Journal of Sports Economics
http://jse.sagepub.com/

Explaining Game-to-Game Ticket Sales for Major League Baseball Games Over Time

Elise M. Beckman, Wenqiang Cai, Rebecca M. Esrock and Robert J. Lemke

Journal of Sports Economics published online 30 June 2011
DOI: 10.1177/1527002511410980

The online version of this article can be found at:
http://jse.sagepub.com/content/early/2011/06/18/1527002511410980

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
The North American Association of Sports Economists

Journal of Sports Economics
000(00) 1-18

DOI: 10.1177/1527002511410980
http://jse.sagepub.com

Explaining Game-to-Game Ticket Sales for Major League Baseball Games Over Time

Elise M. Beckman[1], Wenqiang Cai[1], Rebecca M. Esrock[1], and Robert J. Lemke[1]

Abstract

Using data from more than 10,000 games from 1985 through 2009, the authors estimate the effect various factors have on attendance at Major League Baseball (MLB) games. As previously found in the literature, interleague and interleague rivalry contests are associated with higher attendances, but this relationship has been weakening over time. Contrary to some of the literature, the authors find that the likelihood the home team will win the contest is inconsistently estimated over time, lending little support for the uncertainty of outcome hypothesis. Generally the effect on ticket sales from many potential factors has generally been weakening over time.

Keywords

Major League Baseball ticket sales, nonlinear pricing, sporting event attendance, uncertainty of outcome hypothesis

Introduction

As originally presented by Neale (1964) and Rottenberg (1956), the financial viability of a sports league depends on attracting fans to the sport. This requires, in part, offering competitive contests as fans are unlikely to turn out to sporting events with

[1] Lake Forest College, Lake Forest, IL, USA

Corresponding Author:
Robert J. Lemke, 555 N. Sheridan Road, Lake Forest, IL 60045, USA
Email: lemke@lakeforest.edu

all but certain outcomes (called the uncertainty of outcome hypothesis [UOH] in the literature). Major League Baseball (MLB), a sports league that generates a substantial proportion of its revenue via local ticket sales (Zimbalist, 2003), in particular has an acute need to entice fans to the ballpark. For this reason, concern for "small market teams" in baseball, more so than for professional basketball or football teams, receives much attention. In 2001, reflecting on when he became the Commissioner of Baseball, Allan H. Selig testified to Congress that he would concentrate his efforts on addressing "competitive balance on the field and the economic stability of the clubs" (Selig, 2001).

This article makes an important contribution to the literature by estimating ticket sales for MLB games by allowing the coefficient estimates to vary across seasons. Using 25 years of data from over 10,000 games, separate coefficient estimates for the empirical specification are obtained for each half decade plus the strike years of 1994 and 1995. The flexibility of the empirical model provides coefficient estimates that not only inform the reader on fan behavior at a particular point in time but also on how behavior has changed over time. When considering the point estimates of the factors thought to influence ticket sales, the main conclusion from this study is that the effect of most factors has waned over time. Factors that are unknown until shortly before the game is played (e.g., quality of starting pitching) have largely become unimportant, and even most factors that are known when the schedule is released in the off-season (e.g., whether the visiting team was a playoff team the previous year) have become less important recently.

Ticket Sales

Following each game, MLB records the game's official box score. Included in the box score is an attendance figure. Originally "attendance" was the turnstile count of the number of people watching the game. Eventually, teams started to report turnstile attendance along with tickets sold. Since the mid 1970s, however, MLB's report of box score "attendance" is actually tickets sold and not turnstile attendance. This is a subtle, though important, distinction. For the purposes of this article, the data are for tickets sold, but "attendance" will be used interchangeably with tickets sold for convenience.

Table 1 shows three important features of MLB ticket sales for a 20% sample of games over the last 25 years. (The data are formally introduced in the Data section.) First, average game attendance steadily increased from 1985 to 1994, dropped in 1995 following the strike, and steadily rose since the strike to exceed 33,000 by 2007. Second, the coefficient of variation of attendance (the standard deviation divided by the average) has been decreasing over time. This has come about mainly because average attendance has been increasing rather than from the standard deviation decreasing. The coefficient of variation must fall (to zero in the extreme) when attendance figures approach stadium capacity for all games. For MLB, however, this

Table 1. Attendance and Prices by Year

Year	Average Attendance	Standard Deviation of Attendance	Coefficient of Variation of Attendance (%)	Average Ticket Price (2008 dollars)	Average Fan Cost Index (2008 Dollars)
1985	22,032	11,837	53.7		
1986	21,710	10,909	50.2		
1987	23,118	11,411	49.4		
1988	24,597	11,339	46.1		
1989	25,492	11,811	46.3		
1990	25,644	11,160	43.5		
1991	27,341	11,872	43.4	$12.96	$118.84
1992	25,532	11,587	45.4	$14.24	$129.52
1993	30,582	12,630	41.3	$14.27	$135.55
1994	30,699	12,645	41.2	$15.39	$139.95
1995	25,805	11,569	44.8	$15.73	$137.84
1996	26,623	11,150	41.9	$15.58	$142.73
1997	29,091	12,356	42.5	$16.34	$143.54
1998	29,657	12,741	43.0	$18.30	$153.94
1999	29,782	12,610	42.3	$19.66	$159.41
2000	29,873	11,597	38.8	$21.23	$168.31
2001	30,905	11,746	38.0	$23.41	$180.81
2002	28,798	11,659	40.5	$22.54	$178.00
2003	28,954	11,170	38.6	$22.41	$173.24
2004	30,355	11,378	37.5	$23.27	$180.35
2005	31,775	10,796	34.0	$23.34	$180.79
2006	32,628	10,776	33.0	$23.95	$183.49
2007	33,686	10,712	31.8	$23.85	$184.07
2008	33,145	11,075	33.4	$25.57	$183.53
2009	30,850	10,346	33.5		

