Evaluating the Effectiveness of Dynamic Pricing Strategies on MLB Single-Game Ticket Revenue

Joseph Xu, Peter Fader, Senthil Veeraraghavan
The Wharton School, University of Pennsylvania
Philadelphia, PA, USA, 19104,
Email: jiaqixu@wharton.upenn.edu, faderp@wharton.upenn.edu, senthilv@wharton.upenn.edu

Abstract

The sports industry has seen a rapid adoption of dynamic pricing practices in recent years. However, there is still limited understanding on the effect of dynamic pricing on revenue in sport event settings and how to execute effective dynamic pricing strategies. In this paper, we address these issues by developing a comprehensive demand model for single-game ticket sales that can be used to predict the revenue associated with a particular pricing strategy over the course of a sport season. We apply the model to actual ticket sales and pricing data from an anonymous Major League Baseball franchise during a recent MLB season, and evaluate the effectiveness of the dynamic pricing policy applied by our partner franchise during that period. We find that the actual dynamic pricing strategy used by this franchise resulted in revenue decrease of 0.79% compared to a pricing policy where prices are fixed over time. We propose alternative pricing policies to help improve revenue and find that an optimized dynamic pricing policy can improve revenue by 14.3% compared to a pricing policy where prices are fixed over time.

1 Introduction

Dynamic pricing of single-game tickets is now a common practice in the North American sport industry. As of 2013, 21 out of 30 MLB clubs used some form of dynamic pricing strategy for their single-game tickets, with clubs in NBA and NHL also beginning to adopt dynamic pricing strategies [1]. Despite the widespread use of this practice, industry practitioners are left in the dark when it comes to evaluating their pricing decisions. Most pricing decisions are guided by a black-box software recommendation from external vendors. This causes validation of pricing decisions and comparison between different pricing strategies to become a significant challenge for industry practitioners. While anecdotal evidence points to the benefits of dynamic pricing, no academic study or analytical framework exists to quantify this benefit or offer guidance on how to design effective dynamic pricing strategies.

Academic research on the effect of dynamic pricing and price discrimination in the entertainment industry has been documented for some settings, such as musical theater and concerts [2, 3], but not for sports events. Related research in the sports industry has been focused on the effect of dynamic pricing in secondary markets [4, 5] and its implication on the primary market rather than focusing on the primary market itself. We seek to fill this gap by investigating the effect of pricing strategy in the primary market. This is a challenge in itself because we must consider a large set of substitutable products for a significant length of selling period. Further complicating this problem is the fact that the underlying value of a stadium seat section is unclear [6]. An effective dynamic pricing policy must also provide proper valuations for various seat sections in the stadium, which can vary from day to day.

In this paper, we develop a comprehensive demand model for single-game ticket sales. We follow the general framework of Moe et al. [7] and model the underlying decision process of a customer seeking to purchase sporting event tickets. We partnered with an anonymous MLB franchise to apply this model to ticket sales data from a recent MLB season to estimate relevant parameters of the model. In addition, we use this model to evaluate the impact of the dynamic pricing strategy adopted by our partner franchise during the particular MLB season. In doing so, we derive insights on the elements of effective dynamic pricing strategies.

We find that the actual dynamic pricing strategy adopted by our partner franchise resulted in 0.79% decrease in overall revenue compared to the baseline flat pricing strategy. We propose three alternative dynamic pricing strategies and show that an optimal dynamic pricing strategy can improve revenue at least 14.3% compared to the baseline flat-pricing strategy.
2 Background and Data Description

2.1 Empirical Background

We partnered with an anonymous MLB franchise and analyzed their ticket sales data from a recent MLB season. The franchise started out the season with a variable pricing policy for single-game tickets. Ticket prices varied across games and across seat sections, but remained constant over time. A dynamic pricing strategy was introduced prior to the All-Star Break, and the franchise started to change prices for individual game-seat section combinations on a daily basis for the rest of the season. Price changes were based on software recommendations from an external vendor, but sometimes with manual overrides to satisfy internal constraints of the franchise. Most of the price changes were at small increments ($1 to $3) and the franchise maintained non-decreasing price paths for a large majority of the game-seat combinations. Figure 1 shows a sample price path for a home game in August.

