MIT Sloan Sports Analytics Conference 2011
March 4-5, 2011, Boston, MA, USA

# Buying tickets: Capturing the dynamic factors that drive consumer purchase decisions for sporting events

Wendy W. Moe
University of Maryland
College Park, MD 20742
Email:
wmoe@rhsmith.umd.edu

Peter S. Fader
University of Pennsylvania
Philadelphia, PA 19104
Email:
faderp@wharton.upenn.edu

Barry Kahn
Qcue, Inc.
Austin, TX 78
Email: bk@qcue.net

## Abstract

Though sports teams have a general intuition about the factors that influence ticket sales, little is understood about the decision-making process that underlies consumers' consideration/purchase activities. We develop a two-stage model to better understand this process. In the first stage, consumers are faced with a "universal" set of options (i.e., all game/ticket-tier combinations) from which they construct a smaller consideration set. In the second stage, consumers choose one option from this reduced set. We consider a variety of factors in each stage of the model, such as: game attractiveness (which is allowed to vary over time as the strength of each opponent changes throughout the season), seating tier, days until game, and ticket prices. We estimate our model for a U.S. professional sports franchise, and our empirical results can allow teams to run plausible scenarios about the impact that price changes (for current and future games) will have on ticket sales.

## Introduction

When sports fans consider buying tickets to see their local team in action, several factors come into play. How well is the home team playing? How good is the opponent they are facing? When is the game? How good are the seats and how much do they cost? In this research, our objective is to examine how these factors influence ticket sales, but we want to do so in a psychologically plausible manner, enabling us to run "what if" scenarios that can accurately capture possible dynamics in the marketplace. We do so by considering the consumer's purchasing decision as a two-stage process.

Previous research has established that, in many different choice settings, consumers reduce the complexity of purchasing decisions by separating the decision process into two stages: a consideration stage and a choice stage [1][2][3][4]. The goal in the consideration stage is to identify a reduced set of options from the universal set available in order to reduce the cognitive effort needed to evaluate each alternative. Because the number of options available in the universal set can be quite large, psychologists have shown that individuals tend to use simplified rules (or heuristics) in this stage of the process [5]. One popular heuristic is the "Elimination by Aspects" (EBA) rule, where options that fail to meet minimum thresholds along one or more dimensions are eliminated from consideration. In the context of ticket purchasing, the EBA stage makes it easy for the consumer to whittle down the large set of game/ticket-tier combinations to a much more manageable set that deserves closer scrutiny.

The second stage is generally more "cognitively expensive" for the consumer as he carefully evaluates each of the options that has successfully passed the screening process in the consideration stage. The option that provides the highest value to the consumer is then chosen. Researchers have typically represented this stage of the decision as a multinomial logit choice model which stochastically identifies the highest-value option based on a given set of attributes that may vary over time [6][7]. A "no choice" option is also included, giving the decision-maker the chance to turn away all of the options (despite that the fact that some may have passed the first-stage screening rule).

For sports fans, buying tickets to see the home team play can be a complex purchasing decision that aligns well with this two-stage decision process described above. At any given point in the season, fans are choosing from a large (but constantly decreasing) number of future games as well as a variety of seating options available for each game, resulting in hundreds of choice alternatives to consider. We develop a two-stage model of ticket sales where only a limited set of game-seating options are considered for purchase at any point in time, but the composition of this set will vary from day to day, depending on the performance of the home team and each of its upcoming opponents. In the consideration stage, alternatives are eliminated based on the attractiveness of the game itself, the inherent value of the seating tier and the number of days until the game is played. This results in a smaller and more manageable set of options to evaluate.

In the second-stage choice model, we construct a value function for each remaining game-seat option that estimates the underlying value of the game separately from the value of the seating tier, while also considering time-dynamics in game attractiveness by incorporating win-loss records of the competing teams. Specifically, we model the attractiveness of a specific game in the schedule to be a function of the visiting team's win-loss record for the season thus far. Consequently, as the season progresses, the attractiveness of any given game changes. We also allow the attractiveness of all games to shift as a function of the home team's record. We do this by allowing the attractiveness of the "no-choice" option to vary as a function of the home team's win-loss record. Therefore, as the home team wins, the no-choice option becomes less appealing and ticket sales for all games increase. The inclusion of win-loss records as covariates is an important aspect of the model as it allows us to quantify the value of a winning record on ticket sales.