Note. Attendance figures are tickets sold and come from box scores reported by Major League Baseball. See the text for the sampling procedure. The coefficient of variation is the standard deviation of average game attendance as a percentage of the mean. Ticket prices and the fan cost index are from *Team Marketing Report*, 2007.

cannot be the most important factor as the great majority of games do not sell out.[1] Third, the cost of tickets and attending games has been increasing over time. Given that attendance has been increasing at the same time prices have been increasing, demand for tickets must have increased over the past 25 years. The potential causes of this increased demand are numerous, including an expanded sports media and culture (e.g., ESPN) and baseball being a luxury good. It is also possible that the entertainment of seeing a baseball game has changed. In particular, with stadiums being built at a rapid pace over the last 15 years (in 2009, 18 of the 30 MLB teams played in stadiums built since 1994), the experience of going to a game can now often include visiting a state-of-the-art stadium offering an array of food and

beverage options including restaurant dining, lounge seating with TV viewing, safe travel to and from the stadium, and scheduled children activities.

Given these patterns of increased attendance and ticket prices, it is unclear how the estimated coefficients from an empirical model predicting ticket sales should change over time. If all games sold out, then none of the potential explanatory variables would have any explanatory power as stadium capacity or a home-team fixed effect would fully capture attendance. Although MLB has not reached this point, greater attendance and more sellouts in general will likely mitigate the effect any potential explanatory variable has on attendance. Moreover, the role of prices should not be overlooked. The cost of seeing games has increased dramatically. As prices increase, it is less likely that going to a baseball game is a casual, spur-of-the-moment decision. As prices increase, fans are more likely to plan their entertainment carefully and budget appropriately. If so, fewer tickets would be sold on game day (i.e., the walk-up gate) and more tickets would be purchased well before game day (i.e., advanced ticket sales). Although the data do not allow us to separately identify the walk-up gate from advanced purchases, the empirical results will be used to try to isolate these two possibilities.

Estimation Strategy

Our empirical specification for estimating ticket sales for MLB games is

$$y_{i,h,t} = \beta_0 + \beta X_{i,h,t} + \mu_t + \nu_h + \varepsilon_i, \quad (1)$$

where $y_{i,h,t}$ is ticket sales to game i for home team h in year t, $X_{i,h,t}$ is a vector of game-team-year specific variables with coefficients corresponding to β, μ_t is a complete set of year dummy variables, ν_h is a complete set of home-team fixed effects, and ε_i is an error term with standard properties.

Before discussing $X_{i,h,t}$ more completely, several aspects of the estimation procedure warrant discussion. First, Equation 1 will be estimated separately for each half decade exclusive of the strike years.[2] Second, no adjustment is made for correlated error terms within series. In other studies, notably Rascher (1999) and Lemke, Leonard, and Tlhokwane (2010), a correction is made to the error structure to take into account that errors in attendance are likely correlated across games, because MLB schedules games in series (e.g., the Red Sox will host the Yankees for a three-game series to be played Friday through Sunday). In those studies, every or almost every game of the season is in the sample, and thus all games in most series are included in the analysis. In our sample, however, as we randomly sample roughly 20% of the games from each season, only 3.2% of the games in our sample represent a second game in the same series. Third, because some games sell out, the structure of the linear model is violated. To address this, each observation is right-censored at the home team's stadium capacity and a censored normal regression is estimated.

The list of potential explanatory variables to be included in X is long when looking at the literature. Because our data cover 25 years, data collection becomes an issue as some variables cannot be recreated back to 1985. For example, although Fortunato (2006), Lemke et al. (2010), and McDonald and Rascher (2000) all find that gameday promotional events or giveaways (e.g., batday) are associated with greater attendance, our empirical specification does not include gameday promotions as promotional dates are unknown for all but the most recent years. Other variables that the literature has considered in the past that we do not include in our reduced-form model are the starting pitchers' race (Hill, Madura, & Zuber, 1982; Rascher, 1999), winning streaks (Fort & Rosenman, 1998, 1999; insignificant in Lemke et al., 2010), game-time temperature (Lemke, et al., 2010; Meehan, Nelson, & Richardson, 2007; Rascher, 1999), and roster turnover from one season to the next (Kahane & Shmanske, 1997).[3]

As the empirical strategy includes home-team fixed effects, we also omit variables that are largely team specific. For example, whereas we do not include a league variable (American vs. National) because this effect is directly included in the fixed effect, we also do not include whether the home team plays in a classic stadium (Fenway, Wrigley, or Yankee Stadium) or in a city with two teams (Chicago, Los Angeles, and New York) because the point estimate would simply take a portion of each team's fixed effect. Similarly, because teams do not often change prices dramatically from season to season, home-team fixed effects largely control for monetary prices as well as for nonmonetary costs of attending games such as typical driving times to the stadium and other local attractions competing for entertainment budgets.[4]

Despite these omissions, the factors considered in the empirical specification are numerous, including scheduling (divisional games, interleague games, month and day of the game, and preseason information on both teams such as last year's record), the quality of the game-day matchup (the starting pitchers' current winning percentages and earned run averages, both teams' current win percentage and position in the playoff race), the probability of the home team winning the game (implied by the betting line), the age of the stadium, and the local average income level and unemployment rate.

