A Choice-Based Dynamic Programming Approach for Setting Opaque Prices

Chris K. Anderson, Cornell University
Xiaoqing Xie, Shanghai University of Finance and Economics

Opaque pricing is a form of pricing where certain characteristics of the product or service are hidden from the consumer until after purchase. In essence, opaque selling transforms a differentiated good into a commodity. Opaque pricing has become popular in service pricing as it allows firms to sell their differentiated product at higher prices to regular brand loyal customers while simultaneously selling to non-brand loyal customers at discounted prices. We use a nested logit model in combination with logistic regression and dynamic programming to illustrate how a service firm can optimally set prices on an opaque sales channel. The choice model allows the characterization of consumer trade-offs when purchasing opaque products while the dynamic programming approach allows the characterization of the optimal pricing policy as a function of inventory and time remaining. We compare optimal prices and expected revenues when dynamic pricing is restricted to daily price changes. We provide an illustrative example using data from an opaque selling mechanism (Hotwire.com) and a Washington DC-based hotel.

Introduction

Using opaque products allows firms to reach price-sensitive consumers by offering discounts given the uncertainty of the seller while simultaneously selling at higher prices to brand loyal consumers on regular non-opaque channels. While opaque selling is most commonly used in the selling of travel services and pioneered by Priceline.com, as discussed in Fay (2008) it offers opportunities for numerous retailers. The pricing of services (rooms, rental cars, airline seats, etc.) online has dramatically changed how service firms reach customers, with online travel sales now exceeding offline (or traditional sales channels). Initial thoughts about pricing online were very positive as firms had new channels to reach customers enabling increased opportunities for segmentation. Over time service providers have

increased efforts to move customers back to company direct distribution channels (company websites and call centers) in an effort to control sales costs and commissions while maintaining direct contact with the customer to facilitate loyalty programs and other marketing efforts. Hotwire and Priceline, unlike other online travel sites such as Expedia, Travelocity, and Orbitz, offer customers opaque products (e.g., hotel rooms, flights, and rental cars) with aspects of the service provider concealed until the transaction has been completed. For instance a customer purchasing a hotel room through Hotwire can only specify check-in/out dates, a subarea within a city, and a star rating. Customers do not know the identity or exact location of their non-refundable choice property until after purchase. Opaque travel sites offer service providers a convenient channel to segment customers and distribute discounted products without cannibalizing or diluting full priced products. The opaque channels naturally segment customers as regular full price paying customers desiring to stay at the hotel of their choice with full cancelation flexibility are different from those willing to purchase the non-refundable opaque product at the unknown service provider. Therefore service providers need to understand better how to optimally set prices on opaque channels. This question is the focus of our study.

In this study we develop a nested logit (NL) model of customer choice using data from transactions occurring at Hotwire.com. The choice model provides insight into the impact of property characteristics that customers can access before purchases (i.e., price, star rating, and neighborhood) on the market share. We also estimate a logistic regression model using data from a Washington DC-based hotel to determine whether or not a service firm is displayed at Hotwire. These two price-dependent models are then combined with dynamic programming to determine optimal prices for the service firm to post at Hotwire.

There is a growing interest in modeling demand using discrete choice models to improve revenue management and pricing. Unlike most choice-based revenue management research, which uses an individual firm's sales data for parameter estimation, our data includes market level purchase transaction information (purchases across all service providers, not a single provider) as well as records of the requests that did not induce an actual purchase. We solve a stochastic dynamic program to set optimal choice-based prices for a service provider to post on the intermediary. We characterize the structure of this optimal policy and benchmark optimal fixed price policies against dynamic pricing policies.

In section 2 we briefly summarize the related literature, section 3 describes our data and estimation of the NL model as well as the logistic regression, with section 4 devoted to the development of the dynamic program for setting dynamic and daily fixed prices. A numerical example in section 5 is

used to illustrate how pricing policies change with capacity and length of selling horizon. Finally, we provide some conclusions and possible future work in section 6.

Literature Review

Most of the early research on opaque pricing focusses on Priceline's Name-Your-Own-Price mechanism or similar opaque bidding channels (e.g., Anderson 2009, Fay 2004, Hann and Terwiesch 2003, Terwiesch et al. 2005, Wang et al. 2009). As we focus on the posted price opaque model of Hotwire versus the bidding format of Priceline we refer readers to Anderson (2009) for a review of this research stream. Here we summarize research on posted price opaque mechanisms. Fay (2008) is the first to investigate how product opacity affects the market. Fay studies two competing service providers selling products to two types of customers (business and leisure) on both an opaque posted price channel and a traditional distribution channel. Fay shows that opaque selling benefits the monopoly service provider when customers have heterogeneous values for products. Similar to Fay (2008), Jiang (2007) also considers heterogeneous customers and studies how online opaque selling, as a new price discrimination technique, improves profits for service providers. Jiang (2007) compares opaque selling and regular selling (selling full-information products), providing insight when to implement opaque selling. Jiang (2007) also investigates the impact of opaque selling on social welfare. Shapiro and Shi (2008) focus on providing a rationale for opaque selling. They explain why service providers are willing to distribute products through opaque travel sites such as Priceline and Hotwire and lose the advantage of product differentiation. Jerath et al. (2010) compare opaque selling with last-minute direct selling and obtain the conditions under which opaque selling is preferred.

Research on the application of choice models is quite diverse in both the marketing as well as the operations literature. Discrete choice models have been employed to study consumer choice behavior and understand consumers' brand preferences, market structure as well as product attributes. The study by Coretjens and Gautschi (1983) provides a general survey of discrete choice models in marketing and a systematic introduction to discrete choice modeling theory. During the past 40 years, the multinomial logit (MNL) model has been the most popular discrete choice model used by marketing researchers (e.g., Fader and Hardie 1996, Gönül and Srinivasan 1993, Guadagni and Little 1983, Krishnamurthi and Raj 1991). MNL models require alternatives to have the same relative probability of being selected independent of the choice set, the so-called Independence of Irrelevant Alternatives (IIA) property. The IIA property of the MNL model has given rise to the development of nested models. Dalal and Klein (1988) give several approaches to relax the IIA assumption and the nested logit model is one

of them. Since the nested logit model is a natural generalization of the MNL model (Ben-Akiva and Lerman 1985, McFadden 1981) and it can be estimated easily, it has been commonly implemented by marketing researchers.

By adding the customer's decision on the selection of the category, Guadagni and Little (1998) extend the MNL formulation in Guadagni and Little (1983), which only models the brand choice for coffee purchase using an NL model that include both decision components (nested on categories). They demonstrate that the NL model allows a more accurate calculation of sales and a better forecasting of market response to store promotion. They use 32 weeks of coffee purchase panel data collected from four Kansas City stores, which contains information such as the date, the price, the item purchased, and the household buying for each coffee purchase. Similarly, our model is based on 6 weeks of booking and shopping data in Washington, DC, from Hotwire, which includes, for each purchase and search, the date, the property's star rating, neighborhood, and price. Because the properties located in the same neighborhood share some characteristics, our NL model is nested on the neighborhood. Guadagni and Little (1998) use a sequential estimation technique (estimate separate levels of the NL tree in sequential order from the bottom to the top), similar to Dubin (1986) and Kannan and Wright (1991). However, we implement the full maximum likelihood estimation, that is, estimating all the levels of the NL simultaneously, as it can yield statistically consistent and asymptotically efficient estimates of the parameters that sequential estimation cannot (Ben-Akiva and Lerman 1985).

Over the past decade, choice-based models have been introduced in the revenue management literature where both price and inventory decisions need to be made. Andersson (1998) using data from Scandinavian Airlines System uses a MNL model to estimate passenger preferences among buying up, being recaptured, and deviating to another airline. Talluri and van Ryzin (2004) study consumer choice behavior among multiple fare classes on a single-leg flight using a MNL model. They formulate the optimal seat allocation policy as a dynamic program (DP). This formulation is extended by Liu and van Ryzin (2008) to a network case, in which their DP problem is approximated by a deterministic linear programming problem. Similarly, Gallego et al. (2004) also use a deterministic approximation to solve a network RM problem where newly defined flexible products are chosen endogenously by customers. Zhang and Adelman (2009) also consider a network revenue management problem where consumer choice behavior is explicitly modeled by a MNL model. They claim that they provide a better approximation to the resulting DP model of Liu and van Ryzin.

Our study is similar to Talluri and van Ryzin (2004), Liu and van Ryzin (2008), and Zhang and Adelman (2009) in that we all use a discrete choice model to characterize consumer choice behavior and

formulate the revenue management problem as a dynamic programming problem. However, we solve the resulting stochastic DP problem versus assuming a priori optimal policy. This difference is due to the fact that we use price as the decision variable in the DP instead of deciding which subsets of fare classes to offer to customers. To our knowledge, this study is the first attempt in the literature to characterize customer choice behavior by using a NL model versus a MNL model.