The evaluation of seating tiers can also be a complex process. Ticket buyers face a tradeoff between the inherent value of a particular seating location versus the price of the ticket. Our model allows us to disentangle these two effects, thereby letting us understand the separate effects of pricing, per se, from the value of each seating tier. This will allow us to examine how ticket prices influence both the decisions of whether a ticket is considered at all and, if so, whether it is ultimately chosen from the consideration set. This is a key benefit since it enables the examination of different pricing scenarios using the model's estimated parameters.

Finally, we recognize that the choice process is not only fairly complex for a given ticket buyer, but also that the nature of this complexity may vary from fan to fan. Thus we allow for differences across consumers by incorporating multiple segments, each of which is characterized by a unique EBA screening rule. So, for instance, some consumers may only consider high-value games, while others may be more interested in obtaining better seats. Capturing heterogeneity using latent segments in this manner will not only improve the statistical performance of the model but will also provide useful diagnostics about how different kinds of fans will react differently to pricing, game attractiveness, and other factors.

We estimate our model on home-game ticket sales for an undisclosed NHL team over the 2008-09 season. The results suggest that ticket buyers do in fact use simplifying decision rules to construct a consideration set. Furthermore, the results reveal noticeable differences across consumer segments in their decision processes. We find that one consumer segment tends to be more opportunistic in that they consider tickets only for high value games within the next two weeks, with a focus on the lower-priced seats. In contrast, a second segment tends to plan farther ahead, considering events up to a month in advance and are less discerning with respect to game attractiveness and seating tiers.

Our results allow us to examine the impact of team performance and ticket prices on ticket sales. This has tremendous revenue implications for the team – our modeling efforts allow us to examine the value of a winning team in terms of the number of tickets sold as well as how price changes will impact ticket sales for various games and seating tiers.

## Model Development

Our modeling framework considers two stages in the decision process. In the first stage, elimination by aspect (EBA) rules are used to construct a consideration set. In the second stage, a multinomial logit model is used to make a final choice decision. We begin with the latter component.

***Choosing among the alternatives in the consideration set***

In the second-stage choice process, we assume that consumers choose one alternative from the set of options resulting from the consideration stage. Amongst these alternatives is the option to not buy any ticket at all. We utilize the following form of a multinomial logit model with a no-choice option:

$$p_{gst}^{LOGIT} = \frac{\left(I_{gst} \cdot \exp\{V_{gst}\}\right)^{1-D} \left(\exp\{\phi_t\}\right)^D}{\sum_{g=1}^{G}\sum_{s=1}^{S}\left(\exp\{\phi_t\} + I_{gst} \cdot \exp\{V_{gst}\}\right)}$$

where $g$ indexes each of the G games, $s$ indexes each of the S seating tiers, $t$ indexes time, $I_{jt}$ indicates whether alternative $j$ was available for purchase and included in the consideration set at time $t$, and D is an indicator variable to represent the no-choice option. The remaining terms, $V$ and $\phi$, are latent values to be estimated. Specifically, $V_{gst}$, represents the value of the game-seat combination at time $t$ and $\phi_t$ represents the value of the no-choice option at time $t$.

The value for each game-seat choice option at time $t$, $V_{gst}$, can be decomposed into the attractiveness of the specific game $g$ at time $t$ ($A_{gt}$), the value of the seating tier $s$ ($\delta_s$), and a covariate effect for the number of days until the game ($T_{gt}$):

$$V_{gst} = A_{gt} + \delta_s + \beta \cdot \ln(T_{gt})$$

For identification purposes, we fix $\delta$ for one of the seating tiers to equal zero. Consequently, the estimated $\delta$ values for all other seating tiers represent their value relative to the baseline seating tier.