Data

The data span 25 seasons, from 1985 through 2009. The data collection process was designed to extract a 20% random sample of games from each season starting on April 14th to avoid opening day effects.[5]

The following stratified random sampling process was repeated for five 5-year intervals starting in 1985. April 14th through September 30th, which is roughly the end of each season each year, is 170 days. These 170 days were divided into 35 consecutive five-day blocks (i.e., April 14–18, April 19–23, . . . , September 26–30).

Each day of each five-day block was then randomly assigned to a year in the 5-year interval. In this way, every date from April 14th to September 30th was assigned to 1 of the 5 years. All games played on the selected dates were sampled for the year. Differences in the sample sizes across the seasons are mostly because of the strike-shortened seasons (1994 and 1995) and the growth of the league (1993 and again in 1998).

For each game, pregame and box score (postgame) data were collected from the *Chicago Tribune* (1985–2009). Pregame data include the time of the game, the teams involved, each team's divisional and (starting in 1994) wildcard standing, and the Glantz-Culver betting line for the game. The box score reports the tickets sold (attendance) for each game.[6] The final sample includes 10,386 games. The descriptive statistics are reported in Table 2. The dependent variable in the analysis is game attendance, which ranged from a low of 1,632 to a high of 73,957 with an average of 28,516.

The probability that the home team will win the game is an important variable in the literature, as there is considerable empirical testing, with mixed results, of the UOH. The most straightforward strategy to test the UOH is to include the probability the home team will win and its square. Following the methods of Knowles, Sherony, and Haupert (1992) and Lemke et al. (2010), we use betting lines to infer the probability of the home team winning the contest. The average probability in the sample is 0.54, indicating that, on average, there is a slight home-field advantage.

Prior to 1994, each league was divided into two divisions—East and West. Starting in 1994, each league was divided into three divisions—East, Central, and West. Almost 42% of the games in the sample represent divisional matchups. Starting in 1997, MLB also started scheduling interleague play. As of 2009, depending on its division, each team plays between 12 and 18 interleague games each season out of a 162 game schedule. Of the 10,386 games, 6.3% represent interleague contests; however, none of these occurred before 1997. From 1997 onward, the sample includes 5,113 games, 658 of which are interleague games. MLB also schedules six games every year for 10 pairs of geographic interleague rivals. Throughout the article, these games are referred to as interleague rivalry games. The interleague rivalries are: Cubs versus White Sox, Reds versus Indians, Marlins versus Rays, Astros versus Rangers, Dodgers versus Angels, Brewers versus Twins, Mets versus Yankees, Giants versus Athletics, Cardinals versus Royals, and Nationals versus Orioles. Of the teams with an interleague rival, about one third of its interleague contests are rivalry games. Of the 658 interleague games in the sample, 134 are interleague rivalry games.

In addition to controlling for divisional and interleague rivalries, mapquest.com was used to determine the distance (in thousands of miles) from the home team's stadium to the visiting team's stadium. On average, the stadiums are 1,283 miles from each other.

Lowry (2006) catalogs information on all current and former MLB stadiums. Following the work of Clapp and Hakes (2005), we construct two dummy

Table 2. Descriptive Statistics: 1985-2009

Variable Description	*M*	*SD*	Min.	Max.
Box Score Attendance (Ticket Sales)	28,516	11,992	1,632	73,957
Probability the home team will win the game	0.542	0.078	0.2500	0.7949
Divisional game	0.417	0.493	0	1
Interleague game	0.063	0.244	0	1
Interleague rivalry game	0.013	0.113	0	1
Distance between teams (1,000s of miles)	1.283	0.854	0.0100	3.299
Stadium is at most 5 years old	0.160	0.367	0	1
Stadium is 6 to 10 years old	0.117	0.321	0	1
Home team's games back in the playoff race	6.479	6.993	0	48
Home team's current winning percentage	50.07	8.93	0	100
Home team qualified for the playoffs last year	0.216	0.411	0	1
Visiting team's games back in the division	7.863	7.914	0	49
Visiting team's current winning percentage	50.00	9.11	0	100
Visiting team qualified for the playoffs last year	0.217	0.412	0	1
Visiting team is the Boston Red Sox	0.034	0.180	0	1
Visiting team is the Chicago Cubs	0.035	0.184	0	1
Visiting team is the New York Yankees	0.036	0.186	0	1
Weekday game (Monday–Friday afternoon)	0.491	0.500	0	1
Weekend game (Friday night–Sunday night)	0.509	0.500	0	1
Game played in April	0.092	0.289	0	1
Game played in May	0.184	0.387	0	1
Game played in June	0.189	0.391	0	1
Game played in July	0.174	0.379	0	1
Game played in August	0.186	0.389	0	1
Game played in September	0.176	0.381	0	1

Note. There are 10,386 observations. See the text for data sources.

variables—one indicating that the home team's stadium has been in use for at most 5 years (16% of the sample) and one indicating that the home team's stadium has been in use for more than 5 years but at most 10 years (11.7% of the sample). We refer to these two classifications as "newly constructed" and "recently constructed" stadiums.

In order to capture increased demand for tickets when the home team is in playoff contention, divisional and league standings are used to calculate how far back the home team is in the playoff race at the start of play on each day in the sample. Prior to 1994, only the four division winners qualified for the playoffs. For these years, "games back in the playoff race" is identical to games back in the division. Starting in 1994, MLB included a single wildcard team from each league in the playoffs. From 1994 onward, therefore, "games back in the playoff race" is calculated as the lesser of games back in the division and games back in the wildcard race. For the entire sample, the home team is, on average, almost 6.5 games back in the playoff

race. To further control for fan interest in the home team, the empirical specification also includes the home team's current winning percentage and whether the home team qualified for the playoffs the previous season.[7]

To control for fan interest in seeing the visiting team play, the empirical specification includes the visiting team's games back in the division, current winning percentage, and whether it qualified for the playoffs the previous season. The specification also includes dummy variables for three visiting teams—the Cubs, Red Sox, and Yankees. These three teams top the list of average road attendance for the sample, with the Red Sox and Cubs attracting an average road attendance of nearly 32,000 and the Yankees attracting an average road attendance of over 33,000.