Zhang and Cooper (2005) study a seat allocation problem when customers dynamically choose among multiple parallel flights. Similar to our model setting, they also assume that there is at most one arrival in each period and the arrivals follow a non-homogeneous Poisson process. They do not model customer choice behavior in a discrete choice framework, similar to this study, they assume that customers choose products from different inventory resources. In contrast, other studies from the literature examine customer choice between products using a common pool of resources.

More recently Vulcano et al. (2010) apply the simulation-based optimization approach used in van

Ryzin and Vulcano (2008) to estimate optimal controls where passenger demand is estimated using choice-based models approximated with the expectation-maximization algorithm owing to their use of censored data. Gallego et al. (2009) use a choice-based model in conjunction with EMSR heuristic, see Belobaba (1987), to set leg level inventory controls. Similar to Andersson (1998) and Algers and Besser (2001), Vulcano et al. (2010) perform an empirical choice model study of estimation and optimization on actual airline data from a major US airline. Both the simulated and real data in all the work above use sales transactions from one service provider. In contrast, our data contain complete market level booking and shopping records, which allows us to use the traditional maximum likelihood method (MLE) estimation. These unique data also allow us to truly and better characterize the price competition, since the competitors' prices are included in our data (not in one company's sales data) and so were captured in the choice probability resulting from the choice model. As far as we know this work is also the first choice-based revenue management research applied to the hotel setting, whereas almost all prior literature focused on airline revenue management systems.

Demand Models for Opaque Products

In the following section we briefly describe our data set and outline the application of a nested logit choice model to this data set.

Data

Our data consist of daily data on all hotel booking requests and purchase transactions made through Hotwire in Washington, DC over an arrival period of 6 weeks. As the first request was made 321 days before arrival (DBA), our data set extends beyond 1 year. Each data record consists of the date the request was made (date consumer shopped online); the check-in and check-out dates; and characteristics of the properties such as price, star rating, and neighborhood within DC. The data also contain information on whether a request led to a purchase transaction.

There are 195,226 requests in total in the data set and 7509 of them lead to a reservation. Looking at 7505 reservations (we drop 4 of the 7509 as these are the only 4 reservations for 4.5 star properties) we get a feel for what sort of properties customers are booking and where they are located within DC. Table 1 displays the percent of reservations or market share on Hotwire by hotel star level. Table 2 shows a similar market share summary by geographic subarea (neighborhood) within Washington, DC. While requests were made as early as 321 DBA, over two-thirds of reservations are made within the last 10 DBA.

Table 1 Percentage of Reservations by Hotel Star Level, Washington, DC

Star rating	Percentage of booking (%)
1	0.1
2	5.2
2.5	4.2
3	18.9
3.5	15
4	56.6

Table 2 Percentage of Reservations by Washington, DC Subarea

Neighborhood	Percentage of booking (%)
Alexandria-Old Town	12.3
Arlington	3.5
Bethesda-Silver Spring	2.4
Chantilly-Dulles Intl Airport IAD South	4.3
Crystal City-Reagan National Airport DCA	2.3
Dupont Circle-Embassy Row	45.5
Georgetown	1.8
Kennedy Center-GW University	0.6
Reston-Sterling-Dulles Intl Airport IAD North	2.5
South of the Mall	0.8
The White House-Capitol Hill	24

From Tables 1 and 2 we see each star and area receiving some fraction of reservations with stars 3, 3.5, and 4 dominating, and Dupont Circle-Embassy Row the most popular area. Figure 1 illustrates some of the price trade-offs customers are making when selecting hotel star levels. In Figure 1 the horizontal axis is a 4 star property's price relative to the 3.5 star property's prices in the same neighborhood. A relative price of 100% means that the 4 and 3.5 star property were priced the same, with 80% indicating the 4 star was 20% cheaper than the 3.5 star. The bars, the primary vertical axis, represent 4 star purchases at each price point as a percentage of all 4 star purchases. As shown in the graph, the majority of the purchases are made at properties with relative prices of 90% or less. The solid line (the secondary vertical axis) corresponds to the market share of 4 star properties at each discount. We see market share decrease with relative price increases. We note that even when the 4 star property is 30% cheaper than a 3.5 star (relative price of 70%), there are still sales made to other star qualities as market share is <1. Figure 1 indicates that while customers are making price trade-offs (price is a primary driver of the purchase decision) there must be other factors motivating purchase choice—motivating our use of choice models for estimating market share or purchase probability. An interesting observation is the existence of price inversion where hotels of higher star rating are priced cheaper than lower rated hotels. Figure 1 shows that 4 star properties tend to have very high market share under price inversion, but that drastic price inversion does not occur that often as indicated by the height of the bars.

The Nested Logit Model and Estimation

A discrete choice model is designed to explain how an individual decision maker makes a choice among a feasible set of alternatives. In this case, the decision makers are the consumers who make requests and book a hotel located in Washington, DC through Hotwire.com. We represent the consumer choice behavior using the nested logit model. Although the MNL model is widely used in modeling choice behavior, it requires alternatives in the choice set to have the IIA property. The IIA property requires that changes in the choice set do not affect the ratio of choice probabilities of any two alternatives. Under this property, all alternatives should have the same measure of similarity. The IIA property is often violated when some alternatives share attributes (are similar), but some do not (are dissimilar). The nested logit model is an extension to the MNL model which was introduced to relax the IIA assumption by grouping similar alternatives.

Figure 1 Distribution of Reservations and Market Share Star Level Price Ratios
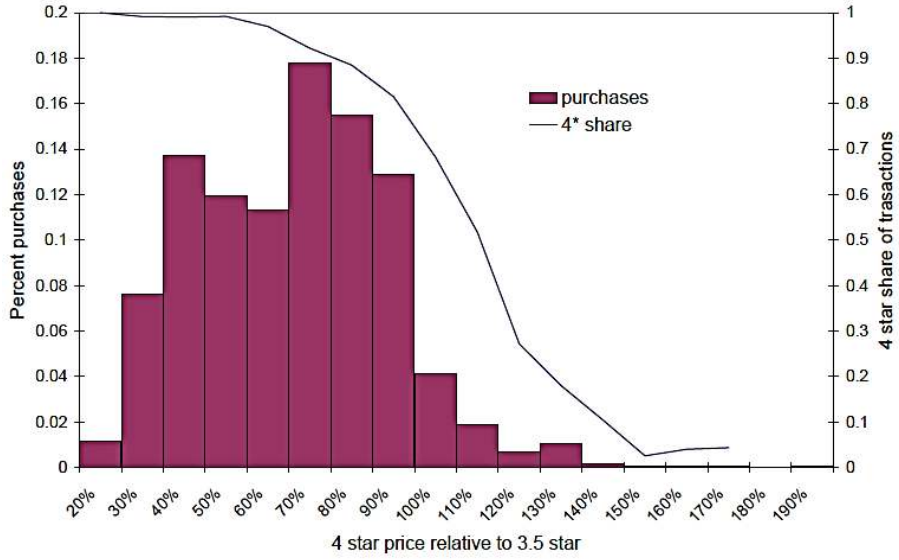
Specifically, consumer *n* is faced with a decision of booking a property at a star level *j* within a neighborhood *i* in Washington, DC. The universal choice set of alternatives denoted by C consists of all the properties identified by an (*i,j*) pair (neighborhood, star). Each consumer n has a corresponding choice set $C_n \subset C$. We partition the alternatives into 11 groups by neighborhood. In other words, the properties in each group are in the same neighborhood. We need to construct a nested logit model to avoid the violation of the IIA property as comparing a 3 star to a 4 star hotel in the same neighborhood is a different trade-off compared to choosing between a 3 star in one neighborhood over a 3 star hotel in another neighborhood. In our two-level NL model, neighborhood is the first-level feature *i* and star is the second-level feature j. The random utility that consumer *n* books a property of star quality j within neighborhood *i* is given by

(1)

$$
\begin{aligned}
U_{nij} &= \beta' \mathbf{x}_{nij} + \varepsilon_{ni} + \varepsilon_{nij} \\
&= \beta_p x_{nijp} + \beta_{s1} x_{njs1} + \beta_{s2} x_{njs2} + \beta_{s3} x_{njs3} \\
&\quad + \beta_{s4} x_{njs4} + \beta_{s5} x_{njs5} + \varepsilon_{ni} + \varepsilon_{nij},
\end{aligned}
$$

where $x_{nij} = (x_{nijp}, x_{njs1}, x_{njs2}, x_{njs3}, x_{njs4}, x_{njs5})'$ is the vector of attributes of a property at star level j in a neighborhood i. x_{nijp} represents the price of alternative (*i,j*) and x_{njsi} are indicator variables for 5 star qualities (1, 2, 2.5, 3, and 3.5 [stars] with the reference a 4 star hotel). $\beta = (\beta_p, \beta_{s1}, \beta_{s2}, \beta_{s3}, \beta_{s4}, \beta_{s5})0$ is the parameter vector that captures the consumer's preferences for those attributes and it needs to be estimated.

