Volume 26, No. 3 ISSN 0167-8116
September 2009

ELSEVIER

International Journal of
Research in Marketing
Official Journal of the European Marketing Academy

CONTENTS

Intern. J. of Research in Marketing 26 (2009) 189–196

Contents lists available at ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Quality-adjusted price comparison of non-homogeneous products across Internet retailers ☆

Wagner A. Kamakura [a,1], Sangkil Moon [b,*]

[a] Fuqua School of Business, Duke University, United States
[b] College of Management, North Carolina State University, United States

ARTICLE INFO

Article history:
First received in June 24, 2008 and was under review for 5 months

Area Editor: Russell Winer

Keywords:
Internet retailing
Non-homogeneous products
Price efficiency
Assortment quality
Information presentation
Information search strategy
Stochastic-frontier model

ABSTRACT

This study compares prices offered by multiple Internet retailers. This task is challenging because e-tailers cannot present their entire assortments to each consumer. Therefore, the quality of the product assortments presented by different e-tailers to each consumer is not directly comparable on an item-by-item basis, resulting in non-homogeneous offerings across retailers. We further consider the interaction between retailers (product information presentation format) and consumers (product information search strategies), which makes price comparisons among the retailers even more non-homogeneous. To grapple with this quality-adjusted price comparison problem for non-homogeneous products, we use a stochastic-frontier hedonic-price regression model to find the "lowest" theoretical price for a product given its characteristics. We then assess the price efficiency of the product as the ratio between this lowest price and the offered market price. This framework allows for the comparison of retailers in their ability to offer the "best deals" even when their actual assortments are not directly comparable in quality. Moreover, this framework provides Internet retailers with a relative measure of price efficiency. This helps them understand when and where they offer competitive prices to consumers. We illustrate our approach empirically in a comparison of price efficiency among three major Internet travel agents on a sample of posted itineraries and airfares. Furthermore, we demonstrate that the price efficiency of an Internet travel agent depends on the format of its website and on consumers' search strategies.

1. Introduction

It was widely expected that Internet retailing would lead to a substantial reduction in price dispersion due to reduced search costs and cost transparency (Degeratu, Rangaswamy, & Wu, 2000; Sinha, 2000). This has been confirmed across a wide range of products such as automobiles (Zettelmeyer, Morton, & Silva-Risso, 2006), insurance (Brown & Goolsbee, 2002), and music CDs and books (Brynjolfsson & Smith, 2000). For the same reasons, conventional wisdom also predicted that the Internet would become a nearly perfect market, where well-informed consumers push prices closer to sellers' costs. However, recently reported empirical evidence still shows substantial price dispersions on the Internet for such products as music CDs and books (Brynjolfsson & Smith, 2000) and airline tickets (Clemons, Hann, & Hitt, 2002).

When consumers can find identical products, that is, homogeneous goods across various retailers, comparing prices can be easily solved by using price comparison websites. This straightforward strategy does not work for non-homogeneous situations, where consumers choose the best alternative considering both price and quality attributes (Soberman & Parker, 2006; Estelami et al., 2001). In particular, since online retailers flexibly display a different product assortment reflecting their retailing and pricing strategies, it is critical to assess the competing online retailers' ability to offer the lowest price at the desired quality level. Furthermore, depending on their search goals, consumers can apply various search strategies on the Internet, further enhancing the non-homogeneous nature of such price comparison tasks by interacting with retailers' information presentation formats.

The purpose of this study is to demonstrate how to compare prices across competing Internet retailers offering *non-homogeneous products* (Levy, Grewal, Kopalle, & Hess, 2004). In our empirical application, we compare three major online travel agents. Although these travel agents have a limited say on each ticket price, we demonstrate that there is considerable price dispersion among agents. Due to consumers' search costs and limited information processing capacity, they present only a small fraction of their vast assortment to each

☆ Financial support for Dr. Moon from the Edwin Gill Research Grant and the Services and Production Innovation Management Initiative's (SPI) Xerox Research Fund, both from College of Management, North Carolina State University, is gratefully acknowledged. Data support by Ross Darrow at Sabre Holdings is also gratefully acknowledged. Finally, constructive and insightful comments from Editor Donald R. Lehmann, one anonymous AE, and two reviewers are greatly appreciated.
* Corresponding author. Tel.: +1 919 515 1802; fax: +1 919 515 6943.
E-mail addresses: kamakura@duke.edu (W.A. Kamakura), smoon2@ncsu.edu (S. Moon).
[1] Tel.: +1 919 660 7855; fax: +1 919 684 2818.

0167-8116/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2009.03.004

consumer. Therefore, retailers' selection of items to present to consumers influences their pricing image.

While search costs have decreased substantially in Internet retailing, the amount of information has increased considerably, leading to information overload. Recent studies have pointed out that consumers use various search strategies according to purchase decision complexities, purchase objectives, and demographic characteristics. Furthermore, these factors influence how much consumers are willing to search as well as how much they are willing to pay (Luo, Feng, & Cai, 2004; Morton, Zettelmeyer, & Silva-Risso, 2003; Oorni, 2004; Ratchford, Lee, & Talukdar, 2003). Contrary to popular belief, empirical research shows that consumers' decision quality suffers as the amount of product information increases (Malhotra, 1982; Payne, Johnson, & Bettman, 1993). Moreover, consumers may respond to increased information in ways undesirable to the seller, such as deferring purchase decisions altogether (Dhar, 1997). Understanding who offers the best alternatives to consumers in the Internet environment is important not only to retailers but to consumers.

