Interleague Play And The Big Mac Attack: Estimating The Within-Season Demand For Major League Baseball

Thomas H. Bruggink, (Email: bruggint@lafayette.edu), Lafayette College
Colin Roosma, Lafayette College

1. Introduction

*I*n most U.S. professional sports the attendance at games is dictated by the capacity of the stadium. In men's basketball, football, and hockey, the majority of the games are sold out, or nearly so. Estimating the demand curve for these sports is impossible. But in baseball the games are generally not sold out. There are eighty-one home games played in stadiums that have an average capacity of 40,000. It is unusual to have a sellout because there are so many games compared to football, basketball, and hockey. Men's basketball and hockey play half as many games in an arena with half as many seats. In football games there are only eight home games. Thus game-by-game variation only occurs in baseball, enabling the estimation of a within-season demand model.

Estimation of game-by-game attendance allows the testing of special attendance factors. In this study the attendance factors of interest are the drawing power of star players (especially homerun hitter Mark McGuire) and particular scheduling circumstances. The scheduling of interleague games (between the separate National and American League teams) is still a league experiment and deserves close scrutiny with respect to its attendance impact. Claims by league officials that interleague play increases attendance have not been subject to standard statistical controls. The opportunity cost of an interleague game was examined by economists before baseball adopted this experiment to increase attendance. As noted in Bruggink and Eaton (1996),

...each game a team plays against a team from the other league reduces the number of games it will play against teams in its own league, which have more impact on its division standings and qualification for post season play. Depending on the schedule, for example, a NL team might substitute a game against a weak AL team starting an unsuccessful pitcher for a game against a strong NL team in its own division starting a strong pitcher. In this case the novelty of interleague play is partially offset by the reduced quality of the game. Although attendance will likely increase in any case, it would increase more if above average teams from the league could always be substituted on the schedule in place of below average opponents from a team's own league. This obviously cannot be true of every interleague contest. (p. 28)

2. Literature Review

Although no attendance studies exist on women's basketball, there are several works on baseball within-season attendance in the 1990s. Knowles, Sherony, and Haupert, (1992), Bruggink and Eaton (1996), Rascher (1999), and Kahane and Shmanske (1997) all estimated within-season demand to test hypotheses on such attendance factors as uncertainty of outcome, quality of teams, television broadcasts, etc. Only recently have the issues of interleague play emerged (Butler, 2001). Butler found that interleague games experience only a 7% increase in attendance. The interesting aspect of these results is that the overall increase in attendance is due mainly to only a few matchups that draw a large attendance. Of the 30 teams, only 7 caused attendance to rise significantly when they were the visiting teams. While attendance rose overall due to interleague play, it is heavily weighted on only a few games. Butler concluded that revenue from ticket sales for interleague games rose "only about one-half of one percent" (Butler, p. 11) from what it would have been had the teams played their regular opponents.

_____

Readers with comments or questions are encouraged to contact the authors via email.

3. Empirical Model

Previous studies on the demand for baseball reveal the opportunity to augment traditional microeconomic theory to reflect day-to-day differences in the quality of the good that is offered. Hence a large number of predictor variables are included. :

$$ATTEN_{ij} = f(M, A, S)$$

where $ATTEN_{ij}$ = attendance at game i for team j, M is a vector of standard microeconomic variables, A is a vector of augmented demand factors, and S is a vector of special factors. The double log functional form applies to the standard microeconomic variables, but some of the additional variables will not be logged. This includes variables stated as percentages, all dummy variables, and counting variables such as the number of All-Stars.

Dependent Variable

        ATTEN$_i$ *attendance at game i for team j*

Core Variables for Demand

POP	population of home city for team j
LOWRES$	the lowest price for reserved seat for team j ($)
AGE	number of years the team j has been in existence in present city
INCOME	per capita income of population of team j's city

Augmented Demand Variables

WP	current season winning percentage of team i before game j
PREVWP	winning percentage of team j in 1999
WPV	current season winning percentage of visiting team before game i
PREVWPV	winning percentage of visiting team in 1999
ALLSTAR	number of all-stars from past three seasons (97-99) on team j
ALLSTARV	number of all-stars from past three seasons (97-99) on visiting team
STADIUM	dummy variable equals 1 if team z has a brand new stadium; 0 otherwise (baseball only)
TURF	dummy variable equals 1 if team z has artificial turf; 0 otherwise (baseball only)
WEEKEND	dummy variable equals 1 if game played on a Friday, weekend. or holiday
JULY-ON	dummy variable equals 1 if game played in July; August, September or October

Special Interest Variables

INTER	dummy variable equals 1 if game is an interleague game between teams not in the same metropolitan area; 0 otherwise
INTERRIV	dummy variable equals 1 if game is an interleague rivalry game (game between two teams in the same metropolitan area); 0 otherwise
BIGMAC1	dummy variable equals 1 if Mark McGuire is not on the disabled list for home game I
BIGMAC2	dummy variable equals 1 if Mark mcQuire is not on the disabled list when the St. Louis Cardinals are the visiting team for home game i

3.1. Sample Selection

In reduce the data requirements, we include only the sixteen National League teams rather than the full thirty major league teams. Despite the fact that the 16 teams play 81 home games each (a total of 1296), there are only 1200 observations because the first few games of each team are not included in the model. This was done to give the variables that depend on results from previous games some validity. It also excludes from the model the abnormally high attendance regularly seen on opening day.