ε_{nij} is the random error term corresponding to the alternative $(i,j) \in C_n$, which follows a Gumbel distribution. ε_{ni} is the random error term corresponding to neighborhood i and is distributed such that $max_{j \in J_{ni}}$ is Gumbel distributed with scale parameter μ^I, commonly known as the dissimilarity coefficient. J_{ni} is defined as the set of all star qualities available for consumer n if he or she chose neighborhood i at the first level.

The probability that consumer n chooses alternative $(i,j) \in C_n$ is given by:

(2)

$$
\begin{aligned}
P_n(i,j) &= Pr[U_{ij} \geq U_{i'j'} \text{ for all } (i',j') \in C_n] \\
&= P_n(i) \cdot P_n(j|i) \\
&= \frac{e^{V_i' \mu^I}}{\sum_{i' \in I_n} e^{V_{i'}' \mu^I}} \cdot \frac{e^{\beta' \mathbf{x}_{nij}}}{\sum_{j' \in J_{ni}} e^{\beta' \mathbf{x}_{nij'}}},
\end{aligned}
$$

where,

(3)

$$
V_{i'}' = \ln \sum_{j \in J_{ni}} e^{\beta' \mathbf{x}_{nij}},
$$

is called the inclusive value and represents the scaled effect of the utilities in the second-level on the first-level's utility. I_n is the set of all neighborhoods in C_n.

We estimate the parameters β_p, β_{si} for i = 1,2…,5 and obtain the choice probabilities by applying the maximum likelihood method using NLOGIT (Greene 2007). Table 3 summarizes parameter estimates. As shown in Table 3 all the estimates of the parameters are significant and have the expected sign. The parameters for price are negative, implying that consumers prefer properties with lower prices. The parameters for star rating's indicator variables are all negative with $\beta_{s1} < \beta_{s2} < \ldots < \beta_{s5}$ indicating customers prefer higher quality hotels. The dissimilarity coefficients are all significantly different from zero and lie in the interval [0,1] consistent with random utility maximization.

Table 3 Nested Logit Parameter Estimates

Parameter	Parameter estimate	Standard error	p-Value
β_p	−0.06941	0.0016	0.0000
β_{s1}	−7.7288	0.3784	0.0000
β_{s2}	−6.2962	0.1259	0.0000
β_{s3}	−5.0331	0.1121	0.0000
β_{s4}	−3.1586	0.0704	0.0000
β_{s5}	−1.9744	0.0592	0.0000
β_{DBA}	−0.0153	0.0037	0.0000

First-level feature neighborhood	Dissimilarity coefficient		
Alexandria-Old	0.6874	0.0187	0.0000
Arlington	0.7376	0.0202	0.0000
Bethesda-Silver Spring	0.7425	0.0217	0.0000
Chantilly-Dulles Intl Airport IAD South	0.9049	0.0243	0.0000
Crystal City-Reagan National Airport DCA	0.6674	0.0183	0.0000
Dupont Circle-Embassy Row	0.4678	0.0141	0.0000
Georgetown	0.5832	0.061	0.0000
Kennedy Center-GW University	0.7361	0.0249	0.0000
Reston-Sterling-Dulles Intl Airport IAD North	0.9700	0.0264	0.0000
South of the Mall	0.6068	0.0178	0.0000
The White House-Capitol Hill	0.4463	0.0129	0.0000

Table 4 Nested Logit Model Fit Diagnostics

LL	−10979.11
LL^*	−29503.89
McFadden Pseudo R-squared	0.6278757

Table 4 summarizes model fit characteristics. The McFadden Pseudo R-squared is an indicator of the overall model fit, which is

$$1 - \frac{L\dot{L}}{LL^*} = 1 - \frac{-10979.11}{-29503.89} = 0.628,$$

where LL is the log-likelihood of the test model and LL^* is the log-likelihood for a reference model that is estimated with constant only, that is, the market shares predicted by the model are what are in the data. The NL model correctly classified 68.9% of choices with the modal category being a 4 star property in Dupont Circle (39%) of purchases resulting in a proportional reduction in classification error of

$$\frac{0.689 - 0.39}{1 - 0.39} = 49.0\%$$

In addition to modeling consumer choices, the NL model can be used to determine consumer price sensitivity or elasticity as well as the monetary value of each star rating. Table 5 summarizes the average (across the 11 neighborhoods in DC) of own and cross price elasticities. Price elasticities from the NL model are the percent change in choice probability given a percent change in price. In Table 5 the

diagonal (negative values) are the own price elasticities and the off diagonal (positive values) are the cross price elasticities. For example in the first row (1 star hotel) the choice probability decreases 1.36% for each 1% increase in the 1 star hotel's price, whereas the 2 star property's choice probability increases 0.02% with this same 1% increase. The higher star value properties tend to be more elastic with the 4 star's price having the largest impact on the other star's choice probabilities (largest cross price effects). The cells which are empty indicate that these two types of properties were never simultaneously listed, for example, a 4 star property was never displayed with a 1 star property.

Table 5 Own and Cross Price Elasticities by Star Level

Star	Star					
	1	2	2.5	3	3.5	4
1	−1.36	0.02	0.01	0.02	0.02	
2	0.13	−2.86	0.11	0.26	0.16	0.22
2.5	0.24	0.12	−4.0	0.2	0.2	0.25
3	0.41	0.75	0.92	−4.83	1.03	0.9
3.5	0.53	0.67	0.53	0.64	−4.67	0.84
4		2.4	2.43	2.05	2.4	−4.55

The choice model can also be used to quantify the financial impact of different star ratings. Focusing on the 7505 reservations in the data set, for each purchased property we can use the parameters from Table 3 to calculate the utility of the selected property. Then decrease the star rating of the selected property by one level (a 4 star becomes a 3.5 star) then we calculate the change in price needed such that this decreased quality property has the same utility as the higher quality property. Table 6 summaries these impacts, for example, if a 2 star was to change to a 1 star it would need to drop its price by $20.58 or 32% to have the same utility (and same market share or choice probability). The higher star ratings tend to have a higher marginal value.

The NL estimates the probability of making a sale assuming the firm is displayed in the choice set, in the following section we estimate a logistic regression model that estimates the probability a firm is displayed.

Probability of Display

Figure 2 shows a subset of the information a hotel (or prospective consumer) would have available about its market for any future arrival date. The figure shows a set of prices for hotels by star class and location within Washington, DC. Figure 3 shows a map of the 11 neighborhoods within DC with each of these 11 neighborhoods potentially having one property at each star level displayed similar to

those in Figure 2. Not all service providers releasing inventory to Hotwire will be displayed in Figure 2. For instance, there are 13 3.5 star properties in Dupont Circle in Washington, DC; all may release inventory to Hotwire, but given the opaque structure of Hotwire only one property is displayed, that is gets an opportunity at a sale on Hotwire. Hotwire provides a daily report to each property which can be used to determine its likelihood of being displayed in the future. A sample daily report is displayed in Figure 4. The report contains summaries on price (firm and competitors), the number of requests made, as well as the number of times the firm was displayed on that day.

Figure 4 summarizes all Hotwire activity made the previous day for all future arrival dates, with each row in the report representing a future arrival date. The data in the middle block of columns (Your Hotel Data for Today) indicate for how many consumer searches the hotel had available inventory (# of Times Avail) as well as the number of times it was displayed (# of Times Disp) and the prices it had posted to Hotwire (Net rate). For example, for an arrival on Sunday, September 30th there were 94 customer searches, of these 94 the hotel had rooms available 87 times, the hotel was displayed 75 times, and 2 reservations were made.

Table 6 Financial Value of Star Rating

Star	Price change ($)	Percentage price change (%)
2	20.58	32
2.5	18.2	30
3	27	35
3.5	17.06	18
4	28.44	30

We use this data from 6 weeks of daily reports to estimate the probability that a hotel will be displayed on as a function of price using logistic regression. Let $P_d(p(t))$ represent the probability of display for a particular arrival date posting a price $p(t)$ at t days prior to arrival, then

(4)

$$P_d(p(t)) = \frac{1}{1 + \exp(g(p(t)))},$$

where $g(p(t))$ is a linear function of $p(t)$.