Another important aspect of Internet retailing is how e-tailers present product information. Previous research reveals that the information presentation mode determines the way information is acquired and used (Painton & Gentry, 1985). In our study, among various presentation characteristics (Brucks, 1988; Kardes & Kalyanaram, 1992; Kleinmuntz & Schkade, 1993), we focus on *sequence* as a factor that influences which products are displayed out of the available assortment. The presentation sequence can determine whether or not specific product offerings are viewed by consumers. Presentation *sequence* is also influenced by consumers' search costs and abilities to process displayed information. So far, marketing researchers have investigated the impact of information presentation mode on information acquisition for single items. They have focused primarily on differences in choice process (brands vs. attributes) using information display boards in an experimental setting (Bettman & Kakkar, 1977; Lehmann & Moore, 1980). By contrast, we investigate the impact of the information presentation format (simultaneous vs. sequential) in a bundled product category (i.e., round-trip flight tickets) in terms of price efficiency. Furthermore, research on information search involving Internet retailing has focused on the increased amount of information due to low search costs for consumers (Lynch & Ariely, 2000). In contrast, we focus on information presentation *sequence* by e-retailers (Hoque & Lohse, 1999; Widing & Talarzyk, 1993).

Because our goal is to compare product prices for non-homogeneous products, we compare them according to their theoretically best prices (the lowest prices after adjusting for differences in quality). That is, we define the hedonic prices at the efficient frontier rather than at the "average" (Hjort-Andersen, 1984; Kamakura, Ratchford, & Agrawal, 1988). For this purpose, we extend traditional hedonic price analysis to a *stochastic-frontier hedonic-price regression* that defines the "lowest feasible" price for a product given its features in the context of all the prices and features of products offered across various competing retailers in the same market.

Our intended contributions are two-fold. First, we combine Rosen's (1974) hedonic price theory and the stochastic-frontier regression methodology (Kumbhakar & Lovell, 2000) to define the "lowest" theoretical price benchmark for assessing the price efficiency of e-tailers selling non-homogeneous products. Although the methodology has been widely used in various applications, we extend it to estimate the lowest price as opposed to the traditional hedonic price regression model based on the estimated average price. Second, we demonstrate that the price efficiency of e-tailers depends on the interactions between retailers' information presentation modes and consumers' search strategies, as these interactions significantly influence the ultimate set of product alternatives considered by consumers. In the next section, we describe our framework for price comparison and show how it assesses the retailer's price efficiency relative to the lowest feasible price frontier.

2. Analysis method

Two methodologies have been commonly used for price comparisons. The first is the well-known hedonic price analysis (Rosen, 1974), which assumes that posted prices and product characteristics trace the joint-envelope of the intersection between consumer utility functions of the product attributes and the supplier's production costs (Berndt, 1991; Epple, 1987; Nerlove, 1995; Working, 1927). However, because hedonic prices are calculated using the least-squares regression line, they reflect average prices rather than the lowest feasible ones. The second methodology (Hjort-Andersen, 1984; Kamakura et al., 1988) attempts to find the lowest feasible price by determining the convex hull enveloping all prices and product attributes with a linear programming tool (known as data envelopment analysis). While this convex hull defines the lowest price at any combination of product features, it does not allow for measurement errors and is therefore highly susceptible to errors and outliers (Fried, Lovell, & Schmidt, 1993). To overcome the problems with the two existing methodologies, we use a *stochastic-frontier hedonic-price regression* that defines a "lowest feasible" price that accounts for random errors, while providing a measure of expected efficiency for each posted price in the market in relation to other alternatives in the same market. Although the stochastic-frontier regression methodology has been widely applied outside of marketing (see Kumbhakar & Lovell, 2000 for a review), our contribution lies in applying it to define theoretical benchmarks based on Rosen's hedonic price theory.

Let y_{ij} be the natural log of the price for product j offered by retailer i, and X_{ij} be a vector of product attributes. Our problem is to estimate the hedonic price regression, not at the expected average, but at the stochastic-frontier (Aigner, Lovell, & Schmidt, 1977; Kalita, 1994; Meeusen & van Den Broeck, 1977; Kumbhakar & Lovell, 2000; Greene, 2003). We define the lowest feasible price as follows:

$$y_{ij} = \beta_0 + \beta_1 x_{1ij} + \ldots + \beta_K x_{Kij} + \varepsilon_{ij}. \tag{1}$$

The disturbance component (ε) is made up of two sub-components:

$$\varepsilon_{ij} = u_{ij} + v_{ij} \text{ where } u_{ij} \sim N^{+}\left(0, \sigma_u^2\right) \text{ and } v_{ij} \sim N\left(0, \sigma_v^2\right). \tag{2}$$

The u term follows a non-negative, half-normal (N^{+}) distribution and captures the inefficiency in the dependent variable, price. The non-negative half-normal disturbance should be between 0 and ∞, and measures the extra charge in the price beyond the lowest theoretical price. The v term follows the usual normal distribution and captures the symmetric disturbance in the model, as well as misspecification errors arising from omitted covariates. The two random components, u and v, are distributed independently of each other and of the regressors. When there is no truncated disturbance component, the proposed model is reduced to the traditional hedonic price regression. Therefore, the stochastic hedonic price frontier model is a generalized version of the original hedonic price regression. Because our main purpose here is to compare retailers in terms of their price efficiency after accounting for product "quality," we specify a single frontier across all retailers rather than allow for unobserved heterogeneity across retailers.

The marginal density function of the disturbance term ε is given by (Kumbhakar & Lovell, 2000, p.140)

$$f(\varepsilon) = 2\sigma^{-1}\phi\left(\varepsilon\sigma^{-1}\right)\Phi\left(\lambda\varepsilon\sigma^{-1}\right), \tag{3}$$

where $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \frac{\sigma_u}{\sigma_v}$ and f and Φ represent the probability density function and the cumulative distribution function of the standard normal distribution, respectively. The density function f is obtained by using the sum of the two values from the normal and truncated normal distributions (Weinstein, 1964). Now, the likelihood of our proposed model is

$$L = \prod_i \prod_j \frac{2}{\sigma} \phi\left(\frac{\varepsilon_{ij}}{\sigma}\right) \Phi\left\{\frac{\lambda \varepsilon_{ij}}{\sigma}\right\} \tag{4}$$