4. Regression Results for Baseball Attendance Model

Ordinary least squares regression provides the results shown below. The double log functional form provides elasticities for the variables population, income, and price. The independent variables that are not logged include all dummy variables, variables that are measured as percentages, and such counting variables as the number of All-Stars on each team. The interpretations given these latter coefficients will be log-linear.[1]   (t-ratios given in parentheses)

$$\text{ln ATTEN} = 3.2 + \underset{(4.6)}{0.09 \text{ ln POP}} - \underset{(-19.9)}{0.67 \text{ ln LOWRES\$}} + \underset{(4.6)}{0.53 \text{ ln INCOME}}$$

$$+ \underset{(5.0)}{0.014 \text{ PREVWP}} - \underset{(-0.22)}{0.004 \text{ PREWPV}} + \underset{(7.0)}{0.01 \text{ WP}} + \underset{(2.1)}{0.003 \text{ WPV}} + \underset{(3.6)}{0.045 \text{ ALLSTAR}}$$

$$+ \underset{(3.0)}{0.022 \text{ ALLSTARV}} + \underset{(24.2)}{0.005 \text{ AGE}} + \underset{(9.4)}{0.41 \text{ STADIUM}} - \underset{(-13.1)}{0.38 \text{ TURF}} + \underset{(15.2)}{0.24 \text{ WEEKEND}}$$

$$+ \underset{(2.5)}{.04 \text{ JULY ON}} + \underset{(1.1)}{0.03 \text{ INTER}} + \underset{(1.0)}{0.08 \text{ INTERRIV}}$$

$R^2 = 0.71$ adj $R^2 = 0.70$ $F = 177.6$

Fans are more concerned with playing ability of their own team before worrying about the strong or weak team that is coming to town. For game-by-game winning, a 1 percent increase in winning for the home team is associated with 1 % increase in attendance. For every 1 percent increase in winning for the visiting team, attendance rose only .3 %. For the previous season, each 1 percent increase in winning by the home team increases by 1.4% while the same increase in visitor winning decreases attendance, although the coefficient is not statistically significant.

The two core variables associated with the fans' expenditures and income are the ticket price variable, LOWRES$, and INCOME. LOWRES$ is negative as expected and significant.   The model says that for each one percent increase in the lowest reserve seating price, attendance will decrease by 0.67%. The demand for baseball is price inelastic. The model suggests that ticket prices could be increased to increase revenue. This apparent under-pricing by clubs is commonly found in baseball attendance studies and is explained by Salant (1992).   INCOME also behaved as expected, as it was positive and significant. Each one percent increase is the fans' per capita income, attendance will rise by 0.53%.

AGE indicates the length of time each team has been in existence in its present city. There are teams that have been in their current city for over 100 years and there are others that have only been in existence for three years. The AGE variable shows if fans like to see new teams or if they enjoy the heritage that some of the older teams have created in their city.   For each year in existence in its present city, attendance at games rose by 0.5%. Over a long stretch of time, for instance the 100 years of the St. Louis Cardinals or Pittsburgh Pirates, the additional fan support is enormous.

ALLSTAR and ALLSTARV are two variables showing the star appeal of the participating teams. They represent the number of players on the roster of the home team and visiting team respectively that have been All-Stars at least once over the last three seasons. Every home team All-Star increases attendance by 4.5%, while visiting team All-Stars add 2.2%.

The STADIUM variable is a dummy variable that denotes whether a team is playing in a brand new stadium or not. The Arizona Diamondbacks and the San Francisco Giants enjoyed new stadiums during the 2000 season, The attendance at Diamondbacks and Giants games was 41% higher, *ceteris paribus*. .   While it was expected that STADIUM would be positive, it was not expected that its coefficient would be that large.   This shows why teams are constantly pushing their local governments for new stadiums.

TURF is a dummy variable that shows when games are played on artificial turf instead of natural grass. For games played on artificial turf instead of grass, attendance was 36% lower, *ceteris paribus*. Most stadiums have grass, but turf is perfect for dome stadiums where there is no sunlight, or multipurpose stadiums that share their facility with football. The seating configuration in a multipurpose stadium is not fan friendly to baseball. Furthermore, when baseballs are hit on the ground in a turf stadium, it rolls much faster than it would on natural grass, which makes for slightly unnatural play of the game. Another big negative is the artificial color.

The WEEKEND variable   is used to separate weekdays from weekend days to see how many more fans attend weekend or holiday games. The coefficient shows a 24% increase for games played on Fridays, Saturdays, Sundays, and holidays. The JULY ON variable captures the baseball season when it does not have to compete with championship series in professional basketball and hockey. In the months of July, August, September, and October, attendance rises by 4%.