We transform price into relative price, a firm's price divided by the displayed comparable price. The displayed price is the price of the currently displayed hotel of the same star level in the same neighborhood (e.g., 3 star hotel in Dupont Circle). As prices change the closer you book relative to the stay date we also include DBA, that is, days before arrival (stay date-search date) as an independent

variable as well as an indicator variable for weekend vs. weekday as weekend prices tend to be lower than weekday.

Figure 2   Sample Hotel Listing on Hotwire.com
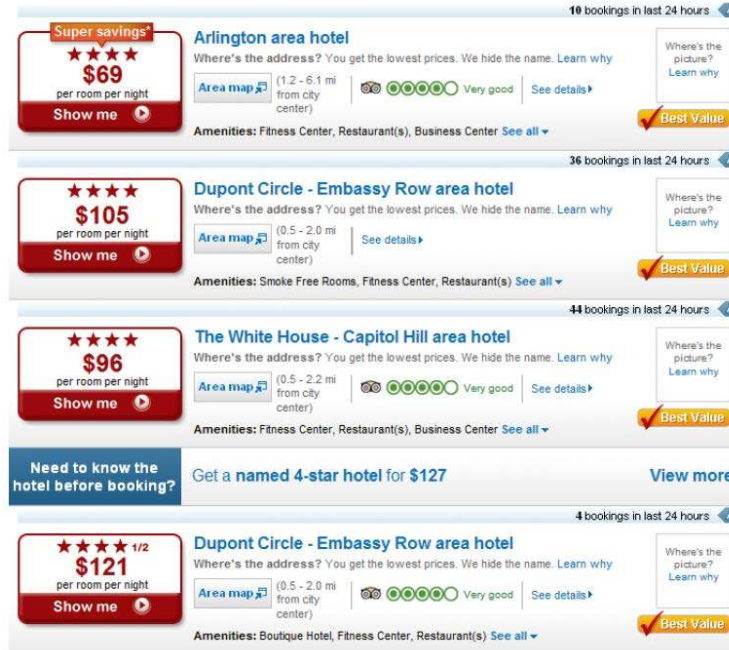
The model specification that we estimated from our data is,

$$g(p(t)) = -\alpha_0 - \alpha_1 \frac{p(t)}{p_{competitor}(t)} - \alpha_2 DBA - \alpha_3 weekend.$$

Therefore $P_d(p(t)) =$

(5)

$$\frac{1}{1 + \exp(-\alpha_0 - \alpha_1 \frac{p(t)}{p_{competitor}(t)} - \alpha_2 DBA - \alpha_3 weekend)}.$$

Table 7 summarizes parameter estimates. Price ratio is the ratio of our price to the comparable price. Weekend is an indicator, a 1 if the date we are pricing is a weekend, 0 otherwise.

Figure 5 shows a set of sample results using Equation (5) and the estimates in Table 7 for a firm posting prices for a weekend (*Weekend* = 1) arrival day (*DBA* = 0) with competitors posting a price of $150.

Fraction of Customers Buying–Making a Sale on Hotwire

Sections 3.2 and 3.3 outlined the estimated nested logit model and logistic regression models. The nested logit model determines the probability a firm makes a sale given a sale is made and the firm

is displayed. The logistic regression estimates the probability a firm is displayed. Figure 6 shows a sample graph of the average number of requests made, referred to as looks, and the percentage of these requests making a reservation (book-to-look ratio).

Figure 3   DC Area Neighborhoods on Hotwire.com

Figure 4   Sample Hotwire.com Report

Figure 6, the nested logit and logistic regression models can be used together to develop a price-dependent customer arrival process. We outline two forms of the arrival process in the following section, one assuming dynamic prices, another for daily fixed prices.

Pricing Model Development

We determine optimal prices under two separate frameworks, later comparing prices and expected revenues. In section 4.1 we assume dynamic pricing over the entire selling horizon. In section

4.2 we assume the firm changes prices each day but that prices remain fixed throughout the day, for example a hotel accepting room reservations for a stay date 3 days from today would post at most four different prices for that room (potentially a different price on each of the arrival date and each day prior to arrival). We evaluate optimal prices under these two frameworks. Dynamic pricing provides increased revenues and allows a characterization of the optimal policy. By contrast firms traditionally use daily fixed prices largely as a result of reservation systems limitations.

Table 7 Logistic Regression Coefficients for Probability of Display Model

Parameter	Coefficient estimate
Constant	0.593[*]
Price ratio	−1.684[*]
DBA	0.110[*]
Weekend	−0.892[*]

[*]Significant at the 0.001 level.

Dynamic Pricing

Customers are assumed to arrive at Hotwire following a non-homogeneous Poisson process with an arrival rate λ_t, t days prior to arrival. We subdivide each day prior to arrival into η_t decision periods of equal lengths such that there is at most one arrival request in each period. More specifically, the value of η_t should be such that the length of each decision period is small enough to have negligible probability ε of more than one arrival. The arrivals for each decision period follow a Poisson process with rate λ_t/η_t and the probability of x arrivals in a decision period m is given by

(6)
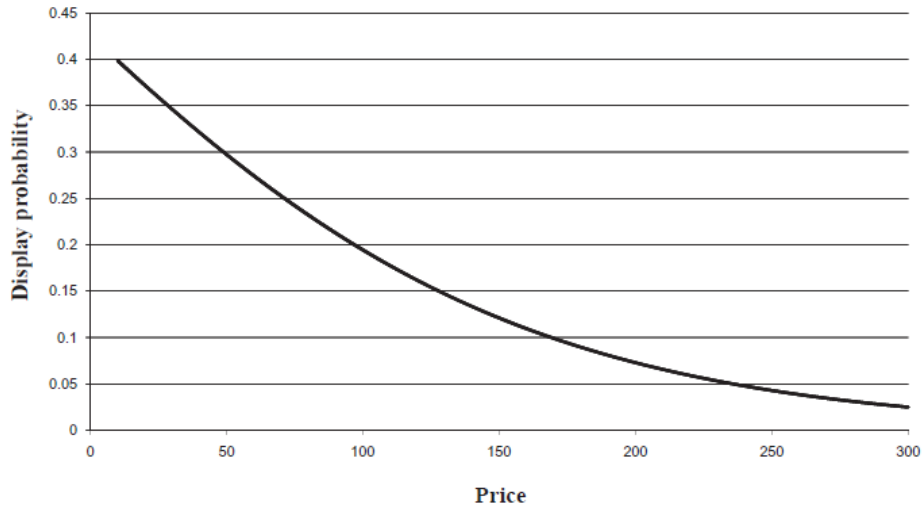
$$P^m(x) = \frac{(\lambda_t/\eta_t)^x \exp^{(\lambda_t/\eta_t)}}{x!},$$

where x = 0,1,2… and 1 - P^m(0) - P^m(1) - ε. For convenience, we denote the arrival probability of one request in decision period m by P^m as,

(7)

$$P^m = (\lambda_t/\eta_t) \exp^{(\lambda_t/\eta_t)}.$$

Figure 5   Display Probabilities from Logistic Regression for Prices Posted on Arrival Day for a Weekend Arrival with a $150 Competitor Rate

As not every arriving customer chooses to book on Hotwire, let $P_b(m)$ be the probability that an arrival makes a reservation (the book-to-look ratio in Figure 6) in decision period m. Consider a property releasing r rooms over the next m periods at price $p(m,r)$ who is displayed on Hotwire with choice probability Pc given by the nested logit model (2).