The estimates of $\varepsilon_{ij} = u_{ij} + v_{ij}$ contain information on the non-negative component u_{ij}. If $\varepsilon_{ij} < 0$, the non-negative component u_{ij} is likely to be very small, suggesting that the product is relatively price-efficient. On the other hand, if $\varepsilon_{ij} > 0$, then u_{ij} is likely to be large since $E(v_{ij})$ is 0. A large u_{ij} value suggests that the product is relatively price inefficient. The problem is to extract the information about efficiency (u_{ij}) contained in ε_{ij}. A solution to the problem is obtained from the conditional distribution of u_{ij} given ε_{ij}, which contains whatever information ε_{ij} might provide concerning u_i (Kumbhakar & Lovell, 2000, pp.141–142). The decomposition developed by Jondrow, Lovell, Materov, and Schmidt (1982) can be used to separate noise from price inefficiency in the residuals through the conditional mean, $E(u_{ij} \mid \varepsilon_{ij})$, as follows.

$$E\left[u_{ij} \mid \varepsilon_{ij}\right] = \frac{\sigma\lambda}{1+\lambda^2}\left[\frac{\phi\left(a_{ij}\right)}{1-\Phi\left(-a_{ij}\right)} + a_{ij}\right], \text{ where } a_{ij} = \frac{\lambda\varepsilon_{ij}}{\sigma}. \tag{5}$$

Then, the price efficiency (PE) from the conditional mean is given by

$$\mathrm{PE}_{ij} = \exp\left(-E\left[u_{ij} \mid \varepsilon_{ij}\right]\right). \tag{6}$$

In our application, a more intuitive measure can be obtained by

$$\text{Price Inefficiency Markup}_{ij} = \exp\left(E\left[u_{ij} \mid \varepsilon_{ij}\right]\right) - 1, \tag{7}$$

which represents the markup of the observed price relative to the hypothetical efficient price frontier. The price inefficiency markup measure, which ranges from 0 to ∞, indicates the extra charge on the posted price beyond the lowest theoretical price. Thus, computing price efficiency against the minimum theoretical price is more logical than doing so against the average price obtained from the traditional hedonic price regression.

3. Application: comparing web travel agents

To illustrate our proposed framework for price efficiency comparison, we focus on the travel industry, where the Internet has become a major source of flight bookings. There are three major web agents in the U.S. (Bergey & Moon, 2009; Nair, Chan, & Cheema, 2006): travelocity.com, expedia.com, orbitz.com. Given that search costs across multiple Internet travel agents are fairly low compared to traditional travel agents, and air travel is sufficiently expensive to justify extensive search, economic theory would suggest that prices would converge to marginal costs, thereby eliminating price dispersions, despite the fact that tickets are horizontally differentiated (Bakos, 1997). However, a recent study found considerable price dispersion in this industry, even after accounting for differences in ticket “quality” (Clemons et al., 2002). The sheer magnitude of the travel industry, the considerable price dispersion, and the growing penetration of web travel agents make such Internet price comparisons important and interesting. Notably, *the main difference in flight itineraries and prices across these agents arises from the different assortment of flights they display to each traveler and from the format used to display these selected flights.*

We consider a sample of 30 travel markets defined by the origin–destination airport pair, including both US domestic and international markets. We use airfare search data obtained from a leading online travel agent's proprietary “shopbot” between November 11th and 15th, 2006, after searching for a variety of itineraries lasting no more than 8 days for each of the 30 airline markets. Selected itineraries were obtained 21, 14, seven, and 3 days prior to various departure dates, varying in terms of the departure times of the departing and returning flights. For each type of trip, we limit our analysis to the 20 best (i.e., lowest fare) itineraries of each agent in conjunction with consumers' various search strategies described below. Our data selection procedure resulted in 95,246 searched itineraries across the 30 markets. The itinerary characteristics are shown in Tables 1 and 2.

In terms of product presentation format, Orbitz uses a one-step (simultaneous) format for round-trips, listing complete itineraries including departing and returning flights on one screen. In contrast, its two main competitors (Expedia and Travelocity) use a two-step (sequential) format, showing the best fares listed by departing flights, allowing travelers to choose returning flights for any of the departing ones. The one-step presentation results in a long list of round-trips, thereby limiting the options considered by the typical traveler. On the other hand, the two-step format requires that the traveler search separate lists for the departing and returning flights, leading to higher search costs and different search strategies, but providing more flexibility in finding the best combination of departing and returning tickets. We assume that the traveler limits her search to 20 itineraries

Table 1
Average itinerary characteristics of various agent-search strategies (3 agents and 13 search strategies).

Agent-search strategy	Number of itineraries	Mean airfare ($)	Search before departure (days)	Trip length (days)	Elapsed time/distance (min/km)	Same airline (%)	Airline category (%)
Orbitz-20	14,320	511	11.3	3.3	.1458	67	MA (81), NMA (15), NA (4)
Expedia-20&1	5868	430	11.5	2.5	.1360	59	MA (86), NMA (12), NA (2)
Expedia-10&2	5970	408	11.3	2.5	.1333	62	MA (85), NMA (12), NA (3)
Expedia-5&4	5933	412	11.3	2.5	.1329	62	MA (86), NMA (11), NA (3)
Expedia-4&5	5911	428	11.3	2.5	.1332	59	MA (86), NMA (11), NA (3)
Expedia-2&10	5790	499	11.3	2.5	.1348	43	MA (85), NMA (11), NA (4)
Expedia-1&20	5567	604	11.3	2.5	.1343	27	MA (86), NMA (10), NA (4)
Travelocity-20&1	9032	457	11.6	3.5	.1523	94	MA (84), NMA (15), NA (1)
Travelocity-10&2	8386	411	11.7	3.4	.1562	92	MA (82), NMA (17), NA (1)
Travelocity-5&4	7774	402	11.5	3.4	.1647	88	MA (80), NMA (18), NA (2)
Travelocity-4&5	7522	395	11.3	3.4	.1674	86	MA (79), NMA (19), NA (2)
Travelocity-2&10	6786	416	11.5	3.4	.1700	85	MA (79), NMA (19), NA (2)
Travelocity-1&20	6387	440	11.6	3.4	.1734	84	MA (78), NMA (20), NA (2)

Notes: Search strategy x&y indicates a combination of the number of outbound itineraries (x) and inbound itineraries (y) observed by the traveler. Same airline indicates the proportion of round-trip itineraries that listed all the flights from an identical airline. Airline category: MA = 6 major American airlines, NMA = non-major American airlines, and NA = non-American airlines.

