The paper empirically tests American attendance factors at Major League Baseball games. The results show that a team’s winning ratio, playoff appearances, stadium capacity, games played, and moving to a new city are positive catalysts for attendance. Surprisingly, the results also indicate a positive attendance response to the gap of years between World Series victories. Regardless of fan motivation, a common behavioral attendance response is estimated. After an extremely prolonged period, the number of gap effect fans can be substantial and therefore valuable to a baseball franchise. Conceivably, the longer a baseball team does not win a World Series, the less incentive there is to win the World Series.
Estimating the World Series Gap Effect on Attendance

Professional athletics in the United States, including Major League Baseball (MLB), are major business enterprises which generate significant revenues for leagues, teams, players, and cities alike. In professional baseball, like many other professional sports, team revenue is predominantly generated through event attendance, broadcasting rights, and merchandising. Maintaining robust fan interest is critical to the long-term sustainability, both for a particular team franchise and the league in which the teams participate.

As sports teams are often large and sophisticated, they contain many layers of organizational management. While the team will have central goals, its size and complexity allows for the potential of smaller units maintaining incentive structures and influences which differ from the stated organizational goal. It is therefore possible that the marketing and promotional aims of a sports team may run counter to the organization goal of victory.

This study focuses on determining fan interest as represented by total yearly attendance. Attendance data for MLB teams are both verifiable and plentiful for empirical testing. A team’s financial success is driven by the fan base it cultivates and the willingness of these fans to monetarily support a franchise. Attendance can be used as a relative proxy for fan interest in the club. If fans appreciate the product and the play on the field, more fans will attend games. If fans do not appreciate the product, attendance will decline. Maintaining fan interest in the sports club is the responsibility of marketing and promotions departments. Their success is driven by attendance which is normally aligned with organizational goal of winning. It is theorized that this might not always be the case.

Literature Review

Studies using major-league baseball attendance began in earnest after Demmert (1973) and Noll (1974) attempted to differentiate individual influences on attendance using cross-sectional data. Over time, this field of analysis expanded as further research focused on an increasing variety of interest areas. These studies examined the presence of fan loyalty and its impact on attendance (Depken 2000, 2001; Won & Lee, 2008). Also of interest were quality-of-play measures such as winning percentage, pitching-staff quality and home-run hitting as examined by Greenstein and Marcum (1981), Marburger (1997) & Horowitz (2007). In addition to winning percentage and quality-of-play, attendance was found to be influenced by the anticipation of postseason play and a higher probability of winning a championship (Whitney, 1998). Additional areas of interest included competitive bal-
ance and outcome uncertainty (Knowles et. al., 1992), and player turnover on a team (Kahane & Shmanske, 1997).

Further market characteristics and the influence of multi-team markets were studied by Pan et. al. (1997). Play disruptions resulting from labor strikes was examined by Schmidt and Berri (2004), while the impact of new stadiums on attendance was studied by Clapp and Hakes (2005) and also by Zygmont and Leadley (2005).

It is assumed that the goal of winning is unified throughout the organization. This study builds on prior research by relaxing this assumption and allowing for the potential of goal diversity within the organization. Research done by Whitney (1998) shows the attendance impact of quality-of-play and postseason pursuit of championships. This study extends this research by examining the implications of time gaps between championships on attendance. If attendance is linked to activities other than winning, the potential exists that the aims of the marketing department may differ from the pursuit of championships. Understanding instances of seemingly atypical fan behavior might provide better insight into the incentives facing those responsible for attendance.

The motivation for the study is the author’s interest in the Chicago Cubs and the ability of a team to maintain a loyal fan base and high attendance despite not winning a championship for an extended period. Lacking championships, prior research should suggest declining fan support and lower attendance. However, attendance has remained strong suggesting the potential for additional and yet unobserved variable influencing fans to remain loyal. The loyalty of these fans, regardless of team victory, provides little incentive for the marketing/promotions group to aspire to victory even though this may be a top organizational goal.

The Chicago Cubs fans have adopted the lovable loser image to anecdotally quantify the loyalty effect. The fans of other teams are far less generous with their labeling of this effect. Unlike the Chicago Cubs, fans of many teams are very vocal in their dissent for teams if lacking success. While possible that some dissenting fans will leave the team and not attend games, it is also possible that another group of fans will bond and commiserate together. While vocal in their displeasure, they still attend games. The shared misery of these fans could be a unifying influence for fans and increase attendance in excess of performance. For the purposes of this study, fan motivation is less relevant. The study is interested in common behaviors and actions of the fans towards attendance as the time gap expands regardless of motivation. The study empirically tests for a consistent World Series gap effect on attendance common to all MLB teams.
Methodology

Using historical baseball data, a model is used to assess the attendance implications of various factors, including a factor to measure the time gap between World Series victories. If a team exhibits a gap effect, the study attempts to quantifiably measure this influence. If the potential value of this effect exceeds attendance benefits of winning the World Series, victory could potentially endanger this effect and reduce the team’s long-term attendance. The result can provide a disincentive to winning the World Series.