Time and calendar factors have long been recognized as being important determinants of attendance. Roughly half of all games in the sample were played Monday through Friday afternoon (weekday) with the other half being played Friday night through Sunday night (weekend).[8] The specification also includes the month in which the game was played. Only 9.2% of the games in the sample were played in April, because the sampling process begins with games starting on April 14th and because teams have more days off in April than during the other months to accommodate for potential bad weather. Fewer games are played in July due to the All-Star break, and fewer games are played in April and September due to the 1994–1995 strike.

Results

Six different sets of censored normal regression results are reported in Table 3. Each specification is for half a decade, as reported at the top of each column, with the exception that separate coefficient estimates are produced for the strike-shortened seasons of 1994 and 1995. Though not reported in the table, each specification includes a full set of year dummies and home-team fixed effects. Each variable enters the specification according to its units as reported in Table 2. Robust standard errors are reported beneath coefficient estimates. A pseudo adjusted R-squared is reported for each time period.

The results in Table 3 are for the reduced form model after eliminating nonsignificant variables from the regression. Variables removed from the specification on this condition were home and visiting pitcher win percentage, interaction terms with age of the stadium, monthly interaction terms with the home team's games back in the playoff race before July, the visiting team's games back for all months, macro variables on local income and unemployment, playing a double-header, and day-of-the-week and time-of-game interactions. Two sensitivity tests were performed on the reduced form model. First, there is no evidence of serial correlation across observations. In other studies that include every or almost every game of a season, serial correlation is an issue (see Lemke et al., 2010; Rascher, 1999) because MLB schedules teams to typically play each other in three game series. Not finding

Table 3. Censored Normal Regression Results

	Regression Models By Years					
	1985-1989	1990-1993	1994-1995	1996-1999	2000-2004	2005-2009
Probability the home team will win	−47,508.9	−132,034.4[a]	−33,846.4	10,223.5	−623.8	−41,808.4[b]
	−9,143.2	37,153.4	37,425.6	21,625.0	14,154.4	17,064.1
Squared	37,906.5	127,105.8[a]	26,264.5	−7,894.8	4,137.6	46,749.6[a]
	36,752.4	34,359.7	34,966.9	20,336.8	13,370.1	15,779.7
Divisional game	922.6[b]	500.4	−277.7	779.2[c]	517.8	512.4
	379.3	408.5	606.0	464.7	387.1	352.4
Interleague game				2,281.9[a]	1,100.2[c]	2,027.7[a]
				747.9	615.8	585.8
Interleague rivalry game				8,388.1[a]	4,505.2[a]	1,957.1[b]
				1,752.9	1,087.9	999.0
Miles (1,000s) between teams	−795.4[a]	−163.8	−1,073.0[a]	−99.2	−412.9[c]	−483.3[b]
	212.8	245.1	363.6	274.0	245.9	212.0
Stadium is at most 5 years old	1,690.3	9,282.0[a]	−8,206.8[a]	6,188.7[a]	6,489.1[a]	−1,841.4[a]
	1,423.3	1,327.2	2,472.1	1,322.1	733.9	676.1
Stadium is between 6 and 10 years old	560.3	7,002.3[a]	−11,082.7[a]	5,605.2[a]	2,251.4[a]	−739.6
	1,015.6	1,267.0	3,950.7	1,382.6	788.3	712.2
Home team's games back in playoff race	−185.5[a]	−113.9[a]	−175.6	−16.9	−85.3	−63.1
	71.9	81.2	163.4	94.7	83.7	88.0
and played in July	−241.1[a]	−230.1[a]	−73.6	−98.1	14.7	−133.6
	79.0	89.3	159.7	102.1	75.9	84.5
and played in August	−258.3[a]	−225.2[a]	−139.5	−92.7	−39.1	−169.4[b]
	72.2	83.8	173.8	89.6	73.5	80.6
and played in September	−191.1[a]	−250.6[a]	−365.7[b]	−128.4	−82.6	−156.6[b]
	70.3	79.8	184.9	90.3	72.7	79.4
Home team's current winning percentage	174.4[a]	216.3[a]	93.3[a]	114.4[a]	130.4[a]	88.8[a]
	27.0	30.2	36.3	34.7	33.1	32.6
Home team made playoffs previous year	2,576.1[a]	4,151.6[a]	5,591.7[a]	2,049.3[a]	2,886.4[a]	2,231.5[a]
	461.0	650.5	1,501.4	534.0	447.2	372.1
Visiting teams' current winning percentage	125.0[a]	84.1[a]	52.4[b]	100.9[a]	102.7[a]	81.7[a]
	18.3	21.9	25.1	22.8	19.0	18.6

(continued)

Table 3 (continued)