Table 2
Stochastic-frontier hedonic-price regression estimates.

Parameter		Estimate	Parameter		Estimate
Intercept		4.461***	Inbound departure time	10 am–3 pm	0.023***
Search day before departure	7 days	−0.036***	(base category − 5 am–10 am)	3 pm–8 pm	0.021***
(dummies: base category − 3 days)	14 days	−0.034***		8 pm–5 am	0.031***
	21 days	−0.398***	Elapsed time (min)/distance (km)		−0.678***
Trip length	2 days	−0.187***	Travel distance (log(km))		0.189***
(dummies: base category − same day trip)	3 days	−0.135***	Number of airline companies		−0.009***
	4 days	−0.099***	Same airline (outbound = inbound)		−0.190***
	5 days	−0.213***	Market category (base category − domestic market)	Hawaiian	0.654***
	6 days	−0.053***		International	0.829***
	7 days	−0.072***	Airline category	Non-major American	0.035***
	8 days	−0.065***	(base category − major American airlines)	Non-American	0.025***
Outbound departure day of the week	Monday	−0.126***	Residuals	λ	3.030***
(dummies: base category − Sunday)	Tuesday	−0.200***		σ	0.536***
	Wednesday	−0.028***	Model goodness of fit: R^2 = .480;		
	Thursday	−0.061***	log likelihood = −34,713.0; BIC = 65,532.3		
	Friday	0.021***			
Outbound departure time	8 am–10 am	0.062***			
(dummies: base category − 6 am–8 am)	10 am–12 pm	0.066***			
	12 pm–2 pm	−0.001			
	2 pm–4 pm	0.018***			
	4 pm–6 pm	0.033***			
	6 pm–6 am	0.051***			

Notes: The dependent variable is the natural log of the airfare. Therefore, a positive estimate value indicates that the characteristic or category is more expensive (i.e. less price-efficient). *** indicates significance at the.01 level. Search day before departure indicates how many days the search was conducted before the departure day. Regarding the departure day of the week dummy variables, there were no Saturday departures. Same airline indicates the proportion of the round-trip itineraries that listed all the flights from the identical airline company. The number of stops variable was excluded because of its high correlations with three other significant variable — elapsed time/distance, travel distance, and market category.

for each upcoming trip. Thus, she can choose the first returning flights for each of the 20 cheapest departing flights (which we label as the *20&1* search strategy) or search the best two returning flights for each of the 10 cheapest departing ones (*10&2*), and so on, down to the 20 top returning flights for the cheapest departing one (*1&20*). As a result, we simulate 13 different search strategies: 6 strategies (*20&1*, *10&2*, *5&4*, *4&5*, *2&10*, and *1&20*) for each of the two sequential format retailers and only one strategy for the single one-step format retailer. We denote these different search strategies by "*x&y*" where x indicates the number of the cheapest departing flights viewed on the initial screen, whereas y indicates the number of the cheapest returning flights viewed for each of the viewed departing flight (refer to Table 1). Because each of these search strategies results in a different assortment of itineraries presented to the traveler, the price efficiency of a travel agent using the two-stage format depends on the particular search strategy adopted by the traveler.

3.1. Direct price comparisons

We present empirical evidence that the actual itineraries presented by the competing travel agents to each traveler are non-homogeneous, even when tickets for specific flights offered at a specific point in time are homogeneous. To do this, we searched through the 48,997 unique itineraries listed in our sample. We found that only 572 specific itineraries were common across the three agents. The very small proportion (1.2%) of directly comparable itineraries demonstrates that the three agents post very different itineraries from the same pool of products, and therefore offer non-homogeneous itineraries. As one would expect, a comparison of prices for the 572 homogeneous itineraries showed negligible price differences. An ANOVA comparing the three agents for the common basket of itineraries failed to show statistically significant differences (p-value = .980). In contrast, a comparison of mean airfares across the 13 different agents and search strategies shows substantial price differences that reflect not only the efficiency of these agents in finding the best prices for a trip but differences in "quality" (e.g. search days before departure, trip length) in the assortments, as shown in Table 1. Accordingly, the main purpose of our proposed model is to define a common benchmark for the three agents in order to compare the agents on a common "quality" level, despite the fact that they offer non-homogeneous travel itineraries.

3.2. The stochastic hedonic price frontier and price efficiencies

The application of our proposed framework to the search data across all markets and travel agents led to the hedonic price frontier estimates presented in Table 2. We also compared our model with traditional hedonic price regression, and found that our stochastic-frontier model performs better in the Bayesian Information Criterion (stochastic-frontier model = 65,532.3; reduced model = 75,265.4). This indicates that the added asymmetric term helps capture price inefficiency in the market. This conclusion is supported by the strong residual term estimates in the stochastic-frontier model, as shown in Table 2.

The estimation results of our stochastic-frontier can be interpreted in a similar way to the results of a linear regression, except that they reflect the impact of each predictor on the lowest theoretical price, instead of the posted price. As one would expect, the best airfares tend to be lower when they are posted well ahead of the departure date, particularly more than 2 weeks ahead of departure. Same-day trips are more expensive than longer trips, with five-day trips being the least expensive. Trips initiated on Fridays or Sundays are generally more expensive than those initiated during mid-week days. The earliest departure times offer the lowest prices (before 8:00 am for departing flights and before 10:00 am for returning flights), although the differences are less substantial for returning flights. This shows that airlines tend to discount early departures more heavily than those during regular business hours because those times are less convenient for travelers.