![Figure 1: Actual Price Path of Select Seat Sections for a Home Game on August](image)

In the empirical setting provides an ideal set-up to study this problem because ticket prices were independent of other variables in the first part of the data. In leveraging this aspect of the data, we overcome the challenge of price endogeneity since there is no interaction between pricing decision and other relevant factors such as time until game and remaining inventory. This allows us to obtain accurate parameter estimates for our model, especially with respect to customer sensitivity to price.

2.2 Data Description and Methodology

Our main data is single-game ticket sales data at the transaction level for a single MLB season, provided by our partner franchise. This includes all single-game tickets sold at the box office and online. For every transaction, we know the time of purchase, number of tickets purchased, price paid and the choice of seat section. We note that these transaction data do not contain identifiable customer ID or additional customer-level data such as demographics. We will focus on transactions that took place after MLB Opening Day because we expect the demand profile for pre-season and the opening game to differ considerably from other regular-season games.

In addition to the transaction data, our franchise partner also provided prices for each seat section and cumulative number of tickets sold for each section at the daily level. The cumulative ticket sales data includes tickets sold from other channels such as season tickets, group tickets, complimentary tickets and discount tickets. Game-specific data such as start time, weather and promotion schedule are also provided. Finally, we supplement our analysis with team performance data in terms of wins and losses during the particular season for the home team and their opponents.

2.3 Methodology

We split the data into two parts. The first part of data is from the early part of the season where prices remained constant and the second part of the data is from the later part of the season where dynamic pricing was applied to the remaining games. We use the first part to estimate the model parameters and use the second part to generate revenue predictions based on the estimates.
These parameter estimates allow us to generate predictions for expected revenue given a particular price path. By generating revenue predictions for different dynamic pricing strategies, we are able to examine how various dynamic pricing strategies impact the overall revenue for our partner franchise.

3 Single-Game Ticket Sales Model

We model customer demand for single-game tickets through a three-stage decision process: game decision, ticket quantity decision and seat section decision. In the first stage, customers decide which game to attend from the remaining home games. In the second stage, customers decide on the number of tickets to purchase for that particular game. Finally, in the third stage, customers decide which seat section to purchase given their choice of game and the number of tickets required. We propose three separate models for each of these decision stages, whose output will be combined to generate predictions for the expected revenue.

3.1 Game Demand Model

We begin by modeling the first stage of customer demand process: game choice. In this stage of ticket purchase, customers choose a game from the remaining home games that they would like to attend. We model this process aggregate using negative binomial regression. Instead of considering each order at the transaction level, we will use daily number of orders as the dependent variable. Using this model, we will predict the expected number of orders (i.e. number of unique transactions in a given day) for the remaining home games.

For each day \( t \) that game \( i \) is on sale, we model the aggregate number of daily orders \( n_{it} \) using negative binomial regression and control for the effect of time, game characteristics, team performance, price and occupancy. We control for effect of time through covariates such as number of days until game \( i \) at time \( t \) and indicator variables for each of the days in the last 8 days before the game. Game characteristics component of the model accounts for various game-specific factors which may affect attendance such as the opponent, price category used in the original variable pricing scheme, time of the game and promotion schedules. Team Performance is measured in terms of wins above .500, which is the difference between the number of wins and the number of losses by a team at time \( t \). This metric is tracked for both the home team and opponents throughout the season. Finally, we include the effect of price and cumulative sales up to the observation date in the model. We define occupancy as cumulative sales to-date divided by the total capacity, expressed in percentage. Stadium-wide average price and average occupancy are used for this stage of the model. The full list of variables used in the game demand model are shown on Table 2 in the Technical Appendix.

The game choice decision is modelled at the aggregate level because of data limitations. In order to have a transaction-level game-choice model, we will need to know the total population of customers. Since our data contains only actual transactions, we cannot accurately estimate a choice model for this stage. On the other hand, aggregate demand model allows the use of the negative binomial model to capture overdispersed count data. The negative binomial regression is particularly useful in this setting, since most of the transactions take place in the last few days before the game. Our specification provides good predictions for the total number of transactions at game-by-game basis (Figure 1 in the Technical Appendix).