	Regression Models By Years					
	1985-1989	1990-1993	1994-1995	1996-1999	2000-2004	2005-2009
Visiting team made playoffs previous year	1,777.9[a]	2,394.5[a]	2,899.4[a]	1,376.2[a]	1,860.9[a]	1,076.2[a]
	419.4	481.0	1,060.3	422.9	368.7	325.0
Visiting team is the Boston Red Sox	4,852.1[a]	−265.2	2,241.3	1,023.3	3,553.9[a]	6,884.0[a]
	793.8	917.9	1,385.5	947.7	837.4	810.1
Visiting team is the Chicago Cubs	685.8	406.2	441.1	2,872.8[a]	7,636.5[a]	4,849.1[a]
	770.8	911.3	1,315.1	951.6	846.9	749.3
Visiting team is the New York Yankees	6,394.5[a]	1,665.5[c]	3,662.7[a]	4,152.4[a]	7,643.6[a]	10,041.7[a]
	777.9	894.5	1,307.6	1,051.8	859.7	854.2
Day game played Monday–Friday	−7,730.0[a]	−6,712.6[a]	−5,433.5[a]	−6,612.1[a]	−6,777.4[a]	−6,166.3[a]
	316.5	358.7	494.2	351.9	297.1	273.0
Game played in May	2,734.0[a]	1,919.3[a]	−910.7	1,594.1[b]	1,271.7[b]	1,799.0[a]
	638.8	687.7	990.5	659.3	595.4	564.3
Game played in June	6,318.0[a]	6,479.8[a]	2,053.1[b]	4,739.4[a]	4,562.2[a]	3,051.0[a]
	665.1	758.1	1,041.8	754.3	665.0	650.4
Game played in July	10,483.3[a]	11,976.7[a]	6,726.5[a]	6,247.7[a]	4,972.4[a]	5,549.4[a]
	779.9	893.9	1,115.3	835.8	718.7	657.1
Game played in August	10,909.8[a]	10,364.9[a]	5,660.1[a]	6,872.3[a]	4,929.6[a]	6,072.9[a]
	738.7	878.8	1,247.9	797.6	684.8	632.4
Game played in September	3,731.7[a]	4,402.1[a]	2,826.1[b]	5,477.8[a]	3,086.7[a]	3,823.9[a]
	760.8	891.2	1,353.2	863.7	700.6	653.9
Constant	17,215.3	35,269.3[a]	25,758.5 b	3,993.6	6,959.4	25,304.9 a
	10,592.1	10,555.7	10,460.8	6,488.2	4,557.0	5,218.7
Number of observations	1,927	1,616	669	1,669	2,223	2,282
Pseudo adjusted R^2	0.7036	0.7120	0.7675	0.6995	0.7836	0.8103

Note. The dependent variable is box score attendance, which is equivalent to ticket sales for each game. All specifications include year dummies and home-team fixed effects. Standard errors are reported beneath coefficient estimates.

[a] 1% significance level.

[b] 5% significance level.

[c] 10% significance level.

evidence of serial correlation in our sample is not surprising, however, as our sample contains 20% of each season's games and less than 4% of our sample is a second game in a series. Second, a variance inflation factor (VIF) test was performed on each regression. Naturally, the probability that the home team will win the contest and its square both have high VIFs as these variables are highly correlated by construction. Because testing the UOH in a way that can be compared to the literature is important, we keep both the linear and the squared term in the reduced form model. The only other variable with a VIF in excess of the standard baseline of 10 is the number of games the home team is back in the playoff race, which barely exceeds 10 in all six regressions. This too is unsurprising as it is highly correlated with the monthly interaction terms. Ultimately, we kept the home team's games back in the playoff race (and three monthly interactions) in the model as these coefficients across time are important to consider. The remainder of this section provides a brief discussion of the results by focusing on six issues important to baseball attendance—testing the UOH, interleague play, playing in new stadiums, the home team's playoff positioning, the attractiveness of seeing the visiting team, and the importance of scheduling. The next section then provides a general discussion of how the process of ticket sales has changed over the past quarter century.

The UOH

There is no support for the UOH in these results. In four of the six time periods, neither coefficient on the probability of the home team winning the game nor its square is statistically significant. In the two time periods in which the coefficients are significant, the estimates suggest that attendance is minimized when the home team's probability of winning is low (0.52 in 1990-1993 and 0.48 in 2005-2009) and attendance continuously increases as the probability of the home team wins increases thereafter. Contrary to what the UOH maintains, and opposite to what Knowles et al. (1992) and Rascher (1999) found for MLB in 1988 and 1996, respectively, the results here suggest that attendance is typically unrelated to the probability the home team wins the game or, if anything, fans purchase tickets to see a home team victory more than they purchase tickets to see a competitive contest.[9]

Interleague Play

As anticipated by MLB, there is strong demand to see interleague games. In the second half of the 1990s, attendance at interleague games was almost 2,300 higher than at comparable non-interleague games, and attendance at interleague rivalry games during this period was additionally almost 8,400 higher than that. Although the point estimate on interleague games is lower after 1999, the *p*-value associated with the test that all three coefficients equal 2000 is 0.4967 suggesting that interleague games have been consistently associated with about 2,000 additional tickets sold per game since the inception of interleague play.[10] The same pattern does not hold for interleague rivalry games. Whereas interleague rivalry games attracted 8,400 more fans

in the first 4 years of interleague play, rivalry games attracted only 4,500 more fans to such contests in 2000-2004, and less than 2,000 more fans to such contests in 2005-2009. These differences are statistically significant, suggesting that the allure of interleague rivalries seems to be wearing off.[11]