As one would expect, trips involving larger distances are more expensive. The elapsed time/distance variable measures how time-efficient the itinerary is given the direct distance between the origin and the destination. Its negative sign indicates that travelers are willing to pay more for more time-efficient itineraries (i.e. shorter travels for the same distance). Markets where more airlines compete for passengers offer lower airfares. Airfares also tend to be lower if the same airline is used for both the departing and returning flights. After discounting for distance and competition, trips to Hawaii and

Table 3
Price inefficiency markup by the agent-search strategy.

Agent-search strategy	Number of itineraries	Airfare ($)	Price inefficiency markup by each search days before departure (%)				
			21 days before	14 days before	7 days before	3 days before	All days
Orbitz-20	14,320	511	48	51	55	52	52
Expedia-20&1	5868	430	48	52	64	59	56
Expedia-10&2	5970	408	43	49	53	51	49
Expedia-5&4	5933	412	45	54	54	53	52
Expedia-4&5	5911	428	53	60	57	54	56
Expedia-2&10	5790	499	82	83	71	63	75
Expedia-1&20	5567	604	128	114	94	79	104
Travelocity-20&1	9032	457	42	88	69	76	68
Travelocity-10&2	8386	411	35	70	57	58	54
Travelocity-5&4	7774	402	34	59	60	51	50
Travelocity-4&5	7522	395	33	54	58	52	48
Travelocity-2&10	6786	416	37	56	68	58	54
Travelocity-1&20	6387	440	46	62	82	65	62

Agent-search strategy	Direct flights		Non-direct flights		Domestic flights		International flights	
	Mean airfare ($)	Mean markup (%)	Mean airfare ($)	Mean markup (%)	Mean airfare ($)	Mean markup (%)	Mean airfare ($)	Mean markup (%)
Orbitz-20	298	44	571	54	381	47	1152	76
Expedia-20&1	316	49	492	59	413	55	1575	116
Expedia-10&2	293	43	480	53	388	48	1588	116
Expedia-5&4	302	44	485	56	395	51	1414	101
Expedia-4&5	327	51	495	59	413	55	1449	106
Expedia-2&10	421	76	543	74	485	74	1466	108
Expedia-1&20	511	101	654	105	591	104	1466	108
Travelocity-20&1	387	66	497	69	434	67	1299	100
Travelocity-10&2	337	51	455	55	389	53	1217	91
Travelocity-5&4	317	43	445	53	375	48	1251	94
Travelocity-4&5	319	43	432	51	367	47	1243	92
Travelocity-2&10	346	50	442	55	387	52	1238	92
Travelocity-1&20	370	55	460	64	409	61	1238	92

Notes: The search strategy *x*&*y* indicates a combination of the number of outbound itineraries (x) and inbound itineraries (y) observed by the traveler. Search days before departure indicates how many days the search was conducted before the departure day. Search date distribution: 21 days ($N=25{,}791$, 27.1%), 14 days ($N=21{,}837$, 22.9%), 7 days ($N=22{,}124$, 23.2%), and 3 days ($N=25{,}494$, 26.8%).

international trips are more expensive than domestic trips. Also, major domestic airlines offer lower airfares than minor domestic and non-American airlines.

From the non-negative residual component (u) in Eq. (2), we obtain estimates of the price inefficiency for each travel agent and search strategy. Based on the estimation results of Table 2, the efficiency measures in Table 3 show the extent to which prices posted by each travel agent (under different search strategies) are higher than the lowest expected price defined by the stochastic-frontier. We also show such price inefficiency "markups" for four different search days because airlines consider this advance search factor in determining ticket prices in their revenue management (Desiraju & Shugan, 1999).

In general, Expedia and Travelocity produce higher price inefficiency markups if viewed in the same way as Orbitz, i.e. by looking only at the cheapest round-trip for each departing flight. For example, viewing only the 20 best fares defined by the departing flight (the *20&1* search strategy) produces inefficiency markups of 48–64% for Expedia and 42–88% for Travelocity compared to 48–55% for Orbitz. Searching the 20 best returning flights for the best departing one (the *1&20* strategy) becomes much less efficient for Expedia, producing inefficiency markups of 79–128%. Instead of these unbalanced "extreme" search strategies (*20&1* and *1&20*), more balanced "mid-way" search strategies (5&4 and 4&5) tend to generate itineraries whose prices are as efficient as the Orbitz-20 case as shown in the comprehensive "all days" column in Table 3. This implies that Orbitz (using the one-step presentation format) chooses well-distributed itineraries in price that are similar to those resulting from the mid-way strategies of its competitors using the two-step presentation format. Therefore, experienced customers will be better off with the two-step format, particularly when they are willing to spend more time searching.

When booking a trip 21 days ahead of departure, a consumer is better off with Travelocity, looking at the five cheapest returning flights for each of the four cheapest departing flights (*4&5* search strategy), which produces fares that are on average only 33% higher than the theoretical minimum prices. On the other hand, if booking the trip 14, 7, or 3 days before departure, the best prices are attained with the *10&2* search strategy at Expedia. Overall, the lowest inefficiency markups (relative to the price frontier) are achieved with Travelocity, using the *4&5* search strategy (48% markup).

