The Attendance Boost Is Over-Rated For Interleague Baseball, And The Big Mac Attack Is A Hit On The Road: All This And More From The Within-Season Demand Model For Major League Baseball

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

1. Introduction

*T*he theory of microeconomic demand is seldom estimated with a rich set of data, yet there is no shortage of statistics in professional sports. Using the sports industry enables economists to extend traditional theory of demand to include products that change daily (the visiting team) as well as the circumstances of their consumption (e.g., the day of the week). In fact the home games of professional sports are never identical from one game to the next. This variation allows one to estimate the quantity response to each of a multitude of factors.

Estimation of game-by-game attendance in major league baseball allows the testing of special attendance factors. Of all sports baseball is particularly well suited for this daily demand curve because there are so many games, and the game-by-game variation in attendance is by far the greatest in all sports. In this study the attendance factors of interest are interleague games and the drawing power of star players (especially homerun hitter Mark McGuire). The scheduling of interleague games between the separate National and American League teams is still a league experiment (it started in 1997) 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. "What game would have taken place had it not been converted into an interleague game?" Economists noted the opportunity cost of an interleague game even 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. Background

In 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 substantive game-by-

game variation only occurs in baseball, enabling the estimation of a within-season demand model.

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. The most recent study we found was by Butler (2001) for the 1999 season. He found that interleague games experience only a 7% increase in attendance for an otherwise identical game, far short of the 20% claimed by league officials who did not adjust for the day of the week, etc. 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.

This study will look at the 2000 season, the fourth year of interleague play. Earlier studies (such as Butler's) might have been influenced by the early excitement that interleague play offered. Fans were treated to seeing teams from the other league at their home stadium—an event that would never have occurred except in the rare event that the two teams were in the World Series.   It is likely that the novelty effect of interleague games has diminished by now, and a more true assessment of its net impact can be measured. Unlike Butler we will not employ a fixed effects model with dummy variables for each team, but rather follow the more common approach of using market and team characteristics to specify the demand for baseball in each city.

The home attendance figures for each National League team are shown in Table 1. Taking a cue from Butler (2001) we divide the interleague games into two groups: those played against the American League "rival" in the same city, and those against all others.

Table 1 Attendance By Team In 2000

TEAM	Average Attendance in the Sample	Inter-City Interleague Average	% Change	Intra-City Interleague Average	% Change
Houston Astros	37,656	38,314	2%		
Atlanta Braves	40,357	43,994	9		
Milwaukee Brewers	19,945	18,447	-7.5		
St. Louis Cardinals	41,157	42,603	3.5		
Chicago Cubs	34,854	38,198	9.6	38,917	12
Arizona Diamond Backs	36,535	35,286	-3.4		
Los Angeles Dodgers	36,969	36,131	-2.3	43,766	18
Montreal Expos	10,935	18,068	65		
San Francisco Giants	40,930	40,930	0	40,930	0
Florida Marlins	15,377	14,525	-5.5		
New York Mets	36,285	28,442	-21.6	54,194	49
San Diego Padres	29,858	26,958	-9.7		
Philadelphia Phillies	20,013	22,915	14		
Cincinnati Reds	22,457	21,948	-2.3		
Colorado Rockies	31,650	40,070	26.6		

The average attendance per game for the 1200 games in the sample is 31,031. For the 126 interleague games played in National League stadiums the attendance is 32,303, an increase of 4% (Note: this would be slightly

higher if we took the interleague games out of the overall average). However, the twelve intra-city interleague games averaged 44,452, more than 13,000 above the season average. Intra-city interleague games all have high attendance yet the non-rival interleague games actually show declines for several teams compared to their season average. At first blush the intra-city interleague games seem a resounding success while the rest of the games the fans appear indifferent if not hostile. However, the increases must be tempered by the fact that 66% interleague games (83 of 126) are played on the more popular days of Friday, Saturday, and Sunday, whereas the regular schedule only places half the game on weekends (e.g. the Rockies played 38 of the 75 games in the sample on the weekend). Furthermore interleague games are all played in June and July, the two most popular months for viewing baseball (no other professional team sports are engaged, with the exception of women's basketball). These confounding factors prevent any conclusion about interleague attendance boosts until proper controls are used. Only in the full regression model will we be able to discern attendance differences due solely to the status of interleague play.

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 the additional variables will not be logged because they are percentages, dummy variables, or counting variables such as the number of All-Stars and the age of the team.

Dependent Variable

$ATTEN_{ij}$ - 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
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, Saturday or Sunday
HOLIDAY – dummy variable equals 1 if game is played on Memorial Day (observed), July 4, or Labor Day.
MAY - dummy variable equals 1 if game played in May
JUNE - dummy variable equals 1 if game played in June
JULY - dummy variable equals 1 if game played in July

AUGUST - dummy variable equals 1 if game played in August

SEPTOCT - dummy variable equals 1 if game played in September or the first week of October

PENRACE - dummy variable equals 1 if the top two teams in the division are in a division race during the last six weeks of the season

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 McGuire is not on the disabled list when the St. Louis Cardinals are the visiting team for home game i

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.