The study gathers historical baseball data on 30 MLB teams. The teams represent the 30 current MLB teams dating back to their origins. The 30 MLB teams are followed over time through various movements between cities, stadiums, and team names. The starting point for the time series is 1903, as this is the date of the first World Series. MLB data are collected from 1903 through the 2013 season (Baseball Almanac, Baseball Reference, 2014).

The focus of the study is attendance and therefore total team yearly attendance variable is the dependent variable. Teams enjoy a measure of fan loyalty as evidenced in prior research. It is reasonable that autoregressive tendencies exist with respect to attendance. The attendance variable is lagged one period and included as a control for many influences including loyalty, season ticket sales, and multi-period macro events such as general public sentiment of the team, local economic conditions, and national events.

The ability of teams to attract crowds and increase attendance is constrained by two limiting factors. The first is the number of games played in a season, and the second is the size of stadium in which they play. These variables potentially could be right-censors on the attendance variable and negatively bias the results. To help compensate, variables are added to the model for the number of games played in the season and stadium capacity as further controls.

The winning ratio of each team is included to illustrate the effect of winning on attendance. The team’s winning ratio is wins divided by games played. A binary playoff variable is included to measure the impact if the team participates in post season play. As teams play better and playoff potentials are realized, it is reasoned that fans would appreciate the product as demonstrated by increases in attendance.

A new city binary variable is included to test the impact of moving a team from one city to another. Beyond the population differences between cities, a move might excite fans and increase attendance. The new city variable is also lagged one period to test for the potential of lingering influence of a city move.

United States Census (2014) data are used for U.S. city populations while World Population Statistics (2014) were used for Canadian cities. Attempts to appropriately link each team with a host city are made. Ideally, a team is associated
with a metropolitan area as the appeal of a team likely extends beyond the city limits. However, Census does not provide metropolitan area population statistics prior to 1950. As Census data are not annual, each Census point is repeated in the subsequent years until the next Census data are available.

A variable is included for a count of other team(s) in the host city. For example, New York has two teams during the 2013 season. The other team variable for New York in 2013 is one. During some periods, New York had three teams. During these years, the other team variable for New York is two. The potential exists in multi-team markets for teams to vie for the same fan base and cannibalize attendance. Including this variable allows for assessment of the interaction of multiple teams in single market.

Finally, a gap variable is calculated as the years between a team’s World Series victories. The year of a World Series victory, the variable is zero. The count starts the year after the World Series win and continues until restarted with another World Series win. The purpose of the variable is to detect the gap effect. Summary statistics for the non-lagged variables are presented in Table 1.

A model is constructed for the data set consisting of yearly data for 30 teams.

\[ Q_{i,t} = \beta_0 + \sum_{v=1}^{10} \beta_v (X_{v,i,t}) + \epsilon \]  

\( Q \) = Dependent (Yearly Attendance),  
\( i \) = Team,  
\( t \) = Year,  
\( v \) = Independent Variables

The MLB data collection is an unbalanced panel data set. Teams have different abilities to draw fans for attendance which is unique to the team. Yearly influences affecting game attendance can also impact the results. To control for team specific and yearly influences, a fixed effects model is used.

The yearly control of the fixed effects model will account for variations over time resulting from such influences as additional use of night games, poor weather seasons, wars, league integration, and weekend games. With over a century of data, approximately normal distributions for home-away schedules, doubleheaders, and home schedules for popular teams can be assumed.

Using a fixed effect model forces the same set of slope-coefficients on each franchise in every year. It is testing for gap effect common across all teams and not a condition specific to a particular MLB team. While motivations of fans may differ, the study examines for common fan behavioral responses for all teams with respect to attendance.
Results

The detailed results for all MLB teams are provided in Table 2. The model is highly significant as a fit for baseball attendance. Examining the individual significance of the independent variables provides interesting insights into fan behavior as it relates to game attendance.