Figure 6   Average Number of Daily Request (Looks) and Percentage Making a Reservation (Book-To-Look Ratio)
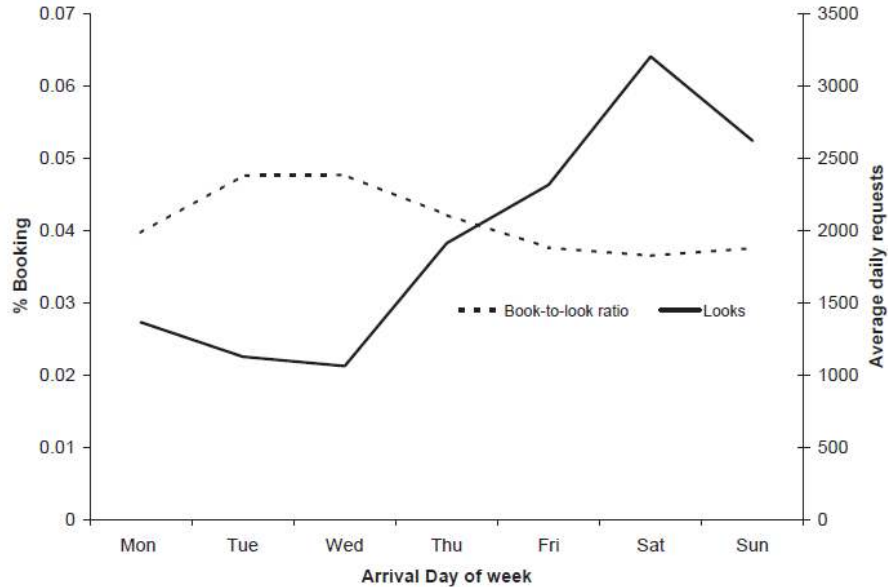
The probability, denoted P_s, that a randomly selected customer will be willing to book a room is a function of its price $p(m,r)$ and given by

(8)

$$P_s(p(m,r)) = P_d(p(m,r))P_b(m)P_c(p(m,r)).$$

The probability, $f_{p(m,r)}$, of making a sale at the price $p(m,r)$ becomes

(9)

$$\vec{f}_{p(m,r)} = P_s(p(m,r))P^m$$
$$= P_d(p(m,r))P_b(m)P_c(p(m,r))P^m,$$

where $P_d(p(m, r))$ is given by Equation (4), $P_c(p(m, r))$ is given by Equation (2), and Pm is given by Equation (7) above. We can use this price-dependent probability of making a sale to determine the optimal price to post via dynamic programming.

Let $V^*(m, r)$ denote the optimal total expected revenue, resulting from price $p^*(m, r)$, that can be generated from the remaining m periods given r remaining rooms, where m = {0,1,2…M} and r = {0,1,2…}. If a sale is made in period m, then $V^*(m, r) = p(m, r) + V^*(m-1, r - 1)$. If there is no sale in that period, then the optimal total expected revenue is given by $V^*(m - 1, r)$. $V^*(m, r)$ is characterized by the following recursion,

(10)

$$V^*(m,r) = \max_{p(m,r)}\Big\{[p(m,r) + V^*(m-1,r-1)]f_{p(m,r)}$$
$$+ V^*(m-1,r)(1-f_{p(m,r)})\Big\} = V^*(m-1,r)$$
$$+ \max_{p(m,r)}\Big\{[p(m,r) - \delta(m-1,r)]f_{p(m,r)}\Big\},$$

where $\delta(m, r) = V^*(m, r) - V^*(m, r-1)$ is the expected marginal value of a room in decision period m given r rooms. The boundary conditions are $V^*(0,\cdot) = 0$, $V^*(\cdot,0) = 0$. We next describe characteristics of the optimal pricing policy.

Optimal Dynamic Pricing Policy

The optimal dynamic pricing policy is described by Theorem 1. Lemmas 1–4 are used in the proof of Theorem 1. Theorem 1 states that the optimal pricing policy is decreasing in capacity and increasing in time. Lemmas 1, 2, and 4 are similar in structure to Lemmas 4, 5, and 3 respectively in Talluri and van Ryzin (2004). We use price as our decision variable whereas Talluri and van Ryzin (2004) characterize an optimal nested inventory allocation policy. The Corollary and Lemma 3 are unique

developments with the corollary naturally following Lemmas 1 and 2, and Lemma 4 dependent upon Lemma 3.

LEMMA 1: The expected marginal value is decreasing in the remaining capacity, that is, $\delta(m, r) \leq \delta(m, r - 1)$.

PROOF: By induction on *m*. For *m* = 0 the boundary condition V*(0,·) = 0 gives us $\delta(0,r) = \delta(0,r - 1) = 0$, the lemma is trivially true for this case.

We assume it is true for period m - 1. *p*(m,r)* denotes the optimal solution for Equation (10). From Equation (10) above, we get

(11)

$$
\begin{aligned}
V^*(m, r-1) = {}& V^*(m-1, r-1) \\
& + \max_{p(m,r-1)} \Big\{ [p(m, r-1) \\
& - \delta(m-1, r-1)] f_{p(m,r-1)} \Big\}.
\end{aligned}
$$

Then Equations (10)–(11) gives

(12)

$$
\begin{aligned}
\delta(m,r) - \delta(m-1,r) = {}& \max_{p(m,r)} \Big\{ [p(m,r) - \delta(m-1,r)] f_{p(m,r)} \Big\} \\
& - \max_{p(m,r-1)} \Big\{ [p(m,r-1) \\
& - \delta(m-1,r-1)] f_{p(m,r-1)} \Big\} \\
= {}& [p^*(m,r) - \delta(m-1,r)] f_{p^*(m,r)} \\
& - [p^*(m,r-1) \\
& - \delta(m-1,r-1)] f_{p^*(m,r-1)}.
\end{aligned}
$$

Similarly Equation (12)–(13)

(14)

$$
\begin{aligned}
\delta(m,r) - \delta(m,r-1) = {} & \delta(m-1,r) - \delta(m-1,r-1) \\
& + [p^*(m,r) - \delta(m-1,r)] f_{p^*(m,r)} \\
& - [p^*(m,r-1) - \delta(m-1,r-1)] \\
& \times f_{p^*(m,r-1)} - [p^*(m,r-1) \\
& - \delta(m-1,r-1)] f_{p^*(m,r-1)} \\
& + [p^*(m,r-2) \\
& - \delta(m-1,r-2)] f_{p^*(m,r-2)}.
\end{aligned}
$$

The optimality of the price *p*(m,·)* implies the following inequalities:

(15)

$$
\begin{aligned}
& [p^*(m,r-1) - \delta(m-1,r-1)] f_{p^*(m,r-1)} \\
& \geq [p^*(m,r) - \delta(m-1,r-1)] f_{p^*(m,r)}
\end{aligned}
$$

(16)

$$
\begin{aligned}
& [p^*(m,r-1) - \delta(m-1,r-1)] f_{p^*(m,r-1)} \\
& \geq [p^*(m,r-2) - \delta(m-1,r-1)] f_{p^*(m,r-2)}.
\end{aligned}
$$

Substituting Equations (15) and (16) into Equation (14) we obtain

(17)

$$
\begin{aligned}
& \delta(m,r) - \delta(m,r-1) \\
& \leq \delta(m-1,r) - \delta(m-1,r-1) \\
& \quad + [p^*(m,r) - \delta(m-1,r)] f_{p^*(m,r)} \\
& \quad - [p^*(m,r) - \delta(m-1,r-1)] f_{p^*(m,r)} \\
& \quad - [p^*(m,r-2) - \delta(m-1,r-1)] f_{p^*(m,r-2)} \\
& \quad + [p^*(m,r-2) - \delta(m-1,r-2)] f_{p^*(m,r-2)},
\end{aligned}
$$

simplifying,

(18)

$$
\begin{aligned}
& \delta(m,r) - \delta(m,r-1) \\
& \leq [\delta(m-1,r) - \delta(m-1,r-1)][1 - f_{p^*(m,r)}] \\
& \quad + [\delta(m-1,r-1) - \delta(m-1,r-2)] f_{p^*(m,r-2)}.
\end{aligned}
$$

By induction, we know that $\delta(m - 1,r) - \delta(m - 1, r - 1) \leq 0$ and $\delta(m - 1,r - 1) - \delta(m - 1,r - 2) \leq 0$. Therefore, we can conclude that $\delta(m,r) - \delta(m - 1,r) \leq 0$.

LEMMA 2: The marginal value is increasing in remaining time, that is, $\delta(m,r) \geq \delta(m - 1,r)$.

PROOF: Recall from Equation (12), we have

(19)

$$\delta(m,r) - \delta(m-1,r)$$
$$= \max_{p(m,r)}\left\{[p(m,r) - \delta(m-1,r)]f_{p(m,r)}\right\}$$
$$- \max_{p(m,r-1)}\left\{[p(m,r-1) - \delta(m-1,r-1)]f_{p(m,r-1)}\right\}.$$

As proved in Lemma 1, $\delta(m - 1,r) \leq \delta(m - 1, r - 1)$. Therefore, the following inequality is always true for any value of the price p(m,r) for all m and r,

$$[p(\cdot,\cdot) - \delta(m - 1, r)]f_{p(\cdot,\cdot)} \geq [p(\cdot,\cdot) - \delta(m - 1, r - 1)]f_{p(\cdot,\cdot)}.$$

Thus,

(20)

$$\max_{p(m,r)}\left\{[p(m,r) - \delta(m-1,r)]f_{p(m,r)}\right\}$$
$$\geq \max_{p(m,r-1)}\left\{[p(m,r-1) - \delta(m-1,r-1)]f_{p(m,r-1)}\right\}$$

which implies $\delta(m,r) \geq \delta(m - 1,r)$.

COROLLARY 1: The optimal total expected revenue V*(m,r) is increasing and concave in both the number of remaining periods m and rooms r.