Stadiums

Recently, Clapp and Hakes (2005) found a significant effect on attendance (on the order of 32 to 37%) during the first year a stadium is open and a positive effect remaining for up to 10 years.[12] Our results are more nuanced. From 1985 to 1989, there is no significant effect on attendance from playing in a newly constructed or recently constructed stadium. From 1990 to 1993, about 9,300 more tickets are sold per game at a new stadium and about 7,000 more tickets are sold per game at a recently constructed stadium. Positive effects are present for the 1996-1999 and 2000-2004 time periods as well, though the predicted effects are smaller in both cases. The coefficients for 1994-1995 are negative, potentially due to the strike. Although new stadiums may invigorate fan support, publicly provided or subsidized stadiums may be even more likely to spurn fan support during a labor dispute between multimillionaire owners and millionaire players. The effect on ticket sales is also negative for newly opened stadiums in 2005-2009. Although this is unexpected at first pass, teams can choose to restrict supply in order to increase price when designing a new stadium. This likely happened in St. Louis in 2006 and for both New York teams in 2009 when all three teams opened a new stadium with lower seating capacity than was available in the previous stadium. For example, in our sample, both New York teams sold more than 7,000 fewer tickets per game in 2009 (the first year of a smaller new stadium) than they did in 2008 (the last year of a larger old stadium).

Home Team's Playoff Positioning

The literature has consistently found that attendance is lower the further back the home team is in the playoff race. Although the results in Table 3 generally support this conclusion, there are differences over time. In particular, from 1985 to 1993 fewer tickets were sold the further the home team was out of the playoff race. Although the point estimates remains negative from 1994 through 2009, the estimates are smaller in absolute value and insignificant during these years. A similar pattern holds for the home team's playoff positioning in July and August. The only exception to the pattern is the interaction term in September for which the point estimates are statistically significant in four of the six time periods. Similar to playoff positioning, the literature consistently finds that attendance increases with the home team's current-year winning percentage. Our results support this conclusion, but Table 3 also shows that the magnitude of the effect has been generally decreasing over time. Finally, ticket sales are consistently higher for teams that qualified for the playoffs the previous season. Although this effect peaked during the first half of the

1990s, the effect is consistently strong, in excess of 2,000 more tickets being sold be game, for the entire quarter century.

Attractiveness of the Visiting Team

Because the visiting team's playoff positioning was consistently insignificant across all time periods, it was removed from the model. The visiting team's quality as measured by its winning percentage, however, is correlated with ticket sales. As with the home team's win percentage, this effect is consistently statistically significant, though it has generally been falling with time. Ticket sales are also higher when the visiting team qualified for the playoffs the previous year, with the effect ranging from a low of almost 1,100 more tickets being sold per game for 2005-2009 to a non-strike year high of almost 2,400 more tickets being sold per game for 1990-1993.

The Cub, Red Sox, and Yankees have attracted the greatest crowds as the visiting team over the last 25 years. The results in Table 3, however, show that each team's effect on ticket sales is different. The Yankees most consistently attracted large crowds when they were the road team. This is especially true in the 2000s when Yankee road games attracted over 7,600 additional fans per game on average in the first half of the decade and more than 10,000 additional fans per game on average in the second half of the decade. The Red Sox increased attendance by over 4,800 as the road team in the second half of the 1980s, but this effect disappeared in the 1990s only to return in the 2000s. This pattern mirrors their World Series appearance in 1986 and their World Series championships in 2004 and 2007. The Cubs, on the other hand, were not associated with a greater road attendance until the second half of the 1990s.

Scheduling

Compared to weekend games, ticket sales to weekday games are consistently lower on average. Although the point estimates fluctuate from 5,433 to 7,730, this effect is consistent over time. Finally, as expected, the estimated coefficients on the monthly dummy variables show that attendance increases in May compared to April, increases more in June, and then peaks in July and August before retreating to May or June levels in September. These patterns are consistent across the years.

Discussion

The nature of ticket sales to baseball games has changed over the past quarter century. Except for the strike years, attendance has steadily risen while the coefficient of variation of attendance has fallen (see Table 1). Relatively speaking, therefore, MLB has increased attendance while distributing attendance more evenly across all games on the schedule. It is unclear whether team owners have a reason to prefer to have

attendance spread more evenly across games or whether owners should only care about total season attendance. Staffing issues may be a concern, but as long as game-to-game attendance is predictable, variable staffing costs probably are not important enough to warrant such a desire.

Rather than stemming from the cost side, however, owners may prefer to have more evenly distributed crowds to maintain a home field advantage or due to demand factors via fan enjoyment. All else equal, fans probably prefer smaller crowds as it makes transportation easier, concession lines shorter, and to some extent seating potentially more spacious. Moreover, Table 3 suggests that lower quality visiting teams attract smaller crowds. But these games are precisely the games that the home team is more likely to win. If owners could attract more fans to these games, the fan experience may be better not just for "small crowd" reasons but also because the fans are more likely to see the home team win the game. This positive experience, then, might lead to those fans choosing to attend more games in the future.

Regardless of what the motives may be, however, the means by which teams are selling tickets indicates that distributing ticket sales more equally across games is a likely profit-maximizing objective. In particular, three recent developments are prevalent in ticket sales: variable ticket pricing, mini-season ticket packages, and online ticket retailers (see Rascher & Schwarz, in press). Variable ticket pricing is a means of nonlinear pricing by which teams charge different prices for the same seat to different games depending on the opponent, month, or day of the week. Assuming there is greater demand for some games than others, it is well known that revenue and profit can be increased by implementing an appropriate nonlinear pricing scheme. Rascher, McEvoy, Nagel, and Brown (2007) estimate variable ticket pricing can increase revenue by up to 7% for some teams. In addition to variable ticket pricing, teams offer mini-season ticket packages in which one can purchase tickets to considerably fewer than all 81 home games for the season (e.g., to only 20 or 28 games). As advertised, mini-season ticket packages allow the customer to see many different visiting teams. From the teams' point of view, however, mini-season ticket packages are a means to require the customer to see many different visiting teams, including the ones for which demand is low. Finally, online ticket brokers are yet another way teams engage in nonlinear pricing. By reserving some tickets for online ticket brokers, teams set some prices on a game-to-game basis, allowing tickets for low-demand games to be sold relatively cheaply while reaping higher prices (sometimes substantially higher prices) for premium games.