Table 1. Summary Statistics Years 1903 to 2013

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Year Attendance</td>
<td>1,390,159</td>
<td>922,145</td>
<td>80,922</td>
<td>4,483,350</td>
</tr>
<tr>
<td>Games Played</td>
<td>156.42</td>
<td>9.14</td>
<td>102(^a)</td>
<td>165(^b)</td>
</tr>
<tr>
<td>Stadium Capacity</td>
<td>42,932</td>
<td>14,606</td>
<td>6,000</td>
<td>93,607</td>
</tr>
<tr>
<td>Winning Ratio</td>
<td>0.5000</td>
<td>0.0823</td>
<td>0.2353</td>
<td>0.7632</td>
</tr>
<tr>
<td>Playoff</td>
<td>0.1685</td>
<td>0.3744</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>New City</td>
<td>0.01</td>
<td>0.10</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Population (Host City)</td>
<td>1,860,456</td>
<td>2,049,900</td>
<td>279,000</td>
<td>8,406,000</td>
</tr>
<tr>
<td>Other Teams (in Host City)</td>
<td>0.46</td>
<td>0.63</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Years Since World Series Win</td>
<td>20.06</td>
<td>19.67</td>
<td>0</td>
<td>105</td>
</tr>
</tbody>
</table>

Note. There are 2,314 Observations
\(^a\) 1981 M.L.B. Strike season
\(^b\) 1962 N.L. Tie breaker (3 games L.A. Dodgers/S.F. Giants)

The lagged attendance variable is a significant indicator of attendance. Autocorrelation factors such as fan loyalty, season ticket sales, and multi-period macro events such as general public sentiment of the team, local economic conditions, and national events are important to yearly attendance. The number of games played in the season and the stadium capacity are also positive and significant influences on yearly team attendance.

The winning ratio is a significant and positive influence on attendance. The results confirm that better play on the field yields increased attendance. A team making the playoffs gains increased attendance, as shown by the positive significance of the playoff variable. The average for all teams is an increase in attendance of approximately 140,996 for a playoff appearance, beyond other controlled variables including winning percentage. The attendance increase likely reflects both the ad-
ditional home games afforded a playoff team and attendance increases towards the end of the regular season due to increasing excitement at the prospect of postseason play.

Table 2. *Fixed Effects Regression Results – World Series Gap Effect*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coef</th>
<th>Std Error</th>
<th>t</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance Lagged 1 Period</td>
<td>0.837</td>
<td>0.008</td>
<td>101.43</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Games Played</td>
<td>12,457.3</td>
<td>621.0</td>
<td>20.20</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Stadium Capacity</td>
<td>1.398</td>
<td>0.505</td>
<td>2.77</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Winning Ratio</td>
<td>1,045,992.0</td>
<td>82,643.1</td>
<td>12.66</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Playoff Appearance (Binary)</td>
<td>140,996.2</td>
<td>17,779.6</td>
<td>7.93</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>New City (Binary)</td>
<td>619,361.9</td>
<td>75,099.7</td>
<td>8.25</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>New City (Binary) Lagged 1 Period</td>
<td>-114,537.0</td>
<td>51,107.8</td>
<td>2.24</td>
<td>0.03</td>
</tr>
<tr>
<td>Population (Host City)</td>
<td>-0.016</td>
<td>0.008</td>
<td>2.13</td>
<td>0.03</td>
</tr>
<tr>
<td>Other Teams (in Host City)</td>
<td>-50,300.5</td>
<td>20,036.4</td>
<td>2.51</td>
<td>0.01 ***</td>
</tr>
<tr>
<td>Gap Effect (Years Since World Series Win)</td>
<td>1.769.1</td>
<td>340.2</td>
<td>5.20</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2,310,350</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. Statistical Significance 1%***

*Groups (Year: Team), F(10,2260) = 2259.58 or Prob > F 0.0000
R-Sqr: within 0.9091, between 0.9807, overall 0.9182
Observations 2300, Groups 30, Rho 0.0912*

A team moving to a new city positively impacts attendance in the year following the move. The coefficient of the new city variable is positive and significant. The impact is in addition to the changes resulting from population differences of host cities. Fans of the new city actively attend games. Averaged across all teams, a moving a team to a new city benefits team attendance by approximately 619,362. However, these benefits are quickly fading. After one season, the attendance benefits of a team in a new city are insignificant as indicated by the lagged city move.
variable. Moving a team between cities provides a novelty benefit for a single year, beyond which typical factors such as the winning ratio dictates attendance.

The population of the host city is not significant. As noted by the prior research (Demmert, 1973; Noll, 1974) it is possible that this is the result of correlation with other explanatory and control variables.