The special interest variables behaved as expected except for the two interleague play and the promotional variables. , Both INTER and INTERIV were positive   While it was originally thought that interleague games of any kind would increase attendance because their novelty An explanation for the lack of statistical significance may be that a majority of interleague games are played between two teams that the fans have no special interest in watching. But even the INTERIV teams did not show statistical significance.[2]

[NOTE: the two variables for Mark McGuire are not yet included in the empirical results—this will be corrected shortly]

4.1. Residual Analysis

The residual plot of the full model shows small residuals for almost every observation. Only four of the 1,200 residuals are outliers. Three of these residuals show that attendance was much lower than usual, while one showed attendance to be much higher. The first abnormally low residual is from Monday, May 22, 2000 when the Milwaukee Brewers hosted the Houston Astros and attendance for this game was a mere 3,913. The game took place on a Monday, which may have lowered attendance. This was a one game series between the two teams and attendance may have been so low due to the Astros' underachieving play. The Astros finished the 1999 season with a winning percentage of .599, but entered this game at .366. Fans in Milwaukee may not have thought it to be worthwhile to watch the Astros play.

The second abnormally low residual is from Thursday, June 1, 2000 when the Chicago Cubs hosted the Atlanta Braves and attendance was 5,267 fans. This game was the third of a three game series with the Braves, in which the Cubs had lost the previous two games. Attendance may have been low because of the Cub's recent poor play or because the game was held on Thursday. After being beaten twice by Atlanta, Chicago fans may not have wanted to watch it happen a third time.

The third low residual is from Thursday, June 8, 2000 when the New York Mets hosted the Baltimore Orioles and attendance was only 9,540. This being an interleague game certainly contributed to the model showing INTER to be statistically insignificant. The game was the third of a three game series and attendance may have been low due to the Orioles poor play. At this point in the season, the Orioles' winning percentage was below .500 and Mets fans may not have cared to see them play.

The final abnormal residual is far above the average. It is much easier to pinpoint the reasoning for this residual as it is from Thursday, September 28, 2000 when the Milwaukee Brewers hosted the Cincinnati Reds and attendance was a phenomenal 56,354. This was the last home game of the season for the Milwaukee Brewers, which meant that it was the last game ever at County Stadium. The Brewers would move into a new stadium for the 2001 season, so fans made sure to attend the final game held at County Stadium.

5. Conclusion on Baseball Attendance Model

In the beginning of the study, the special interest variables were of importance to the model because they represented new ideas that could affect attendance. The INTER and INTERRIV variables, designated for interleague play and interleague rivalries respectfully, were placed into the model to show how emerging interleague play boosted attendance. Although the media reported large attendance at many interleague contests, these casual observations do not hold other attendance variables constant. When the quality of team, the day of the week, etc., are held constant, the statistical significance of the interleague factor drops off. In conclusion, this study shows that interleague play are over rated as attendance factors.

Playing conditions are another option for teams attempting to boost attendance. First, new stadiums are a big draw. Second, artificial turf has been shown in the model to lower attendance. Its removal and the addition of natural grass can help a team augment attendance. This has been evident during the first month of the 2001 season. The stadiums for the Cincinnati Reds and Philadelphia Phillies used artificial turf in 2000, but have changed for 2001. In Cincinnati, natural grass has replaced the unnaturally appearing Astro turf. In Philadelphia, the unnatural looking turf has been replaced with an alternative style of artificial turf. This new turf looks and acts just like grass, which adds to the aesthetic value and the play of the game. 📖

Bibliography

1. Bruggink, Thomas H., Eaton, James W. "Rebuilding Attendance in Major League Baseball: The Demand for Individual Games" in *Baseball Economics: Current Research,* Larry Hadley, Elizabeth Gufstafson, and, editors Praeger Press (1996)
2. Butler, Michael R. "Interleague Play and Baseball Attendance." Presented at the Midwest Economics Association March 2001.
3. Gaquin, Deirdre A., DeBrandt, Katherine A., eds. *2000 County and City Extra: Annual Metro, City, and County Data Book*. Lanham, MD: Bernan Press (2000)
4. Kahane, Leo, Shmanske, Stephen. "Team Roster Turnover and Attendance in Major League Baseball." *Applied Economics* 29 (1997) 425 - 431.
5. Knowles, Glenn, Sherony, Keith, Haupert, Mike. "The Demand for Major League Baseball: A Test of the Uncertainty of Outcome Hypothesis." *The American Economist* 36 (1992) 72 - 80.
6. Rascher, Daniel. "A Test of the Optimal Positive Production Network Externality in Major League Baseball." *Sports Economics: Current Research* Larry Hadley, Elizabeth Gufstafson, and John Fizel, editors Praeger Press (1999).

Endnotes

[1] Special adjustments for the log-linear interretations of the dummy variable coefficients have not been made, so consider the coefficient interpretations as approximate.

[2] Although not statistically significant at conventional levels, each t-ratio was over 1 and thus these variables did contribute to the overall fit of the model.

Notes