The implication from MLB teams engaging in more forms of nonlinear pricing should be that ticket sales are spread more evenly across games.[13] And the empirical implication from having ticket sales being spread more evenly across games is that factors that once affected ticket sales should now have less affect.[14] More precisely, the decrease in influence on ticket sales should be greater for unforeseeable (at the start of the season) factors such as the home team's current winning percentage than for foreseeable factors (such as whether the visiting team qualified for the playoffs the previous year) as nonlinear pricing schemes reduce the quantity of tickets sold on

game day. The results from Table 3 largely support this conclusion on several fronts. There is less variation in ticket sales by month in the late 1990s and 2000s than there was in the late 1980s and early 1990s. The negative effect associated with trailing in the playoff race and the positive effects of the home team's and the visiting team's current winning percentages are likewise less in the late 1990s and 2000s than was the case in the late 1980s and early 1990s. On the other hand, the effect on ticket sales from foreseeable factors such as either team qualifying for the playoffs the previous year has not fallen with time, and ticket sales for games when the Cubs, Red Sox, and Yankees are visiting are even greater in the 2000s than at earlier times.

Conclusion

Using a sample of over 10,000 MLB games from 1985 through 2009, an empirical model of ticket sales with home-team fixed effects is estimated using censored (at stadium capacity) normal regression. This is the first study that allows coefficient estimates to vary over time. Although the signs of the estimates are relatively stable across MLB seasons, their magnitude and statistical significance frequently depend on the seasons under consideration.

By allowing estimates to change over time, five important empirical results emerge. (a) There is no evidence in support of the UOH. (b) Attendance at interleague games is greater than at league games, but the original large effect on attendance at interleague rivalry games has greatly dissipated since the inception of interleague play. (c) The relationship between playing in a newly constructed stadium and ticket sales is one of the most erratic relationships over time. In most years, playing in a newly constructed stadium is associated with a significant increase in attendance, and this effect lessens as the stadium ages. (d) Ticket sales are greater the closer the home team is to making the playoffs but are unrelated to the playoff position of the visiting team. Ticket sales are also higher when either team qualified for the playoffs during the previous season. (e) The New York Yankees perennially attracted large road crowds for all 25 seasons in the sample, while the Boston Red Sox attracted large road crowds during their World Series years, and the Cubs attracted significantly higher road crowds since 1996.

As a likely consequence of MLB teams being more actively engaged in nonlinear pricing, the coefficient of variation of game attendance has steadily fallen over the last quarter century while average attendance has steadily risen. Empirically this is evidenced through many factors that contributed to predicting attendance in the late 1980s and early 1990s having less of an influence on ticket sales starting in the second half of the 1990s and continuing to the present.

Authors' Note

Kelly DeRango, Dan Rascher, Isaac Rischall, and Holly Swyers provided comments on a previous draft. All errors are solely the authors' responsibility.

Acknowledgment

The authors thank Leo Kahane and two referees for comments that substantially improved the article. The authors also thank Keith Glantz for explaining how the format of betting lines for baseball games has changed over the years and Thomas Gordan for sharing his knowledge of MLB stadiums.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: Financial assistance has been provided by the Richter Apprentice Scholarship Program at Lake Forest College.

Notes

1. About 1.3% of games sold out from 1985 to 1999, and about 8% of games sold out since then.
2. MLB experienced a players' strike from August 12, 1994 through April 2, 1995, though games did not resume until April 25. Schmidt and Berri (2002) show that yearly attendance is affected by strikes. Though we do not control for the effect of a strike in the same way, we do control for the potential effects of a strike by including year dummy variables in each regression and by estimating the parameters separately for different time periods.
3. Although the omission of these variables could lead to omitted variable bias, the omitted variables should not be correlated with the scheduling variables in the model (e.g., day, month, visiting team, interleague play, etc.). Intuitively, the omitted variables (such as the pitchers' race) should not be correlated with home team characteristics like winning percentage or games back in the playoff race. The bias stemming from these omitted variables, therefore, should be minimal.
4. When average ticket prices or the total fan cost index is included in the specification, the coefficient estimates are essentially unchanged, and the estimated coefficients on prices are always positive.
5. April 29 is the first day in the sample for 1995 when the strike delayed the start of the season.
6. For a small number of games, we used the *Chicago Sun Times*, the *Boston Globe*, or www.baseball-reference.com to replace missing values. We used the Benjamin Eckstein line in place of the Glantz-Culver line whenever the latter was unavailable.
7. The strike-shortened 1994 season failed to have postseason play. Therefore, no team in 1995 was associated with having qualified for the playoffs during the 1994 season.
8. Games played on Memorial Day, Labor Day, and July 4th, when it falls on a Friday or Monday, are also considered weekend games.
9. If the squared term is omitted from the model as the variance inflation factor test suggests may be warranted, the estimated coefficient on the probability of the home team winning the contest is positive, though it is only statistically significant in the 2005-2009 regression.

