policy. *Marketing Sci.* **6**(4) 358–374.

Ho, T.-H., C. S. Tang, D. R. Bell. 1998. Rational shopping behavior and the option value of variable pricing. *Management Sci.* **44**(December) 145–160.

Hoch, S. J., X. Dreze, M. E. Purk. 1994. EDLP, Hi-Lo, and margin arithmetic. *J. Marketing* **58**(October) 16–27.

Hotelling, H. 1929. Stability in competition. *Econom. J.* **39** 41–57.

Kreft, I. G. G., J. De Leeuw. 1998. *Introducing Multilevel Modeling*. Sage Publications, Thousand Oaks, CA.

Lal, R., R. C. Rao. 1997. Supermarket competition: The case of every day low pricing. *Marketing Sci.* **16**(1) 60–80.

Morrison, S. A., C. Winston. 1990. The dynamics of airline pricing and competition. *Amer. Econom. Rev.* **80**(2) 389–393.

Morton, F. S., F. Zettelmeyer, J. Silva-Risso. 2001. Internet car retailing. *J. Indust. Econom.* **49**(4) 501–519.

Mussa, M., S. Rosen. 1978. Monopoly and product quality. *J. Econom. Theory* **18** 301–317.

Neels, K. 2000. Congestion pricing and the economic regulation of airports. Transportation Research Board, The National Academies. http://onlinepubs.trb.org/onlinepubs/circulars/ec027/ec027.pdf.

Ortmeyer, G., J. Quelch, W. Salmon. 1991. Restoring credibility to retail pricing. *Sloan Management Rev.* **33**(1) 55–66.

Pan, X., B. T. Ratchford, V. Shankar. 2004. Price dispersion on the Internet: A review and directions for future research. *J. Interactive Marketing* **18**(4) 116–135.

Pedhazur, E. J. 1997. *Multiple Regression in Behavioral Research: Explanation and Prediction*, 3rd ed. Harcourt Brace College, Fort Worth, TX.

Pratt, J. W., D. A. Wise, R. Zeckhauser. 1979. Price differences in almost competitive markets. *Quart. J. Econom.* **93**(2) 189–211.

Ratchford, B. T., X. Pan, V. Shankar. 2003. On the efficiency of Internet markets for consumer goods. *J. Public Policy Marketing* **22**(1) 4–16.

Raudenbush, S. W., A. S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage, Newbury Park, CA.

Raudenbush, S. W., A. S. Bryk, Y. F. Cheong, R. Congdon. 2001. *HLM5: Hierarchical Linear and Nonlinear Modeling*, Scientific Software International, Lincolnwood, IL.

Salop, S. 1979. Monopolistic competition with outside goods. *Bell J. Econom.* **10** 141–156.

Salop, S., J. E. Stiglitz. 1977. Bargains and ripoffs: A model of monopolistically competitive price dispersion. *Rev. Econom. Stud.* **44**(3) 493–510.

Shankar, V., R. N. Bolton. 2004. An empirical analysis of determinants of retailer pricing strategy. *Marketing Sci.* **23**(1) 28–49.

Smith, M. D., J. Bailey, E. Brynjolfsson. 2000. Understanding digital markets: Review and assessment. E. Brynjolfsson, B. Kahin, eds. *Understanding the Digital Economy*. MIT Press, Cambridge, MA, 99–136.

Sorensen, A. T. 2000. Equilibrium price dispersion in retail markets for prescription drugs. *J. Political Econom.* **108**(4) 833–850.

Stavins, J. 2001. Price discrimination in the airline market: The effect of market concentration. *Rev. Econom. Statist.* **83**(1) 200–202.

Steenbergen, M. R., B. S. Jones. 2002. Modeling multilevel data analysis. *Amer. J. Political Sci.* **46**(1) 218–237.

Stigler, G. J. 1961. The economics of information. *J. Political Econom.* **69**(3) 213–225.

Varian, H. R. 1980. A model of sales. *Amer. Econom. Rev.* **70** (4) 651–659.

Venkatesan, R., K. Mehta, R. Bapna. 2006. Understanding the confluence of retailer characteristics, market characteristics and online pricing strategies. *Decision Support Systems* **42**(3) 1759–1775.

Verlinda, J. A., L. Lane. 2004. The effect of the Internet on pricing in the airline industry. Working paper, http://ssrn.com/abstract=965788.

Walter, Z., A. Gupta, B.-C. Su. 2006. The sources of on-line price dispersion across product types: An integrative view of on-line search costs and price premiums. *Internat. J. Electronic Commerce* **11**(1) 37–62.

Weatherford, L. R., S. E. Bodily. 1992. A taxonomy and research overview of perishable-asset revenue management: Yield management, overbooking, and pricing. *Oper. Res.* **40**(5) 831–844.

Zettelmeyer, F., F. S. Morton, J. Silva-Risso. 2001. Cowboys or cowards: Why are Internet car prices lower? NBER working paper, http://ssrn.com/abstract=288601.