PROOF: V*(m,r) increases with m and r is obvious and straightforward, so we only focus on the concavity of the expected revenue. From Lemma 1 we know $\delta(m,r) \leq \delta(m,r-1)$, which implies

$$V^*(m, r) - V^*(m, r - 1) \leq V^*(m, r - 1) - V^*(m, r - 2).$$

Hence, V*(*m,r*) is concave in *m*.

Also,

$$V^*(m, r) - V^*(m - 1, r) = [p^*(m, r) - \delta(m - 1, r)]f_{p^*(m,r)},$$

and

$$V^*(m-1,r) - V^*(m-2,r) = [p^*(m-1,r) - \delta(m-2,r)]f_{p^*(m-1,r)},$$

and Lemma 2 gives that δ(m - 1,r) ≥ δ(m - 2,r). Therefore, using the same argument as for proving Equation (20) we obtain

$$V^*(m,r) - V^*(m-1,r) \leq V^*(m-1,r) - V^*(m-2,r).$$

This means V*(m,r) is concave in r as well. Lemma 4 shows that the optimal price in this period is positively related to the marginal value of the previous period. The following lemma is required in the development of Lemma 4.

LEMMA 3: Assume the property's display probability given by Equation (5) and choice probability given by Equation (2) are decreasing in price p(m,r). Then, the probability of making a sale $f_{p(m,r)}$ also decreases in price p(m,r).

PROOF. From the expression of $f_{p(m,r)}$ in Equation (9), we have

(21)

$$f_{p(m,r)} = P_d(p(m,r))P_b(m)P_c(p(m,r))P^m.$$

It is straightforward to see that given $P_b(m)$ and P^m do not depend on price *p(m,r)* as well as both $P_d(p(m,r))$ and $P_c(p(m,r))$ are decreasing in *p(m,r)*, one can conclude that $f_{p(m,r)}$ also decreases in price p(m,r).

LEMMA 4: The optimal price p*(m,r) increases with the marginal value δ(m - 1,r).

PROOF: We prove this statement by using contradiction. For ease of notation, set δ = δ(m - 1,r) and p* = p*(m,r). Assume δ_1 and δ_2 are two marginal values that satisfy $\delta_1 > \delta_2 \geq 0$. P*$_1$ and p*$_2$ are the optimal prices that maximize the term [p(m,r) δ(m -1,r)]$f_{p(m,r)}$ in Equation (10) when δ(m - 1,r) equals δ_1 and δ_2 respectively. $f_{p_i^*}$, i = 1; 2 is the corresponding probability of making a sale at optimal prices p_i^*.

Now suppose $p_1^* < p_2^*$, we attempt to get a contradiction later.

From Lemma 3, if $p_1^* < p_2^*$ then $f_{p_1^*} > f_{p_2^*}$.

On the other hand, the optimality of p_1^* implies

$$p_1^* f_{p_1^*} - \delta_1 f_{p_1^*} \geq p_2^* f_{p_2^*} - \delta_1 f_{p_2^*},$$

then

(22)

$$p_1^* f_{p_1^*} - p_2^* f_{p_2^*} \geq [f_{p_1^*} - f_{p_2^*}]\delta_1 \geq [f_{p_1^*} - f_{p_2^*}]\delta_2.$$

Then $p_1^* f_{p_1^*} - \delta_2 f_{p_1^*} \geq p_2^* f_{p_2^*} - \delta_2 f_{p_2^*}$. This inequality contradicts the optimality of p_2^*. Therefore, we must have $p_1^* < p_2^*$ if $\delta_1 > \delta_2$.

By combining Lemmas 1, 2, and 4, we obtain the next theorem that characterizes the optimal pricing policy.

THEOREM 1: Consider a property that releases r rooms over the next m periods at price p(m,r). Assume the property's display probability given by Equation (5) and choice probability given by Equation (2) are decreasing in price p(m,r). Then, the optimal pricing policy for Equation (10), p*(m,r), is decreasing in the remaining number of rooms r and increasing in the remaining selling periods m.

We next show an illustration using our data for which both $P_d(p(m,r))$ and $P_c(p(m,r))$ are decreasing in p(m,r). For ease of notation, we simply denote p(m,r) by p.

Recall from Equation (5) that the display probability $P_d(p)$ is given by

$$P_d(p) = \frac{1}{1 + \exp(-\alpha_0 - \alpha_1 \frac{p}{p_{competitor}} - \alpha_2 DBA - \alpha_3 weekend)}$$

Since the estimate for the price ratio's coefficient α_1 is negative, it is straightforward to see that the display probability $P_d(p)$ decreases in price p.

Recall from expression (2), the choice probability $P_n(i,j)$ for property (i,j), that is, P_c, is

$$P_c := P_n(i,j) = \frac{e^{V_i' \mu^I}}{\sum_{i' \in I_n} e^{V_{i'}' \mu^I}} \cdot \frac{e^{\beta' x_{nij}}}{\sum_{j' \in J_{ni}} e^{\beta' x_{nij'}}},$$

where,

$$V_{i'}' = \ln \sum_{j \in J_{ni}} e^{\beta' x_{nij}},$$

$$x_{nij} = (x_{nijp}, x_{njs1}, x_{njs2}, x_{njs3}, x_{njs4}, x_{njs5})'.$$

For simplicity, we now define

$$E = e^{\beta' x_{nij}}; C = \sum_{j' \neq j; j' \in J_{ni}} e^{\beta' x_{nij'}}; H = e^{V_i' \mu^I}; D = \sum_{i' \neq i; i' \in I_n} e^{V_{i'}' \mu^I}.$$

Hence,

$$P_c = \frac{E}{E+C} \cdot \frac{H}{H+D}.$$

Take the derivative of the terms defined above with respect to the chosen property's price variable x_{nijp} and get,

$$\frac{\partial E}{\partial x_{nijp}} = E\beta_p; \quad \frac{\partial C}{\partial x_{nijp}} = 0; \quad \frac{\partial H}{\partial x_{nijp}} = \frac{HE\mu^I}{E+C}\beta_p;$$

$$\frac{\partial D}{\partial x_{nijp}} = 0.$$

Following simplification, we further obtain the derivative of P_c with respect to x_{nijp} is

$$\frac{\partial P_c}{\partial x_{nijp}} = \left[\frac{HE}{(E+C)^2(H+D)}(C + \frac{DE\mu^I}{H+D}) \right] \beta_p$$

As shown in Table 3, β_p is negative and the terms in the brackets are all positive. Negative first derivative with respect to price implies that the choice probability decreases with prices. Therefore, we conclude the assumption in Theorem 1 that the property's display probability and choice probability are decreasing in price p(m,r) is satisfied based on our data set.

We use Theorem 1 in the numerical implementation of Equation (10) as it reduces the set of prices that need to be searched over in the determination of p*. We illustrate a numerical example in subsequent sections. In the following section we restrict daily prices to be fixed vs. variable.

4.2. Daily Fixed Pricing

Similar to dynamic pricing we use Equation (8) to define the probability that a randomly selected customer will purchase a room, $P_s := P_s(p(t,r)) = P_d(p(t,r)) \, P_b(t)P_c(p(t,r))$. We use t to represent days before arrival, with m used to define periods in the dynamic pricing formulation. If we assume N customers arrive over day t, then the number of people who will make a purchase at a firm with posted price p(t,r) follows a binomial probability distribution. We denote this number by a random variable $D_n(p(t,r))$ and, for $0 \le x \le n$,

(23)

$$P(D_n(p(t,r)) = x) = \binom{n}{x} P_s(p(t,r))^x (1 - P_s(p(t,r)))^{n-x}.$$

As the number of arrivals n is unknown with probability P(N = n), the probability, $f_{p(t,r)}$, of the number of people willing to buy at the price p(t,r) equal to x becomes

(24)

$$f_{p(t,r)}(x) = \sum_{n=0}^{\infty} P(D_n(p(t,r)) = x)P(N = n).$$

As with dynamic pricing we assume customer arrivals follow a Poisson process then,

(25)

$$f_{p(t,r)}(x) = \sum_{n=0}^{\infty} P(D_n(p(t,r)) = x)P(N = n)$$

$$= \sum_{n=0}^{\infty} \binom{n}{x} P_s^x (1 - P_s)^{n-x} \frac{\lambda_t^n e^{-\lambda_t}}{n!}$$

$$= \sum_{n=x}^{\infty} \binom{n}{x} P_s^x (1 - P_s)^{n-x} \frac{\lambda_t^n e^{-\lambda_t}}{n!}$$

$$= \sum_{k=0}^{\infty} \frac{(k+x)!}{k!x!} P_s^x (1 - P_s)^k \frac{\lambda_t^{(k+x)} e^{-\lambda_t}}{(k+x)!}$$

$$= \frac{\lambda_t^x e^{-\lambda_t}}{x!} P_s^x \Sigma_{k=0}^{\infty} (1 - P_s)^k \frac{\lambda_t^k}{k!}$$

$$= \frac{\lambda_t^x e^{-\lambda_t}}{x!} P_s^x e^{\lambda_t(1-P_s)} = \frac{(\lambda_t P_s)^x e^{-\lambda_t P_s}}{x!}.$$

With daily fixed prices a property displayed releases r rooms over the next T days, t = 0,1,2,…,T. The firm changes its daily price with the remaining capacity in order to maximize the total expected revenue. Define V*(t,r) as the maximum achievable expected revenue with r rooms and t days remaining till the arrival day. Then V*(t,r) is characterized by the following recursion,

(26)

$$V^*(t, r) = \max_{p(t,r)} \left\{ \sum_{x=0}^{\infty} [p(t, r) \min (r, x) \right.$$

$$\left. + V^*(t - 1, (r - x)^+)] f_{p(t,r)}(x) \right\}.$$

The boundary conditions are V*(0,·) = 0,V*(·,0) = 0.