Table 3. Fixed Effects Regression Results – World Series Gap Effect Without Chicago Cubs Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coef</th>
<th>Std Error</th>
<th>t</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance Lagged 1 Period</td>
<td>0.834</td>
<td>0.009</td>
<td>97.53</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Games Played</td>
<td>12,555.6</td>
<td>640.7</td>
<td>19.60</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Stadium Capacity</td>
<td>1.308</td>
<td>0.515</td>
<td>2.54</td>
<td>0.01 ***</td>
</tr>
<tr>
<td>Winning Ratio</td>
<td>1,067,057.0</td>
<td>85,514.0</td>
<td>12.48</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Playoff Appearance (Binary)</td>
<td>142,626.5</td>
<td>18,351.9</td>
<td>7.77</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>New City (Binary)</td>
<td>617,702.1</td>
<td>75,875.9</td>
<td>8.14</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>New City (Binary) Lagged 1 Period</td>
<td>-117,319.8</td>
<td>51,640.5</td>
<td>2.27</td>
<td>0.02</td>
</tr>
<tr>
<td>Population (Host City)</td>
<td>-0.148</td>
<td>0.008</td>
<td>1.90</td>
<td>0.06</td>
</tr>
<tr>
<td>Other Teams (in Host City)</td>
<td>-58,716.8</td>
<td>20,596.5</td>
<td>2.85</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Gap Effect (Years Since World Series Win)</td>
<td>1,312.8</td>
<td>372.9</td>
<td>3.52</td>
<td>0.00 ***</td>
</tr>
<tr>
<td>Intercept</td>
<td>-2,304,837</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Statistical Significance 1%***
Groups (Year: Team), $F(10,2150) = 2096.7$ or Prob $> F 0.0000$
R-Sqr: within 0.9070, between 0.9798, overall 0.9167
Observations 2189, Groups 29, Rho 0.0944

The inclusion of other teams in a market variable is negative and significant. Multiple teams in a single host city are shown to have some cannibalizing effect on each other’s fan base. The presence of each additional team in a market has a relatively modest cost of approximately 50,300 fans in yearly attendance.
The results are largely consistent with the prior noted research when assessing the same or similar influences on yearly attendance. Successful play on the field and postseason play yield positive attendance results. In an attempt to better understand strong attendance despite lack of championships, the study adds a variable to assess the impact of years between World Series victories.

The measure of years since World Series win variable is significant and positive. While the number of the gap effect fans is initially very small, they grow marginally over time at a rate of about 1,769 fans per year. After a prolonged period without a World Series victory, the gap effect fans can represent a substantial portion of the team’s overall fan base.

It is possible that the results are biased by the Chicago Cubs. The length of time between World Series victories and the easy adoption of the lovable loser image by the Chicago fans could be distorting the results. These fans could be outliers and no such effect exists outside Chicago Cubs franchise due to the uniqueness of the fans. To test for this possibility, the model was run without the Chicago Cub data. The results are consistent both with and without the Chicago Cubs data, results provided in Table 3. The gap effect is shown to be consistent across MLB teams.

The attendance influence of a gap effect is estimated. For MLB teams with a shorter time gap since a previous World Series win, the benefits of a World Series victory and resulting attendance buildup far outweigh the nominal number of gap effect fans which may be lost. Victory for these teams is optimal. However, once the time gap reaches the extreme of many decades, the percentage of fan base represented by these effect fans becomes important for the team franchise.

**Discussion**

The time gap between World Series victories for the vast majority of MLB teams is relatively modest. For such teams, the gap effect will be trivial and can be ignored. However, for teams which have developed an extreme time gap since a World Series victory, the gap effect creates a perverse incentive against winning. The Chicago Cubs have not won a World Series in over a century. The gap effect helps explain strong game attendance for the Chicago Cubs despite the lack of a championship. The potential loss of the gap effect and its impact on attendance could make winning the World Series a financial sub-optimal outcome for the Chicago Cubs in terms of attendance.

Management of sport organizations must understand and coordinate the diverse interests of the many units within the team. While winning might be an organizational goal, marketing and promotional units are responsible for maintaining attendance. In the instance of the Chicago Cubs, the organizational goal of winning is potentially at odds with the incentives facing marketing/promotions.
Conclusion

MLB teams are in the business of engaging fans and encouraging them to purchase tickets and attend games. A model is constructed to examine various elements which influence fans’ decisions to attend games. The ability of a team to win and make playoff appearances are large catalysts of fan attendance. The average benefit of a playoff appearance afforded all MLB teams is an average increase in attendance of approximately 140,966.

Moving a club to a new city yields an increased average attendance of approximately 619,362 while controlling for city population differences. However, the benefits only last one season as lingering effects are insignificant. In markets with multiple teams, it is determined that a cannibalization effect exists which causes a modest reduction in yearly attendance of approximately 50,300 per additional team in the market.

Finally, it is found that the gap in years between World Series wins is a significant and positive influence on attendance. While few in early years, the gap effect grows marginally at a rate of approximately 1,769 fans per year. A World Series victory has the potential to negatively impact this effect. Identifying this influence helps explain fan behavior of teams with historical lack of championships. The motivation from the fan’s perspective might be different. However, the fans are shown to have similar attendance behaviors.

References


Address for correspondence:

Timothy E. Zimmer,
School of Business, University of Indianapolis.
1400 East Hanna Avenue, Indianapolis, IN 46227
Email: zimmert@uindy.edu