To incorporate time dynamics, we include covariates for win-loss records in both the attractiveness of each game and the no-choice option. We assume that the home team's record will systematically shift the no-choice option as the season progresses, which in turn will affect all game-tier alternatives accordingly. That is, as the home team wins, the no-choice option should become less attractive. Because of the multinomial logit structure of the model, this makes all other alternatives more attractive and ticket sales for all games will increase.

$$\phi_t = b_0 + b_1 \cdot HomeWin\%_t$$

However, one would also expect that the attractiveness of each game is also subject to time dynamics. Specifically, we assume that the win-loss records of the visiting team in each game will influence the attractiveness of that game. For example, a home game against a Stanley Cup hopeful is sure to attract larger crowds than a game against a less successful team. Therefore, we model the latent attractiveness of each game as a function of the win-loss record of the visiting opponent in that game. This covariate varies over time as team records evolve over time.

$$A_{gt} = a_{g0} + a_1 \cdot VisitorWin\%_{gt}$$

To accommodate unobserved heterogeneity across games and to ensure identification, we assume that $a_{g0}$ is distributed across games according to a Normal distribution with mean $\mu$ and standard deviation $\sigma$.

***Modeling the consideration stage***

Now we move backwards to the specification of the model for the consideration set. Specifically, we assume that potential buyers screen out alternatives using a EBA rule and consider only those available choice alternatives that satisfy the following criteria:

1. Game-seat combination must be available for purchase (i.e., it is not sold out)
2. Game attractiveness is greater than a minimum threshold (i.e., $A_{gt} > A^*$).
3. Seating tier value must be greater than a minimum while not exceeding a maximum threshold (i.e., $d_1^* < \delta_s < d_2^*$ ).
4. The game day is within a certain future time period to allow for adequate planning (i.e., $T_1^* < T_{gt} < T_2^*$ ).

The specific cutoffs used for these criteria may vary across consumers. That is, different consumer segments may have different criteria for what they will or will not consider purchasing. Therefore, we allow for consumer heterogeneity by estimating latent consumer segments that vary in terms of their threshold parameters ($A^*$, $d_1^*$, $d_2^*$ , $T_1^*$, $T_2^*$).

***Modeling the effect of price***

The above model specification provides latent estimates for the value of each seating tier. These seating tiers vary in terms of both the quality of the seat as well as price. To examine the role of price in the ticket-buying process, we further model these latent seating tier values to be a function of (1) a tier-specific baseline constant that represents the inherent value of the seat itself and (2) the effect of price (i.e., the face value of the ticket). Because seating values are estimated relative to a baseline tier, we also represent price as a value indexed against that baseline tier.

$\delta_s = C + \gamma * RFV_s + \varepsilon_s$ where RFV (relative face value) is defined as $RFV_s = \frac{FV_s}{FV_0} - 1.$

## Data and Estimation

We estimate our model on home-game ticket sales for an undisclosed NHL team over the 2008-09 season, covering choices made for 41 home games and 16 seating tier options for each game, i.e., a total of 656 game-seat alternatives. Ticket prices ranged from $15 for an upper-level seat (which we will use as our baseline tier) to $300 for the best seats..

We collected win-loss records for the home team and all of their home game opponents. Throughout the season, these records change as games are played. For each day in our selling period, we use the percent of all games won at that point in the season as our covariate measure.

To estimate the model, we use a Metropolis-Hastings MCMC algorithm. This simulation based estimation allows us to estimate the specified model without providing a closed form likelihood function [8].

## Results

To illustrate our model and highlight some empirical findings, we present the results associated with a three-segment model specification. In this specification, one consumer segment is assumed to use no first-stage screening rule, i.e., all consumers within it will consider all available game-seat options. The other two consumer segments use EBA screening rules with different threshold criteria. Table 1 presents these results.

Starting with the multinomial logit choice model estimates, the negative effect of the home team's record on the no-choice option (-1.57) indicates that overall ticket sales increase as the team wins more games. The positive effect of the visiting team record on game attractiveness (.24) suggests that win-loss records in the league can also affect ticket sales. That is, a specific game becomes more attractive if the visiting team has been winning. These results effectively quantify the value of a winning record in terms of ticket sales.