We can define the expected marginal value of a room at day t as the following

$$\delta_x(t,r) = \frac{1}{x}[(t,r) - V * (t,r - x)].$$

On each day, the property posts a price p(t,r) and sells x, up to r rooms. The following day the hotel updates its price, selling up to r - x rooms, until such a time as it has no rooms left to sell for that arrival day, V*(·,0) = 0, or the day has passed and the rooms are valueless, V*(0,·) = 0.

Unlike dynamic pricing we cannot characterize all the properties of daily fixed pricing, but the following theorem greatly reduces the search space for finding optimal prices. The theorem is similar in nature to Theorem 3 from Lee and Hersh (1993). Lee and Hersh (1993) develop optimal booking limits in the presence of batch or group bookings. Lee and Hersh do not use price as a decision variable, but the sale of multiple items over a single decision period (batch bookings) is analogous to our daily fixed pricing policy.

THEOREM 2: For fixed x and r, $\delta_x(t,r)$ increases in t.

PROOF: We first show it is true for the case that x = 1, that is, $\delta_1(t,r) \geq \delta_1(t - 1,r)$.

(27)

$$
\begin{aligned}
V^*(t,r) - V^*(t-1,r) &= [V^*(t,r-1) + v^t] \\
&\quad - [V^*(t-1,r-1) + v^{t-1}] \\
&= [V^*(t,r-1) - V^*(t-1,r-1)] \\
&\quad + [v^t - v^{t-1}],
\end{aligned}
$$

where v^t denotes the expected revenue that can be generated by selling one room over the remaining t days. It is obvious that the expected revenue that a room can generate in t days is no less than the expected revenue that a room can generate in t - 1 days. That means $v^t - v^{t-1} \geq 0$. Therefore, it follows from Equation (27) that

$$
V^*(t,r) - V^*(t-1,r) \geq V^*(t,r-1) - V^*(t-1,r-1),
$$

then,

$$
V^*(t,r) - V^*(t,r-1) \geq V^*(t-1,r) - V^*(t-1,r-1),
$$

resulting in

$$
\delta_1(t,r) \geq \delta_1(t-1,r).
$$

Then, when x > 1, we can write

(28)

$$V^*(t,r) - V^*(t,r-x) = [V^*(t,r) - V^*(t,r-1)]$$
$$+ [V^*(t,r-1) - V^*(t,r-2)]$$
$$+ \cdots + [V^*(t,r-x+1)$$
$$- V^*(t,r-x)].$$

From Equation (27) above we note that

(29)

$$\delta_1(t,r) \geq \delta_1(t-1,r) \Longrightarrow$$
$$V^*(t,r) - V^*(t-1,r) \geq V^*(t,r-1) - V^*(t-1,r-1) \Longrightarrow$$
$$V^*(t,r) - V^*(t,r-1) \geq V^*(t-1,r) - V^*(t-1,r-1)$$
$$\text{for any } r > 1.$$

Thus, combining Equations (28) and (29) gives

$$V^*(t,r) - V^*(t,r-x) \geq V^*(t-1,r) - V^*(t-1,r-x),$$

that is $\delta_x(t,r) \geq \delta_x(t-1,r)$ for any t given the fixed value of x and r.

Note that given the positive relationship between the optimal price and the marginal value, we can characterize that the optimal pricing policy is also increasing in the remaining selling time.

Numerical Example

We next present a numerical example to illustrate the application of our choice-based dynamic programming model under both daily fixed pricing and dynamic pricing as well as the revenue gains from moving to dynamic pricing policies. In this example, we consider a 3 star hotel in Chantilly DC planning to sell five rooms over the last 3 days prior to a Monday arrival.

Table 8 summarizes the average number of requests received by day before arrival, the book-look ratios, and the number of subperiods for $\varepsilon = 0.05$ (probability of more than 1 request in a subperiod ≤ 0.05).

We use the parameters from Table 3 to build the nested logit model for estimating the probability of sale P_c given display with parameters from Table 7 used for the logistic regression to estimate the probability of being displayed P_d. The book-to-look ratios in Table 8 provide estimates of P_b. For the dynamic pricing model of section 4.1 the average requests (from Table 8) divided by the number

of subperiods provides the arrival rate for arrivals. The average requests provides the rate for the fixed daily prices of section 4.2.

Table 9 displays optimal daily fixed prices. For example, the firm would post a price of $56 if it had five rooms to sell 3 days prior to arrival. If two reservations were accepted on the third day, they would increase prices to $62 with 2 days left (and three rooms remaining). Consistent with the properties of dynamic pricing, Table 9 shows that prices are decreasing with increasing capacity and with decreasing time.

Figure 7 displays the same fixed prices as in Table 9 for the firm setting prices on the arrival date for one to five rooms with series labels F-1, F-2, F-3, F-4, and F-5 (F for fixed - number of rooms). The figure also displays optimal variable prices (in dollar price steps) throughout the arrival day, DBA 0 (period 958 the start of the day, 0 the end). If we look at the series for 1 room (F-1 for fixed and D-1 for dynamic) we see the dynamic price starts higher than the fixed decreasing gradually below the fixed price. For situations with more than one room the prices behave differently, with the dynamic price always less than the fixed.

Table 8 Model Parameters—Monday Arrival 3 Star Hotel in Washington, DC

	Days before arrival			
	0	1	2	3
Book-to-look ratio	0.041	0.039	0.040	0.037
Average requests	341	385	233	197
No. of subperiods	958	1089	656	554

Table 9 Optimal Daily Fixed Prices

	Days before arrival			
Rooms	0	1	2	3
1	55	72	78	82
2	49	64	69	72
3	44	58	62	65
4	41	53	57	60
5	38	49	53	56

Figure 8 shows the fixed price for the arrival day with five rooms to sell (F-5). The figure also displays five separate series for the dynamic pricing. Series D-5 assumes dynamic pricing at the start of

the day with five rooms with D-4 assuming a room is reserved at the start of subperiod 750. The remaining series assume another reservation is accepted every 200 subperiods. Unlike daily fixed pricing, dynamic pricing sets prices knowing the firm will react to future sales, potentially raising prices as capacity is decreased.

Dynamic pricing offers increased revenue opportunities. As an estimate of these revenue gains we simply evaluate Equation (10) using optimal fixed daily prices vs. optimal dynamic prices resulting in revenue gains of 5.0% from dynamic pricing.

Summary

Incorporating customer choice behavior into revenue management has been an active research area over the past several years. In this study, we estimate a nested logit model on data from firms selling hotel rooms at Hotwire.com to understand the customers' choice preferences on hotels located in Washington, DC. Our study is unique in that it is the first choice-based revenue management model that uses market level demand data (vs. a single firm's sales data) to estimate the underlying demand models.

We model demand as a function of a property's prices using the result from the nested logit model and develop a choice-based dynamic programming model with pricing as the decision variable. We propose two stochastic dynamic programming formulations, one with daily fixed prices and one with fully dynamic prices. We provide a complete characterization of optimal dynamic prices and a partial characterization for optimal daily fixed prices. We show numerically that optimal fixed prices are consistent with the properties of dynamic prices. We then use these two characterizations to estimate revenue impacts from daily fixed vs. fully dynamic pricing policies.