10. An F test that all three interleague coefficients equal 0 is rejected with a p-value of less than .0001, while an F test that all three interleague coefficients are equal is associated with a p-value of .3830.
11. Testing that the coefficient estimates of 8,388 and 4,505 are equal is associated with a p-value of .0475. Testing that 4,505 equals 1,957 is associated with a p-value of .0905.
12. Coates and Humphreys (1999), Coffin (1996), McEvoy, Nagel, DeSchriver, and Brown, (2005) and others have also estimated a positive honeymoon effect associated with the construction of a new stadium.
13. A good example of this is shown in the point estimates on the monthly dummy variables in Table 3. From 1985 to 1993, July and August games attracted over 10,000 more fans than April games, but from 1994 to 2009, these summer months attracted only about 5,000–6,000 more fans. This change, however, is not due to lower summer attendance figures. Rather, before 1994, attendance at all April games barely averaged 20,000. From 1994 onward, however, April games averaged 26,000. This kind of redistribution of fans to previously less desirable games is exactly what should be the result of teams actively engaging in nonlinear pricing, such as offering 20-game mini-season ticket packages that likely include two or three April games.
14. One likely result from certain factors becoming less important determinants of ticket sales over time would be for the fixed-effect terms to increase in magnitude (and statistical significance). For the teams in existence for the entire sample period, the average fixed effect increased from 7,240 before the strike to 8,502 after the strike (and the average t statistics for the fixed effect increased from 5.75 before the strike to 7.00 after the strike).

References

Chicago Tribune. (1985–2009). Daily issues April 14 through September 30.

Clapp, C. M., & Hakes, J. K. (2005). How long a honeymoon? The effect of new stadiums on attendance in Major League Baseball. *Journal of Sports Economics*, *6*, 237-263.

Coates, D., & Humphreys, B. R. (1999). The growth effects of sport franchises, stadia, and arenas. *Journal of Policy Analysis and Management*, *18*, 601-624.

Coffin, D. A. (1996). If you build it, will they come? Attendance and new stadium construction. In J. Fizel, E. Gustafson, & L. Hadley (Eds.), *Baseball economics: Current research* (pp. 33-46). Westport, CT: Praeger.

Fort, R., & Rosenman, R. (1998). Winning and managing streaks for baseball. *American Statistical Association, 1998 Proceedings of the Section on Statistics in Sports*, 11-15.

Fort, R., & Rosenman, R. (1999). Streak management. In J. Fizel, E. Gustafson, & L. Hadley (Eds.), *Sports economics: Current research* (pp. 119-134). Westport, CT: Praeger.

Fortunato, J. A. (2006). Scheduling promotional events in Major League Baseball: Examining team and sponsor desires. *International Journal of Sports Marketing and Sponsorship*, *8*, 104-114.

Hill, J. R., Madura, J., & Zuber, R. A. (1982). The short run demand for major league baseball. *Atlantic Economic Journal*, *10*, 31-35.

Kahane, L., & Shmanske, S. (1997). Team roster turnover and attendance in major league baseball. *Applied Economics*, *29*, 425-431.

Knowles, G., Sherony, K., & Haupert, M. (1992). The demand for major league baseball: A test of the uncertainty of outcome hypothesis. *American Economist*, *36*, 72-80.

Lemke, R. J., Leonard, M., & Tlhokwane, K. (2010). Estimating attendance at Major League Baseball games for the 2007 season. *Journal of Sports Economics*, *11*, 316-348.

Lowry, P. J. (2006). *Green cathedrals: The ultimate celebration of major league and negro league ballparks.* New York, NY: Walker.

McDonald, M., & Rascher, D. (2000). Does bat day make cents? The effect of promotions on the demand for Major League Baseball. *Journal of Sport Management*, *14*, 8-27.

McEvoy, C. D., Nagel, M. S., DeSchriver, T. D., & Brown, M. T. (2005). Facility age and attendance in Major League Baseball. *Sport Management Review*, *8*, 19-41.

Meehan, J. W., Jr., Nelson, R. A., & Richardson, T. V. (2007). Competitive balance and game attendance in Major League Baseball. *Journal of Sports Economics*, *8*, 563-580.

Neale, W. (1964). The peculiar economics of professional sports. *Quarterly Journal of Economics*, *78*, 1-14.

Rascher, D. A. (1999). A test of the optimal positive production network externality in Major League Baseball. In J. Fizel, E. Gustafson, & L. Hadley (Eds.), *Sports economics: Current research* (pp. 27-45). Westport, CT: Praeger.

Rascher, D. A., McEvoy, C. D., Nagel, M. S., & Brown, M. T. (2007). Variable ticket pricing in Major League Baseball. *Journal of Sport Management*, *21*, 407-437.

Rascher, D. A., & Schwarz, A. D. (in press). Illustrations of price discrimination in baseball. In L. Kahane & S. Shmanske (Eds.), *The oxford handbook of sports economics volume 2: Economics through sports.* Oxford, UK: Oxford University Press.

Rottenberg, S. (1956). The baseball players' labor market. *Journal of Political Economy*, *64*, 242-258.

Schmidt, M. B., & Berri, D. J. (2002). The impact of the 1981 and 1994–1995 strikes on Major League Baseball attendance: A time series analysis. *Applied Economics*, *34*, 471-478.

Selig, A. H. (2001). Testimony before the Committee on the Judiciary, United States House of Representative, December 6.

Team Marketing Report. (2007). *Fan cost index* (MLB). Retrieved from www.teammarketing.com

Zimbalist, A. S. (2003). *May the best team win: Baseball economics and public policy.* Washington, DC: The Brookings Institution Press.

Bios

Elise M. Beckman is an undergraduate student at Lake Forest College.

Wenqiang Cai is an undergraduate student at Lake Forest College.

Rebecca M. Esrock is an undergraduate student at Lake Forest College.

Robert J. Lemke is associate professor of economics at Lake Forest College.