Turning to the consideration stage estimates, we see that one segment of EBA consumers considers only highly attractive games (A>.86) that are occurring within approximately 2 weeks. This group constitutes 45% of the potential ticket buying population. At the start of the season, this attractiveness threshold value eliminates many of the games. The games that survive this screening rule tend to be against opponents that are seen as attractiveness at the start of the season, and maintain an adequately high winning percentage through the season. Furthermore, when these consumers buy tickets, they consider only seating tiers with latent value greater than .1 and less than 2.66. These cutoffs result in the consideration of lower priced tickets only (see Figure 1 below). The second segment (20% of the population), in contrast, is less discriminating in terms of what games they consider (A>-.38) and considers virtually all seating tiers. These ticket buyers also plan farther ahead – considering all games (that are sufficiently attractive) within a 33 day window.

Table 1. Model Results

| | | |
|---|---|---|
| Model fit statistics | | |
| Log-Marginal Density (LMD) | -442,435 | |
| Deviance Information Criteria (DIC) | 884,870 | |
| | | |
| Multinomial logit choice stage estimates | | |
| No choice constant | 4.39 | |
| Effect of home team record on no-choice option | -1.57 | |
| Effect of visiting team record on game attractiveness | .24 | |
| Effect of days until game on choice | -.78 | |
| | | |
| Consideration stage estimates | EBA Segment 1 | EBA Segment 2 |
| Game attractiveness threshold | .86 | -.38 |
| Minimum threshold for seating tier | .10 | -4.90 |
| Maximum threshold for seating tier | 2.66 | 2.63 |
| Minimum threshold for days until game | >1 day | >0 days |
| Maximum threshold for days until game | <16 days | <33 days |
| | | |
| Segment proportion in population | 45% | 20% |
| | | |
| Price effect on value of seating tier | | |
| Baseline constant | .49 | |
| Effect of relative face value price | -.29 | |

The negative price effect (-.29) of relative face value indicates that, as price increases, the value of the seat decreases. This comes as no surprise. However, this estimate allows us to measure the impact of any price change on total ticket sales and revenues. Traditional models that only assume a single-stage choice model may underestimate the impact of such price changes as it allows other dimensions of the purchase to compensate for the price increase. However, our results show that consumers do not necessarily follow a single-stage, purely compensatory decision process. Instead, options can be eliminated from consideration if specific aspects of the option are unacceptable. As a result, it may be the case that a price increase will eliminate a ticket from consideration altogether regardless of how attractive other aspects of the game may be. Therefore, ticket sellers should be cautious as to not manipulate prices to the extent that significant segments of the potential ticket buying population no longer consider purchasing. The model estimates pertaining to the price effect provides guidance as to where that breaking point lies. For example, seating tier 4 in our data was priced at $99, resulting in a perceived value below the minimum threshold for consideration by segment 1. However, our results show that segment 1 consumers would consider this seating tier if its price were dropped to $85. It is easy to how this kind of finding can be of great value to managers.

Finally, we look at the dynamics in game attractiveness as the season progresses. Figure 2a shows the initial levels of attractiveness for each game throughout the season, as well as the cutoff thresholds for both segments. But as the season moves ahead, these levels will change as each opponent plays well or poorly; we show an example of these dynamics (for game 39) as the season unfolds. At the start of the of the season, this game would not be attractive enough for consideration by consumers in segment 1, but as that opponent gets hot at various times during the season, the game merits consideration. Several games drift in and out of consideration for the very picky consumers in this segment; likewise, a number of generally unattractive games occasionally drop below the threshold for consideration by the less-choosy consumers in segment 2. These kinds of choice dynamics seem very plausible and, combined with the other elements of the model described above, offer a very rich "laboratory" for testing and optimization.