Obviously, searching for the five best returning flights for each of the four best departing ones offered by Travelocity is more cumbersome than simply comparing the 20 best itineraries presented by Orbitz on a single screen (52% markup). Then, one must question whether the extra search costs with the two-step search are compensated by the additional savings of average 48% instead of 52% inefficiency markups (Hoque & Lohse, 1999; Wu & Rangaswamy, 2003). Previous research by Card, Moran, and Newell (1983) shows that, on average, pointing to a menu item using a mouse takes 1.10 s and clicking a mouse button takes 0.36 s. Based on these statistics, it would take 5.12 s to click on one departing flight and look at its 5 best returning flights (4 "pointings" and 2 "clicks"), which results in a total of 20.48 s for the Travelocity—*4&5* strategy. By comparison, it takes a total of 8.30 s (1 "pointing" and 20 "clicks") to look down with the Orbitz-20 strategy. The time calculations do not include content reading based on the assumption that travel agents provide the same itinerary information, which requires the same reading effort. Therefore, the incremental time cost resulting from the Travelocity *4&5* strategy relative to the Orbitz-20 strategy is only 12.18 s, which appears to be more than the 4% inefficiency markup advantage of an average $450 airfare (unless the traveler has an opportunity cost greater than $5000 per hour). However, in this simple calculation, we

Table 4
Comparison of price inefficiency among the three agents.

Itinerary characteristic		Using all search strategies			Using search strategy 5&4 or 4&5		
		Expedia	Travelocity	Most efficient agent	Expedia	Travelocity	Most efficient agent
Intercept		.401***			.405***		
Search day before departure	7 days	.034***	.056***	Orbitz	.018**	.073***	Orbitz
(base category — 3 days)	14 days	.034***	.063***	Orbitz	.022**	.063***	Orbitz
	21 days	.040***	−.104***	Travelocity	−.001	−.073***	Travelocity
Trip length	2 days	.073***	.047***	Orbitz	−.024	.008	No difference
(base category — same day trip)	3 days	.098***	.048***	Orbitz	−.008	.018*	Orb. = Expedia
	4 days	.104***	.060***	Orbitz	.104***	.038**	Orbitz
	5 days	.115***	.022*	Orbitz	.115***	.022	Orb = travel
	6 days	.180***	−.093***	Travelocity	.078**	−.117***	Travelocity
	7 days	.034	.084***	Orb. = Expedia	−.091***	.079***	Expedia
	8 days	.028***	.065***	Orbitz	.031***	.030***	Orbitz
Outbound departure day of the week	Mon.	.033***	.006	Orb. = travel.	.051***	.012*	Orbitz
(base category — Sunday)	Tues.	−.048***	.144***	Expedia	−.060***	.154***	Expedia
	Wed.	−.037***	.054***	Expedia	−.033***	.053***	Expedia
	Thu.	−.064***	.024**	Expedia	−.060***	.024*	Expedia
	Fri.	.051***	−.060***	Travelocity	.055***	−.035***	Travelocity
Outbound departure time	8 am–10 am	−.044***	.020***	Expedia	−.021***	.009*	Expedia
(base category — 5 am–10 am)	10 am–12 pm	−.035***	.002	Expedia	−.017**	−.018**	Travelocity
	12 pm–2 pm	−.036***	−.004	Expedia	−.027***	−.010	Expedia
	2p–4 pm	−.030***	−.052***	Travelocity	−.011	−.054***	Travelocity
	4–6 pm	−.059***	−.009	Expedia	−.073***	.001	Expedia
	6 pm–6 am	.003	−.008*	Travelocity	−.024	−.012*	Travelocity
Inbound departure time	10 am–3 pm	.022***	.003	Orbitz	.005	.010**	Orb. = Expedia
(base category — 5 am–10 am)	3 pm–8 pm	−.015***	.017***	Expedia	−.018***	.040***	Expedia
	8 pm–5 am	−.023***	−.012***	Expedia	−.006	−.011*	Travelocity
Market category	Hawaiian	.001	.154***	Orbitz	−.030***	.149***	Expedia
(base category — domestic)	International	.212***	.212***	Orbitz	.234***	.245***	Orbitz
Travel distance (log(km))		.002	.002	No Difference	.008**	−.029***	Travelocity
Same airline (outbound = inbound)		−.138***	.168***	Expedia	−.085***	.125***	Expedia

Notes: The dependent variable is the conditional expected value of the non-negative residual ($E[u_{ij} \mid \varepsilon_{ij}]$) in Eq. (5) with itinerary characteristics as covariates. Base agent category = Orbitz. A positive estimate value indicates that the agent is less price-efficient.
Regarding the departure day of the week dummy variables, there were no Saturday departures.
* Significant at .10 level.
** Significant at .05 level.
*** Indicates significance at the.01 level.

do not consider the non-price and non-time aspects (e.g. mental effort, search enjoyment, search effectiveness) of a consumer's search.

The second page of Table 3 shows how two key itinerary characteristics (direct vs. non-direct flights and domestic and international flights) can be correlated with price inefficiency. We found that more expensive tickets involving more product characteristics (e.g. non-direct or international flights) tend to be less price-efficient. This comparison reveals that price inefficiency is magnified for more complex goods with more product characteristics. We believe this pattern emerges when retailers or price-setters take advantage of situations where it is harder for consumers to comprehend the relationship between quality and price.

3.3. Price efficiency analysis of itinerary characteristics

The price inefficiency markup averages presented in Table 3 compare the three travel agents. Our proposed framework for measuring the price efficiency of web agents generates an efficiency measure for each posted price, providing managers with a more detailed assessment of their price (dis)advantages in relation to the price frontier. To further understand the price competitiveness of the travel agents, we fit a linear regression model. We take the conditional expected value of the non-negative residual ($E[u_{ij} \mid \varepsilon_{ij}]$) in Eq. (5) as the dependent variable with some itinerary characteristics as covariates, as shown in Table 4. In this table, we evaluate Expedia and Travelocity across their six different search strategies using Orbitz as the basis for comparison. Because the dependent variable in this analysis ($E[u_{ij} \mid \varepsilon_{ij}]$) represents the log-inefficiency for each posted fare, a positive regression coefficient indicates that the agent is less efficient (i.e. more expensive) than Orbitz and vice-versa.