The parameter estimates (Table 3 in the Technical Appendix) match our intuitions very well: higher price lowers demand, promotions increase demand, and opponents tend to affect demand. In addition to these intuitive results, we also find that stadium occupancy and team performance have a significant effect on demand. The model estimates suggest that, all else being equal, a 1% increase in stadium occupancy has the same effect on demand as lowering the price $3.47 across the entire stadium. Similarly, an extra win for the home team would have the same effect on demand as lowering the price $1.09 across the entire stadium. These results suggest that stadium occupancy and team performance can be an important factor in dynamic pricing decisions.

3.2 Ticket Quantity Model

The second stage of the customer decision process involves customers deciding how many tickets to purchase for the game chosen in the first stage. We model this process with negative binomial regression. This specification is again useful because the number of tickets is a count data that shows significant variation. In order to keep the models consistent, we use the same set of variables as the game demand model.
We use the same specification as the game choice model. We use negative binomial regression to model the expected number of tickets $x_{it}$ for an order $i$ placed at time $t$, controlling for effects of time, game characteristics, team performance, price and occupancy. In this stage of the model, stadium-wide average price and average occupancy are used. We subtract one from the observed number of tickets $x_{it}$ (zero-truncation) in the regression because our model population is now the set of customers who has purchased at least one ticket. Estimates for some of the relevant model parameters are shown on Table 3 in the Technical Appendix. The results are consistent with the game demand model. We find negative effect of price on demand as well as positive effect of stadium occupancy and home team performance on demand.

### 3.3 Seat Section Choice Model

The last stage of the customer demand process is the seat section choice. Given their previous choice of which game to attend and how many tickets to purchase, customers now select a location in the stadium. We model this process using multinomial logistic regression. An outside option does not need to be considering in this case, since we already assume that these customers are committed to purchasing tickets. Customers are allowed to choose from the fourteen seat sections in the stadium (Table 1 in the Technical Appendix). We normalize our model using the bleacher section as the reference level.

The specification of the model is as follows. For a seat section $i$ in stadium $S$, the probability of a customer choosing this seat section given price $p_i$ and occupancy $o_i$ is:

$$ P[Y_i = 1] = \frac{e^{\alpha_i + \theta_1 p_i + \theta_2 o_i}}{\sum_{j \in S} e^{\alpha_j + \theta_1 p_j + \theta_2 o_j}} $$

The price and occupancy used in this stage of the model is the seat section-specific price and occupancy. The $\alpha_i$ term in the equation is the seat section-specific intercept of how likely it is to be chosen by a customer. We incorporate additional factors such as time until game and number of tickets required into the intercept to appropriately capture some empirical observations. In particular, we note that the best seat sections tend to be sold earlier than other seat sections. We also note that customers requiring more tickets would be less likely to purchase from more expensive seat sections. Therefore, the intercept for section $i$ with time until game $t$ given demand with $n$ tickets would be: $\alpha_i + r_i \cdot \log(t + 1) + \eta_i \cdot n$. Our parameter estimates show that for the expensive seat sections, time coefficient $r_i$ are the most positive while number of tickets coefficient $\eta_i$ are the most negative. The opposite is true for cheaper seat sections. Parameter estimates for the seat section model are shown on Table 4 in the Technical Appendix.

### 3.4 Revenue Evaluation

The three main components of the model are combined to generate predictions for the expected daily revenue for each seat section from every game. However, evaluating the total revenue for each game during the entire selling period is still challenging because past sales affect future sales in our model (from the occupancy term). The expected revenue $R$ given remaining selling period $t$ and the vector of remaining inventory in each seat section $\bar{x}$ is:

$$ R(t, \bar{x}) = E_d[\bar{p} \cdot \bar{s} + R(t - 1, \bar{x} - \bar{s})] $$

where $\bar{p}$ is the vector of price for each seat section and $\bar{s}$ is the vector of sales for each seat section. Given the length of selling period (up to three month) and the number of seat sections under consideration, it is impractical to evaluate this function directly. Instead, we will approximate the revenue function using the following function:

$$ \bar{R}(t, \bar{x}) = \bar{p} \cdot E_d[\bar{s}] + \bar{R}(t - 1, \bar{x} - E_d[\bar{s}]) $$