Two teams had new stadiums in 2000: the San Francisco Giants and the Houston Astros. Four teams had artificial turf: the Cincinnati Reds, the Philadelphia Phillies, the Pittsburgh Pirates, and the Montreal Expos.

Due to injuries, Mark McGuire was on the disabled list from July 7 to September 7. Before he was injured he generally started in every game, but after coming off the disabled list he did not regularly start. Although he is listed as participating in 89 games, we have him marked for only the 69 games in which he started.

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)

The overall fit and individual fits of this regression model are very good. All signs are correct and statistically significant except for   PREVWPV, BIGMAC1, INTER and INTERRIV. There is little evidence of multicollinearity or heteroskedasticity. To accommodate the time series aspect of the data a Cochrane-Orcutt correction for serial correlation was run but there were no essential differences in the results. The same can be said for running the regression without the San Francisco Giants, a team that sold out every game. In both runs the signs and statistical significance of each coefficient was unchanged. The magnitudes of the coefficients did change, of course, but not by much for the coefficients that were statistically significant. [2] A short discussion on residual outliers will follow.

We will now discuss each variable's impact on attendance. The three microeconomic core variables, all logged < are ticket price, income, and market size. 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 0.68%. 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).[3] INCOME also behaved as expected, as it was positive and significant. For each one percent increase in the

fans' per capita income, attendance will rise by 0.47%. For each 1%increase in population, attendance rises 0.1%.

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 percentage point increase in winning for the home team is associated with 1.0% increase in attendance. For every 1 percentage point increase in winning for the visiting team, attendance rose only 0.13 %. This last coefficient is not only small but also statistically insignificant. For the previous season, each 1 percent increase in winning by the home team increases by 1.3% while the same increase in visitor winning decreases attendance by 0.013%, although the coefficient is not statistically significant.

---

ln ATTEN = 3.71 + 0.103 **ln POP** - 0.68 **ln LOWRES$** + 0.47 **ln INCOME**
                        (4.89*)   (-19.9*)   (4.06*)

+ 0.013 **PREVWP** – 0.00013 **PREVWPV** + 0.010 **WP** + 0.0013 **WPV** + 0.045 **ALLSTAR**
     (4.54*)   (-0.08)   (7.28*)   (1.03)   (3.44*)

+ 0.024 **ALLSTARV** + 0.005 **AGE** + 0.44 **STADIUM** - 0.38 **TURF** + 0.25 **WEEKEND**
     (3.44*)   (23.2*)   (9.61*)   (-13.3*)   (15.7*)

+ 0.16 **HOLIDAY** + 0.045 **MAY** + 0.107 **JUNE** + 0.186 **JULY** + 0.085 **AUGUST**
    (2.82*)   (1.44***)   (3.28*)   (5.63*)   (2.58*)

+ 0.045 **SEPTOCT** - 0.024 **INTER** +   0.0038 **INTERRIV** – 0.048 **BIGMAC1**
     (1.44***)   (0.05)   (-0.95)   (0.95)

+ 0.21 **BIGMAC2** + 0.047 **PENRACE**
    (4.02*)   (1.30***)

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

---

   significant at a1% level in a one-tail test
  ** significant at a 5% level in a one-tail test
*** significant at a10% level in a one-tail test

We will now discuss each variable's impact on attendance. The three microeconomic core variables, all logged< are ticket price, income, and market size. 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 0.68%. 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).[4] INCOME also behaved as expected, as it was positive and significant. For each one percent increase in the fans' per capita income, attendance will rise by 0.47%. For each 1%increase in population, attendance rises 0.1%. 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 percentage point increase in winning for the home team is associated with 1.0% increase in attendance. For every 1 percentage point increase in winning for the visiting team, attendance rose only 0.13 %. This last coefficient is not only small but also statistically insignificant. For the previous season, each 1 percent increase in winning by the home team increases by 1.3% while the same increase in visitor winning decreases attendance by 0.013%, although the coefficient is not statistically significant.

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%.

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 2.4%.

STADIUM is a dummy variable that denotes a team is playing in a brand new stadium. The Houston Astros and the San Francisco Giants enjoyed new fields of play during the 2000 season, and attendance was 44% 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. Because the Giants sold out every game, the regression was re-run without the team. This makes STADIUM a location dummy for Houston. The new coefficient was essentially the same as were all other coefficients and goodness-of-fit measures.