The estimates reported in the middle columns of Table 4 were obtained using all 13 search strategies and, therefore, represent the overall performance of Expedia and Travelocity compared to Orbitz. These results show that Travelocity and Expedia post higher fares than Orbitz for the same type of itinerary when the search is done less than 3 weeks prior to departure. When the search is done 3 weeks prior to departure, Travelocity posts the lowest fares for the same type of itineraries. Orbitz also tends to produce lower fares for most trip lengths, except for trips lasting 6 days, for which Travelocity produces the lowest fares. By contrast, Expedia tends to be more efficient in most of the outbound departure days of the week as well as the outbound and inbound departure times, with a few exceptions. Orbitz appears dominant in the market category, whereas Expedia is the best for itineraries involving the same airline on a round-trip. In addition, some non-price aspects are not flexible to certain travelers. For example, for a business traveler who must leave by 7 a.m. to be at a meeting, any trip outside the time period is irrelevant.

On the other hand, the results shown on the right side of Table 4 were obtained under the assumption that customers use either a 5&4 or 4&5 search strategy on Expedia or Travelocity. As one would expect, it is now more favorable to use both Expedia and Travelocity. These "middle-way" search strategies are more efficient than the two agents' other "more extreme" search strategies.

4. Discussion

In this study, we propose a stochastic-frontier hedonic-price framework to assess the price competitiveness of retailers offering non-homogeneous assortments where prices cannot be directly compared on an item-by-item basis. The online retailing environment (Shankar, Smith, & Rangaswamy, 2003) leads to substantial differences in product assortment offerings among competing retailers, because retailers are constrained by the number of products actually seen by each consumer, even when the total inventory of itineraries may not differ substantially across retailers. Moreover, these assortments are affected by the consumer's search strategy. Our proposed framework for a price efficiency comparison accounts for not only the non-homogeneous nature of product assortment but also consumers' heterogeneous search strategies. With the sequential information format, consumers who can search effectively are better off taking more balanced “mid-way” search strategies (5&4 or 4&5) as opposed to the extreme search strategies (20&1 or 1&20), because mid-way search strategies cover a wider range of products in both price and non-price attributes. In contrast, individual search strategies are less important with the more definitive simultaneous format because retailers select a product assortment that covers a wide range of products. However, the simultaneous format lacks the assortment depth produced by the two-step format because the number of products presented in the former is more limited.

Even though our empirical illustration focused on Internet travel agents, our approach is applicable to other products or markets where retailers vary in product assortment. One potentially useful application to consumers would be the price comparison of durable goods, especially cases where products from competing retailers are not directly comparable due to different product features and accessories. For example, an increasingly common strategy used by traditional retailers to avoid direct price comparisons against more efficient “e-tailers” is to require manufacturers to supply slightly modified models unique to their stores. Another common practice among retailers is to advertise low prices for a common basket of well-known products and brands, while making their profits on other, non-homogeneous items. The application of our framework based on the minimum theoretical price to both traditional and Internet retailers would allow consumers to assess the relative price efficiency of non-homogeneous items across retailers.

From the retailer's perspective, our study demonstrates how its information display strategy, in conjunction with consumers' information search strategy, can affect price efficiency. In particular, the Internet allows online retailers to combine and present related products (e.g. computer, monitor, and printer) as a bundle with negligible incremental costs (Kannan & Kopalle, 2001). Therefore, retailers can present either each component of the bundle sequentially or the whole bundle simultaneously. Our empirical analysis allows the retailer to measure the interaction effect between the product presentation mode and the consumer's search strategy and choose the more price-efficient product presentation mode for a specific competition situation.

Given the fact that airfares are primarily set by airlines rather than travel agents, readers might wonder why we assessed price dispersion in price efficiency across the retailers. In our comparison, we emulated the search behavior of typical travelers, who will not consider all available displayed itineraries. Since the presentation sequence for the selected itineraries vary across competing travel agents, the actual assortments shown to travelers vary considerably across agents, resulting in non-homogeneous itineraries. By considering the presentation formats by the agents and manipulating itinerary sequences, Internet retailers can effectively establish a unique pricing image that they can highlight in their advertising.

Lastly, due to the lack of such data, our study considered only prices posted by shopping agents and did not include actual bookings and prices paid by consumers. We follow Rosen's (1974) notion of the hedonic price function as the joint-envelope of the intersection between consumers' utility functions of the product attributes and the suppliers' costs. Since we use posted market prices and itinerary features, our data represent all points where supply and demand functions intersect, rather than the trace of either of these two functions. Our main goal here is to identify the price frontier from posted prices and itinerary features, rather than to estimate the demand or supply functions. Demand-side data would allow researchers to better understand the whole picture of price efficiency by simultaneously and explicitly considering both supply and demand.

References

Aigner, D. C., Lovell, A. K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*, 21−37.

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

Bergey, P. K., & Moon, S. (2009). Conditional efficiency, operational risk and electronic ticket pricing strategies for the airline industry. *International Journal of Electronic Marketing and Retailing*, *2*(3), 239−255.

Berndt, E. R. (1991). *The practice of econometrics: Classic and contemporary.* Reading, Massachusetts: Addison-Wesley Publishing Company.

Bettman, J. R., & Kakkar, P. (1977). Effects of information presentation format on consumer information acquisition strategies. *Journal of Consumer Research*, *3* (March), 233−240.

Brown, J. R., & Goolsbee, A. (2002). Does the Internet make markets more competitive? Evidence from the life insurance industry. *Journal of Political Economy*, *110*(5), 481−507.

Brucks, M. (1988). Search monitor: An approach for computer-controlled experiments involving consumer information search. *Journal of Consumer Research*, *15*(1), 117−121.

Brynjolfsson, E., & Smith, M. D. (2000). Frictionless commerce? A comparison of Internet and conventional retailers. *Management Science*, *46*(4), 565−585.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human–computer interaction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Degeratu, A., Rangaswamy, A., & Wu, J. (2000). Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes. *International Journal of Research in Marketing*, *17*, 55−78.

Desiraju, R., & Shugan, S. M. (1999). Strategic service pricing and yield management. *Journal of Marketing*, *63*(January), 44−56.