This approximation will use only the daily expected sales for each seat section and adjust the remaining inventory on a daily basis. The effect of randomness in daily sales will be ignored by this approximation, but this approach significantly reduces computation time and yields similar results as Monte Carlo simulation. Table 1 shows the predicted total revenue versus observed total revenue for the second half of the data. The predicted total revenue for the 48 remaining home games in the second half is $4,158,746 while the observed total revenue for the 48 remaining home games was $4,517,901. The error is -$359,155 and is 7.95% of the observed revenue. However, the last home game contributes a significant amount of error (-$156,499). Once we exclude the last home game, we find that a large portion of the game-by-game error evens out over the course of the season, leading to a good prediction.
for the total revenue through the season. The comparison of predicted revenue versus actual revenue on a game-by-game basis are shown in Figure 2 in the Technical Appendix.

<table>
<thead>
<tr>
<th>Total Revenue ($K)</th>
<th>Error ($K)</th>
<th>Error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>4,334</td>
<td>-</td>
</tr>
<tr>
<td>Predicted</td>
<td>4,132</td>
<td>-202</td>
</tr>
</tbody>
</table>

### 4 The Effect of Dynamic Pricing on Revenue

We now use the model described in the previous section to evaluate counterfactual scenarios with new dynamic pricing strategies. We first compare two strategies: actual price versus baseline flat price. Actual price strategy refers to the dynamic pricing strategy applied by our partner franchise (Figure 1). The baseline flat price strategy uses the price observed on the day before the implementation of dynamic pricing for the entire selling period for all games. In the second part of our data, our partner franchise was able to dynamically price 49 home games. In this calculation, we consider the remaining 47 games (the last home game is excluded and another game is excluded due to data integrity issue) and compare the expected revenue from these two dynamic pricing strategies.

Table 2 shows the expected revenue from the actual price strategy and the baseline flat pricing strategy. We find that revenue from actual price strategy was lower than revenue from baseline flat pricing strategy by 0.79%. On game-by-game level, there were 14 games that experienced revenue increase through dynamic pricing, with average increase of 1.37%. However, 27 games experienced revenue decrease from dynamic pricing, with average decrease of 1.20%. 6 games were unaffected because prices remained constant under the actual dynamic pricing strategy.

<table>
<thead>
<tr>
<th>Total Revenue ($K)</th>
<th>Change ($K)</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Flat Pricing</td>
<td>4,165</td>
<td>-</td>
</tr>
<tr>
<td>Actual Dynamic Pricing</td>
<td>4,132</td>
<td>-33</td>
</tr>
</tbody>
</table>

We find that 8 out of 14 games with revenue increase took place before mid-July while only 2 out of 27 games with revenue decrease occurred took place during the same period. From this point onward, the home team experienced significant underperformance on the field. Since mid-July, the home team had a sub-.350 record for the last 70 games of the year. The team performance during the prediction period was very different from the team performance observed during the calibration period, possibly affecting the final result. At the same time, our parameter estimate suggests that the correct response during this period would be to lower ticket prices. Estimates from the game demand model suggest that an extra loss for the home team has the same effect on demand as raising the ticket price by $1.09 across the entire stadium. We believe that the overall revenue was negatively affected because ticket prices did not go down even though losses were piling up for the home team; however, our partner franchise had strategic long-term rationale related to consumer behavior and pricing integrity that lead them to not lower prices for these games.

### 5 Analysis of Alternative Dynamic Pricing Strategies

In this section, we will propose and evaluate the effectiveness of alternative dynamic pricing strategies. We consider three strategies: optimal variable pricing with look-ahead on team performance, monotone myopic dynamic pricing and unrestricted myopic dynamic pricing. The expected revenue from these strategies will be compared with the expected revenue from the baseline flat pricing strategy described in the previous section. We find that despite the simplicity of these strategies, significant revenue improvement can be achieved. In addition, these strategies can be computed quickly using standard optimization packages and are therefore easily implementable in practice.