Currently our formulation considers the firm selling capacity at the opaque reseller, ignoring the fact that they may simultaneously release inventory across numerous channels—both opaque like Hotwire (e.g., Priceline's Name-Your-Own-Price) and fully transparent (e.g., Expedia or Marriott.com). While our formulation can be extended to incorporate the opportunity cost of releasing rooms on other channels by including dual prices from standard revenue management models as in Anderson (2009), it currently does not directly include consumer channel choice behavior. An interesting direction for future research is to consider consumer choice behavior both across service providers as well as channel selection.

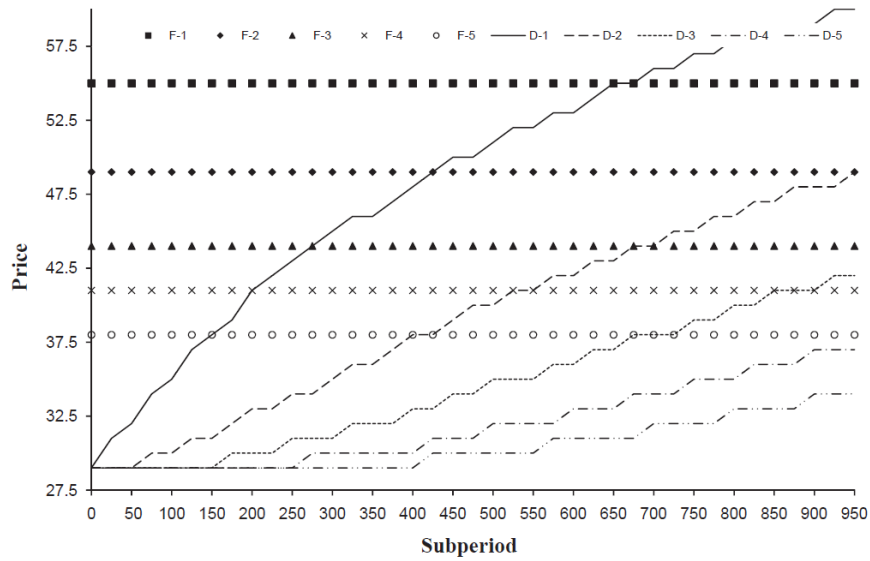**Figure 7   Comparing Daily Fixed Prices Vs. Dynamic Prices for DBA=0 and Five Rooms**
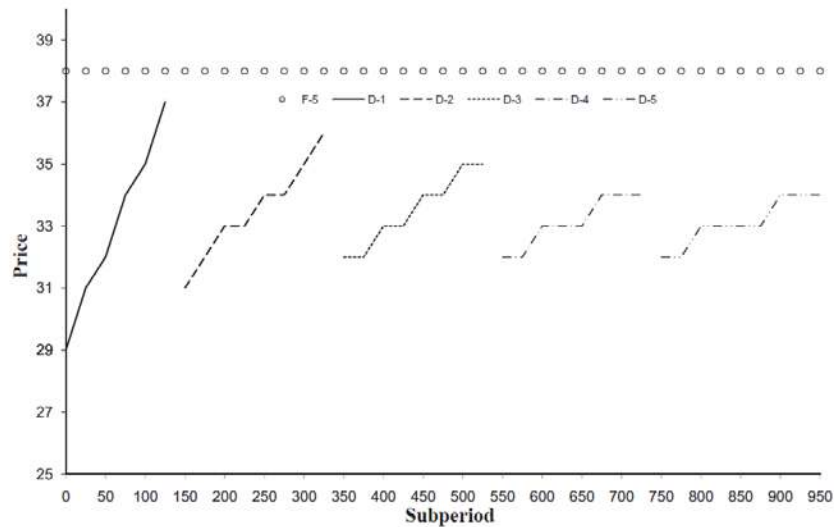

Figure 8   Daily Fixed Prices Vs. Dynamic Prices for DBA=0 and Five Rooms

References

Algers, S., M. Besser. 2001. Modeling choice of flight and booking class: A study using stated preference and revealed preference data. *Int. J. Serv. Technol. Manage.* 2(1): 28–45.

Anderson, C. K. 2009. Setting prices on priceline. *Interfaces* 39(4): 307–315.

Andersson, S. E. 1998. Passenger choice analysis for seat capacity control: A pilot project in Scandinavian Airlines. *Int. Trans. Oper. Res*. 5(6): 471–486.

Belobaba, P. P. 1987. Air travel demand and airline seat inventory management. Dissertation, Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Ben-Akiva, M., S. R. Lerman. 1985. *Discrete Choice Models: Theory and Application to Travel Demand.* The MIT Press, Cambridge, MA.

Coretjens, M. L., D. A. Gautschi. 1983. Formal choice models in marketing. *Mark. Sci.* 2(1): 19–56.

Dalal, S. R., R. W. Klein. 1988. A flexible class of discrete choice models. *Mark. Sci.* 7(3): 232–251.

Dubin, J. A. 1986. A nested logit model of space and water heat system choice. *Mark. Sci.* 5(2): 112–124.

Fader, P. S., B. G. S. Hardie. 1996. Modeling consumer choice among SKUs. J. *Marketing Res.* 33(4): 442–452.

Fay, S. 2004. Partial repeat bidding in the name-your-own-price channel. *Mark. Sci.* 23(3): 407–418.

Fay, S. 2008. Selling an opaque product through an intermediary: The case of disguising one's product. *J. Retailing* 84(1): 59–75.

Gallego, G., G. Iyengar, R. Phillips, A. Dubey. 2004. Managing flexible products on a network. *CORC Technical Report TR- 2004-01*, Department of Industrial Engineering and Operations Research, Columbia University, New York.

Gallego, G., L. Li, R. Ratliff. 2009. Choice-based EMSR methods for single-leg revenue management with demand dependencies. *J. Rev. Pricing Manage.* 8(3): 207–240.

Gönül, F., K. Srinivasan. 1993. Modeling multiple sources of heterogeneity in multinomial logit models: Methodological and managerial issues. *Mark. Sci.* 12(3): 213–229.

Greene, W. H. 2007. *LIMDEP, Version 7.0: User's Manual*. Econometric Software, INC, Bellport, NY.

Guadagni, P. M., J. D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci*. 2(3): 203–238.

Guadagni, P. M., J. D. C. Little. 1998. When and what to buy: A nested logit model of coffee purchase. J. Forecasting 27(3): 303–326.

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

Jiang, Y. 2007. Price discrimination with opaque products. *J. Rev. Pricing Manag*. 6(2): 118–134.

Jerath, K., S. Netessine, S. K. Veeraraghavan. 2010. Revenue management with strategic customers: Last-minute selling and opaque selling. *Manage. Sci.* 56(3): 430–448.

Kannan, P. K., G. P. Wright. 1991. Modeling and testing structured markets: A nested logit approach. *Mark. Sci.* 10(1): 58– 82.

Krishnamurthi, L., S. P. Raj. 1991. An empirical analysis of the relationship between brand loyalty and consumer price elasticity. *Mark. Sci.* 10(2): 172–183.

Lee, T. C., M. Hersh. 1993. A model for dynamic airline seat inventory control with multiple seat bookings. *Transportation Sci.* 27(3): 252–263.

Liu, Q., G. J. van Ryzin. 2008. On the choice-based linear programming models for network revenue management. *Manuf. Serv. Oper. Manage.* 10(2): 288–310.

McFadden, D. 1981. Econometric models of probabilistic choice. C. Manski, D. McFadden, eds. *Structural Analysis of Discrete Data*, MIT Press, Cambridge, MA.

Shapiro, D., X. Shi. 2008. Market segmentation: The role of opaque travel agencies. *J. Econ. Manage. Strategy* 17(4): 803–837.

Talluri, K. T., G. J. van Ryzin. 2004. Revenue management under a general discrete choice model of consumer behavior. *Manage. Sci.* 50(1): 15–33.

Terwiesch, C., S. Savin, I. H. Hann. 2005. Online haggling at a name-your-own-price retailer: Theory and application. *Manage. Sci*. 51(3): 339–351.

van Ryzin, G. J., G. Vulcano. 2008. Computing virtual nesting controls for network revenue management under customer choice behavior. *Manuf. Serv. Oper. Manage.* 10(3): 448–467.

Vulcano, G., G. J. van Ryzin, W. Chaar. 2010. Choice-based revenue management: An empirical study of estimation and optimization. *Manuf. Serv. Oper. Manage.* 12(3): 371–392.

Wang, T., E. Gal-Or, R. Chatterjee. 2009. The name-your-own-price channel in the travel industry: An analytical exploration. *Manage. Sci.* 55(6): 968–979.

Zhang, D., D. Adelman. 2009. An approximate dynamic programming approach to network revenue management with customer choice. *Transportation Sci.* 43(3): 381–394.

Zhang, D., W. L. Cooper. 2005. Revenue management for parallel flights with customer choice behavior. *Oper. Res.* 53(2): 415– 431.