Dhar, R. (1997). Consumer preference for a no-choice option. *Journal of Consumer Research*, *24*(September), 215−231.

Epple, D. (1987). Hedonic prices and implicit markets: Estimating demand and supply functions for differentiated products. *Journal of Political Economy*, *95*(1), 59−80.

Estelami, H., Lehmann, D. R., & Holden, A. C. (2001). Macro-economic determinants of consumer price knowledge: A meta-analysis of four decades of research. *International Journal of Research in Marketing*, *18*, 341−355.

Fried, H. O., Lovell, A. K., & Schmidt, S. S. (1993). *The measurement of productive efficiency: Techniques and applications.* New York, NY: Oxford Press.

Greene, W. H. (2003). *Econometric analysis*, 5th Ed. : Prentice Hall.

Hjort-Andersen, C. (1984). The concept of quality and the efficiency of markets for consumer products. *Journal of Consumer Research*, *11*, 708−718.

Hoque, A. Y., & Lohse, G. L. (1999). An information search cost perspective for designing interfaces for electronic commerce. *Journal of Marketing Research*, *36*(3), 387−394.

Jondrow, J. C., Lovell, A. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier product function model. *Journal of Econometrics*, *19*, 233−238.

Kalita, J. K. (1994). Measuring product market efficiency: A new methodology. *Marketing Letters*, *5*(1), 77−89.

Kamakura, W. A., Ratchford, B., & Agrawal, J. (1988). Measuring market efficiency and welfare loss. *Journal of Consumer Research*, *15*, 289−302.

Kannan, P. K., & Kopalle, P. K. (2001). Dynamic pricing on the Internet: Importance and implications for consumer behavior. *International Journal of Electronic Commerce*, *5* (3), 63−83.

Kardes, F. R., & Kalyanaram, G. (1992). Order-of-entry effects on consumer memory and judgment: An information integration perspective. *Journal of Marketing Research*, *29*(3), 343−357.

Kleinmuntz, D. N., & Schkade, D. A. (1993). Information displays and decision processes. *Psychological Science*, *4*(4), 221−227.

Kumbhakar, S. C., & Lovell, A. K. (2000). *Stochastic frontier analysis.* Cambridge, UK: Cambridge Press.

Lehmann, D. R., & Moore, W. L. (1980). Validity of information display boards: An assessment using longitudinal data. *Journal of Marketing Research*, *17*(November), 450−459.

Levy, M., Grewal, D., Kopalle, P. K., & Hess, J. D. (2004). Emerging trends in retail pricing practice: Implications for research. *Journal of Retailing*, *80*, 13−21.

Luo, M., Feng, R., & Cai, L. A. (2004). Information search behavior and tourist characteristics: The Internet vis-à-vis other information sources. In Juline E. Mills & Rob Law (Eds.), *Handbook of Consumer Behavior, Tourism, and the Internet* (pp. 15−25). : The Haworth Hospitality Press.

Lynch, J. G., & Ariely, D. (2000). Wine online: Search costs affect competition on price, quality, and distribution. *Marketing Science*, *19*(1), 83−103.

Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research*, *8*(4), 419−430.

Meeusen, W., & van Den Broeck, J. (1977). Efficiency estimation from Cobb–Douglas production function with composed error. *International Economic Review*, *18*(2), 435−444.

Morton, F. S., Zettelmeyer, F., & Silva-Risso, J. (2003). Consumer information and discrimination: Does the Internet affect the pricing of new cars to women and minorities? *Quantitative Marketing and Economics*, *1*, 65−92.

Nair, C.T.R., Chan, T.Y., Cheema, A. 2006. Modeling online browsing and purchase of airline tickets. Working paper, Washington University.

Nerlove, M. (1995). Hedonic price functions and the measurement of preferences: The case of Swedish wine consumers. *European Economic Review*, *39*, 1697—1716.

Oorni, A. (2004). Consumer objectives and the amount of search in electronic travel and tourism markets. In Juline E. Mills & Rob Law (Eds.), *Handbook of Consumer Behavior, Tourism, and the Internet* (pp. 3—14). : The Haworth Hospitality Press.

Painton, S., & Gentry, J. W. (1985). Another look at the impact of information presentation format. *Journal of Consumer Research*, *12*(September), 240—244.

Payne, J. W., Johnson, E. J., & Bettman, J. R. (1993). *The adaptive decision maker.* Cambridge: University Press.

Ratchford, B. T., Lee, M., & Talukdar, D. (2003). The impact of the Internet on information search for automobiles. *Journal of Marketing Research*, *40*(May), 193—209.

Rosen, S. (1974). Hedonic prices and implicit markets. *Journal of Political Economy*, *82*, 34—55.

Shankar, V., Smith, A. K., & Rangaswamy, A. (2003). Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing*, *20*, 153—175.

Sinha, I. (2000). Cost transparency: The net's real threat to prices and brands. *Harvard Business Review March–April* (pp. 43—50).

Soberman, D. A., & Parker, P. M. (2006). The economics of quality-equivalent store brands. *International of Journal of Research in Marketing*, *23*, 125—139.

Weinstein, M. A. (1964). Query 2: The sum of values from a normal and a truncated normal distribution. *Technometrics*, *6*(1), 104—105.

Widing, R. E., & Talarzyk, W. W. (1993). Electronic information systems for consumers: An evaluation of computer-assisted formats in multiple decision environments. *Journal of Marketing Research*, *30*(2), 125—141.

Working, E. J. (1927). What do statistical “demand curves” show? *Quarterly Journal of Economics*, *41*, 212—235.

Wu, J., & Rangaswamy, A. (2003). A fuzzy set model of search and consideration with an application to an online market. *Marketing Science*, *22*(3), 411—434.

Zettelmeyer, F., Morton, F. S., & Silva-Risso, J. (2006). How the Internet lowers prices: Evidence from matched survey and automobile transaction data. *Journal of Marketing Research*, *43*(May), 168—181.