We first consider the optimal variable pricing strategy with look-ahead on team performance. This strategy allows prices to vary between games and seat sections, but not across time. For each game under consideration, we will use our three-stage model described in Section 3 to compute expected revenue for the entire selling season. We will then search through a range of possible prices to maximize the expected revenue for the entire selling period of each game. We allow optimization to have information on the future team performance variable. This is done to illustrate the maximum upside from a variable pricing strategy, but in practice one would generate a performance forecast or eliminate this factor by setting the values for these variables to zero in the optimization.
In the monotone myopic dynamic pricing strategy, seat section prices are determined to maximize the daily expected revenue. This strategy allows prices to vary not only between games and seat section but also across time. For each game, we will use our three-stage model described in Section 3 to compute prices that would maximize expected daily revenue for each day that the game is on sale. In the monotone case, we assume that prices cannot be lower than prices from the previous day (i.e. monotone increasing). This is a sensible assumption since many organizations try to avoid making customers feel regret from purchasing an expensive ticket early on.

In the unrestricted myopic dynamic pricing strategy, we use the same approach as the monotone myopic dynamic pricing strategy but with no restriction on the price changes day-to-day. Prices are free to increase and decrease, as long as they remain within a pre-determined upper and lower bound. These two myopic strategies under consideration are attractive since their computational simplicity allows price recommendations to be generated far into the future very quickly.

We impose several constraints on the optimization. First, we set the minimum price for each seat section to be the price observed on the very first day of dynamic pricing. Maximum price constraints for each seat section are imposed at a reasonable price point based on discussions with our partner franchise. We also maintain certain price-quality relationships so that high-quality seats are always more expensive than seats with lower quality. For example, we will always try to maintain the prices so that dugout seats are always more expensive than outfield seats. These relationships are described in Table 1 of the Technical Appendix. We use DEoptimR package in the R Programming Language for optimization.

Table 3 shows the total revenue from baseline flat pricing, actual dynamic pricing, optimal variable price strategy, myopic monotone dynamic pricing and unrestricted myopic dynamic pricing strategy. We find that the optimal variable pricing strategy improves revenue by 13.5% over the baseline flat pricing strategy. This suggests that even adopting a variable pricing strategy for the second half of the season can improve revenue significantly if these prices can be chosen appropriately.

<table>
<thead>
<tr>
<th>Total Revenue ($K)</th>
<th>Change ($K)</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day 0 Flat Pricing</td>
<td>4,165</td>
<td>-</td>
</tr>
<tr>
<td>Actual Dynamic Pricing</td>
<td>4,132</td>
<td>-33</td>
</tr>
<tr>
<td>Optimal Variable Pricing</td>
<td>4,729</td>
<td>564</td>
</tr>
<tr>
<td>Optimal Monotone Myopic Dynamic Pricing</td>
<td>4,310</td>
<td>145</td>
</tr>
<tr>
<td>Optimal Unrestricted Myopic Dynamic Pricing</td>
<td>4,760</td>
<td>595</td>
</tr>
</tbody>
</table>

Figure 2 shows the optimal variable price for several seat sections for the particular home game in Figure 1. We note some pricing discrepancy in the original pricing strategy and the optimal variable pricing strategy. For example, we find that Outfield 2 section, which was originally valued near $20 should be priced around $35. This result helps us identify the correct valuation of a seat section based on past customer demand.

We omit the price path for the optimal monotone myopic dynamic pricing strategy. This strategy leads to a flat price over time given our particular data. However, monotone myopic dynamic pricing strategy has significantly lower revenue compared to the optimal variable pricing strategy. We find that revenue from monotone myopic dynamic pricing strategy increases revenue by 3.48% compared to the baseline flat pricing strategy. One may find it rather strange that the increase from dynamic pricing is less than the increase from variable pricing. This peculiarity is caused by a rather unusual team performance. The myopic dynamic pricing strategy had set high initial prices for each seat section at the beginning, when the home team had a winning record. These prices remained fixed due to the monotonicity restriction, even when home team suffered losses in the later part of the season. The optimal variable pricing strategy, on the other hand, were allowed to look ahead on future team performance and was thus able to have lower initial price for each seat section.