TURF is a dummy variable marking games are played on artificial turf instead of natural grass.. 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. For games played on artificial turf instead of grass, attendance was 38% lower, *ceteris paribus.*

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 25% increase for games played on Fridays, Saturdays, and Sundays. HOLIDAYS added 16% to attendance for Memorial Day, July 4, and Labor Day. The monthly dummies showed increases in attendance compared to April. The coefficient for May was small and statistically significant only at a 10% level. June and July showed large increases (12% and 20%) and then the monthly effects tapered off with 10% for August and 5% for September and the first week of October. Also in the late part of the season, division races develop that can add to attendance. For the year 2000 this occurred between the New York Mets and the Atlanta Braves, and attendance rose an additional 7% for these teams' home games during the last six weeks of the season.

In his second year after demolishing Roger Maris' long-standing season home run record, Mark McGuire spent considerable time on the disabled list. An unfortunate injury allows a natural experiment for those interested in assigning an attendance coefficient to his appearance in the starting lineup. He was in the starting lineup sixty-nine times in the 162 game season. Thirty-nine of those were at home, but surprisingly there was no attendance impact. The coefficient was small, negative, and statistically insignificant. He was a hit on the road, however, as attendance increased 21%, ceteris paribus. Perhaps the home fans, assuming he would be around all season, made no special effort to see him play early in the season (before he was injured). Whereas when the Cardinal team was visiting the hometown fans in other cities had greater urgency, knowing that even a healthy McGuire would be coming back once or at most twice more. This home/away dichotomy of impact for a star player is not without precedent. Hall of Fame pitcher Nolan Ryan had a large attendance impact at opposing teams' stadiums during his last year of baseball but no impact at home (Bruggink and Eaton, 1996).

It was originally thought interleague games of any kind would increase attendance because their novelty. But the coefficient of the inter-city interleague games is a negative 2.4%, although it is statistically insignificant. An explanation for the lack of 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 intra-city interleague games, although positive at 0.4%, did not show statistical significance. When the Giants were removed from the regression the coefficient for intra-city interleague games increased to 4%, but it was still not statistically significant.

The increases in attendance for intra-city interleague games shown in Table 1 melts away when it is recognized that 10 of these 12 games were played on weekends, and all were played in July. When you factor in the 25% weekend boost and the 20% July boost that normally occurs, the 10,000 more fans for the interleague rivals are near-

ly accounted for by simply denoting the day and month. Schedule these same dates with National League teams of comparable quality, and you will have essentially the same attendance.

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 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. 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. 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

This study confirms the important role that individual players make on attendance, particularly when the superstar is on the visiting team. Mark McGuire's shattering of the Roger Maris season home run record made him quite a celebrity. The attendance at games when the Cardinals were the visiting team and Mark McGuire was not on the disabled list show attendance rising 21%, ceteris paribus.

Playing conditions are another option for teams attempting to boost attendance. First, new stadiums are a big draw. Second, artificial turf is associated with substantially lower attendance. Each park that had artificial turf in 2000 was also a multi-purpose stadium of some vintage. No doubt this contributed to the lower demand. Nonetheless improvements in the turf can help a team augment attendance. 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 improves the aesthetic value and the play of the game.

The INTER and INTERRIV variables, designated for interleague play and interleague city 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. This study shows that interleague play is over-rated as an attendance factor for the National League in 2000. In fact it may be time to re-consider the wisdom of such a policy. Observed increases in attendance are completely explained by the day of week and month of season attendance factors for an otherwise identical quality team. The decision to continue the interleague game experiment should be based on factors other than boosting attendance.

References

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.   Halvorsen, Robert, and Raymond Palmquist "The Interpretation of Dummy Variables in Semilogarthmic Equations," American *Economic Review* 70:474-475 (1980)
5.   Kahane, Leo, Shmanske, Stephen. "Team Roster Turnover and Attendance in Major League Baseball." *Applied Economics* 29 (1997) 425 - 431.
6.   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.
7.   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 interpretations of the dummy variable coefficients have not been made, so consider the coefficient interpretations as approximate. The adjustments for "relative effect" are calculated as exp(estimated coefficient) – 1. In most cases the differences are small. For details see Halvorsen and Palmquist(1980).

[2] Regression results from additional specifications are available from the authors. There was very little sensitivity to a variety of specifications with the exception of the Stadium and Turf variables. When these were dropped from the equation the R-squared fell to 55% and several coefficients changed signs and significance.

[3] Among the theories proffered are: owners are not profit-maximizers but sportsmen, a lower ticket price encourages greater attendance and thus greater expenditure on concessions and parking, a lower ticket price serves as an insurance policy to keep the fans loyal year after year despite fluctuating performance.

[4] Among the theories proffered are: owners are not profit-maximizers but sportsmen, a lower ticket price encourages greater attendance and thus greater expenditure on concessions and parking, a lower ticket price serves as an insurance policy to keep the fans loyal year after year despite fluctuating performance.