***Figure 1.***

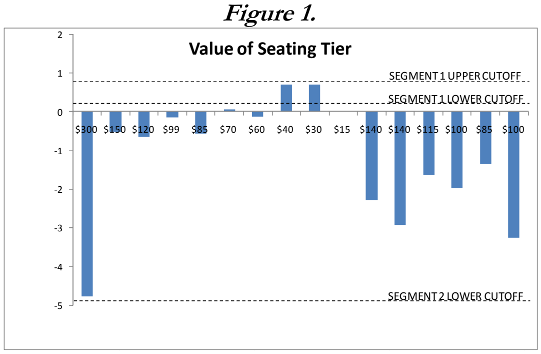



***Figure 2a.***

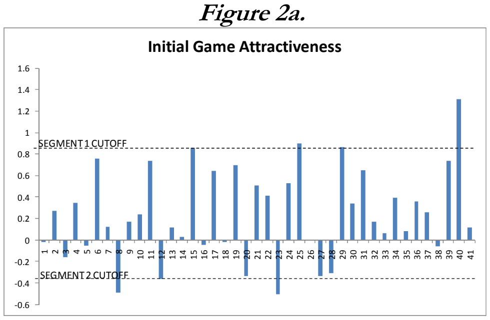
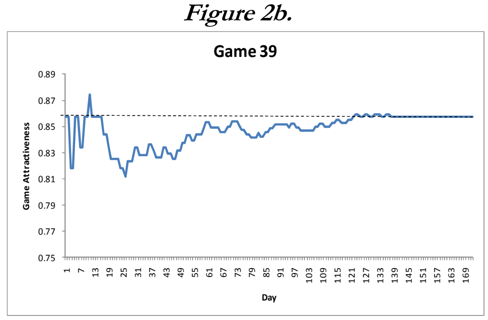



***Figure 2b.***




# Conclusions

In this research, we have developed a two-stage consideration and choice model for ticket sales. Our results convey a high level of face validity for these results, and a number of useful diagnostics emerge quite clearly. But estimating and describing these models is just a starting point. Further testing should be conducted (e.g., applications to other teams/seasons, and the incorporation of additional factors such as media coverage). Once the model is fully validated , it will then be possible to play out different scenarios as briefly mentioned above. For instance, the team can explore the profit implications of different face values for each seating tier, first in the static case, and then allowing for time dynamics. Then it would make sense to look for different heuristics to guide these price dynamics, for instance, coming up with an "early warning" indicator that a game is becoming attractive to the picky consumers in segment 1 – and then to make the appropriate price changes to certain kinds of seats. These kinds of data-driven price dynamics are becoming more common in the professional sports sector, but it is vitally important that the "engine" that underlies these decisions has solid psychological principles associated with it – and we feel that this paper is a useful step forward in that regard.

## References

[1] Shocker, Allan D., Moshe Ben-Akiva, Bruno Boccara, and Prakash Nedungadi, "Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models and Suggestions," Marketing Letters, vol. 2, 181–97, August 1991.

[2] Gensch, Dennis H., "A Two Stage Disaggregate Attribute Choice Model," Marketing Science, vol. 6, 223–31, Summer 1987.

[3] Gilbride, Timothy J. and Greg M. Allenby, "A Choice Model with Conjunctive, Disjunc-tive, and Compensatory Screening Rules," Marketing Science, vol 23, no. 3, 391-406, 2004.

[4] Moe, Wendy, "An Empirical Two-Stage Choice Model with Decision Rules Applied to Internet Clickstream Data," *Journal of Marketing Research*, vol. 43, no. 4, 680-692, 2006.

[5] Bettman, James R., Eric J. Johnson, and John W. Payne, "A Componential Analysis of Cognitive Effort in Choice," Organizational Behavior and Human Decision Processes, vol. 45, no. 1, 111–39, 1990.

[6] McFadden, D., "Conditional logit analysis of qualitative choice behavior," in P. Zarembka Frontiers in Econometrics , Academic Press, New York, 1974.

[7] Guadagni, Peter M. and John D.C. Little, "A Logit Model of Brand Choice Calibrated on Scanner Data", Marketing Science, Vol. 2, No. 3, pp. 203-238, Summer 1983.

[8]Gilbride, Timothy J. and Greg M. Allenby, "Estimating Heterogeneous EBA and Economic Screening Rule Choice Models," Marketing Science, vol. 25, no. 5, 494-509, Sept. 2006.