Finally, we find that significant improvement can be achieved with unrestricted myopic dynamic pricing strategy. The unrestricted myopic dynamic pricing strategy improves revenue by 14.29% over the baseline flat pricing strategy. Figure 3 shows the sample price path for several seat sections for the particular home game shown in Figure 1. We note that even though Outfield 2 section is priced around $35 (the same value as optimal variable pricing strategy), the price of this section is aggressively lowered near the game day. The pricing flexibility allows prices to react properly to changes in exogenous demand factors such as team performance, allowing the franchise to reap the benefit of dynamic pricing.
We note several features from these optimization that can serve as a general guideline for implementing dynamic pricing. First, we find that there is little value to changing the price early in the selling period. While there are fluctuations in the optimal prices for a few sections, these suggested price changes are too small to be of significant value. Most of the pricing changes should happen near the end of the selling period when the demand begins to increase sharply. On the other hand, there is significant value in setting a good initial price to the seat sections, as evidenced by the optimal variable pricing policy. Second, we find that not all seat sections need to follow the same type of price path. In Figure 3, the optimal price path involved increasing price path for Upper Deck 1 section and a decreasing price path for Outfield 2 section. The right type of price path for each seat section depends on various factors such as time until game, team performance and remaining seats in each section. Finally, we note the effect of internal policies such as monotonicity restrictions with respect to dynamic pricing. The value of dynamic pricing is derived from pricing flexibility and the ability to react to external demand shocks through prices. A policy restricting pricing flexibility will necessarily reduce the value of dynamic pricing. At the same time, sport teams may wish to implement certain pricing policies for various reasons (e.g. maintaining consistent marketing message, managing customer relationships). Our model enables teams to make more informed decisions regarding the adoption of pricing policies by helping them compute the revenue impact.

Finally, we note that we were able to achieve 14.3% improvement obtained even with significant amount of optimization constraint. We did not allow seat section prices to go below our initial observation and maintained a strict price-quality relationship between seat sections. If any of these restrictions are relaxed, we believe that the expected gain from dynamic pricing can be greater than the amount documented in this section.
6 Conclusion

We build a customer demand model for single-game ticket sales that can be used to analyze the effect of pricing on revenue. The model is based on three stages of customer decision process in ticket purchase: game choice, ticket quantity choice and seat section choice. The model is able to generate good predictions for total revenue over the course of the season, with prediction error within 4.68% of the actual observed revenue.

Using the customer demand model, we are able to evaluate the effectiveness of the dynamic pricing strategy applied by our partner MLB franchise. We find that revenue would have been higher under the baseline flat pricing strategy. The actual dynamic pricing strategy lowered revenue by 0.79%, and we attribute this loss to the non-decreasing price restriction. We propose three alternative dynamic pricing strategies and evaluate their effectiveness compared to the baseline flat pricing strategy. The first strategy, optimal variable pricing strategy (assuming perfect knowledge of future performance), improves revenue by 13.5%. The second strategy, monotone myopic dynamic pricing strategy increases revenue by 3.49%. Finally, the third strategy, unrestricted myopic pricing policy, improves revenue by 14.3%. As pricing strategy directly affects customers and many internal stakeholders (e.g. communications, ticketing, marketing, sales), these quantitative revenue estimations are critical information to incorporate with qualitative data and longer-term strategic, behavioral, and perception considerations for pricing decision-making.

We contribute to the understanding of how to effectively design and execute dynamic pricing strategies by constructing a theoretical model based on customer decision processes. We are also able to estimate key parameters of this model, leading to a better understanding of how customers respond to various factors such as price, stadium occupancy and team performance. While our model is applied to baseball settings, this general framework can be applied to many other sporting event settings. We believe that our modeling framework would be particularly valuable as dynamic pricing develops into a standard practice within the sports industry.

7 Acknowledgements

We would like to thank our partner MLB franchise for providing us with the raw data. We also received numerous feedback and insight through our communication with members of their business analytics office and are deeply appreciative of these contributions.

8 References


