Journal of Sports Economics

http://jse.sagepub.com/

Estimating Attendance at Major League Baseball Games for the 2007 Season

Robert J. Lemke, Matthew Leonard and Kelebogile Tlhokwane

Journal of Sports Economics 2010 11: 316 originally published online 6 August 2009
DOI: 10.1177/1527002509337212

The online version of this article can be found at:
http://jse.sagepub.com/content/11/3/316

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

On behalf of:
The North American Association of Sports Economists

Journal of Sports Economics
11(3) 316-348

DOI: 10.1177/1527002509337212
http://jse.sagepub.com

Estimating Attendance at Major League Baseball Games for the 2007 Season

Robert J. Lemke[1], Matthew Leonard[1], and Kelebogile Tlhokwane[1]

Abstract

Using games from Major League Baseball's 2007 season, individual game attendance is estimated using censored normal regression with home-team fixed-effects. Included in the model are several factors affecting attendance, such as divisional and interleague rivalries, that to date have been omitted from the literature. The relationship between attendance and game characteristics is shown to be fundamentally different between small market and large market teams. Attendance is also shown to steadily increase as the probability that the home team will win the game increases, which stands in contrast to the uncertainty of outcome hypothesis that predicts that attendance will eventually decrease if the home team's chance of winning the game gets too large.

Keywords

attendance, major league baseball, competitive balance, uncertainty of outcome hypothesis

1. Introduction

Recently, Major League Baseball (MLB) has been concerned that the league has become less competitive on the field, and thus less popular and less financially

[1] Lake Forest College

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

sound. Commissioner of Baseball Allan H. (Bud) Selig testified to Congress, "When I was elected Commissioner on July 9, 1998, I pledged to concentrate on two areas that were especially troubling to our clubs: competitive balance on the field and the economic stability of the clubs and Baseball as a whole." (Selig, 2001). Selig commissioned the *Blue Ribbon Report Panel on Baseball Economics* (Levin, Mitchell, Volcker, & Will, 2000) to investigate these issues. The panel's final report highlighted concerns that a noncompetitive league would lose fans and revenue. In a review of the *Blue Ribbon Report*, however, Eckard (2001a) called for more investigation into the degree of competition in baseball before the league took action to address the elusive condition of competitive imbalance. To this end, several researchers have attempted to quantify the degree to which competitive balance has improved (Butler, 1995; Eckard, 2001b; Fort & Quirk, 1995; Horowitz, 1997; Schmidt, 2001; Schmidt and Berri, 2001; Utt & Fort, 2002) or worsened (Depken, 1999; Hadley, Ciecka, & Krautmann, 2005; Humphreys, 2002; Levin et al., 2000) in MLB over the last several decades.

The literature focuses on three aspects of competitive balance: season-to-season, within season, and game-to-game.[1] Although season-to-season competition certainly matters to fans, game-to-game competition is particularly important in baseball as almost 80% of team revenues come from local revenues, mostly in the form of ticket sales and media contracts (Levin et al., 2000; Zimbalist, 2003b). Moreover, game-to-game attendance drives total season attendance figures which in turn facilitate the season-to-season comparisons. The literature, therefore, remains interested in better understanding the determinants of game-to-game attendance.

El-Hodiri and Quirk (1971), Neale (1964), and Rottenberg (1956) among others are credited with developing the first economic theories of sports leagues (see Fort and Quirk, 1995, for a review). These theories, which are grounded in profit-maximization by teams that join together to form a sports league but then compete on the field, stress the importance of generating fan interest in the league's games. In particular, a league's financial viability requires attracting fans to the sport and enticing them to purchase tickets to its games. Fan interest, in turn, depends in part on competition between teams and the overall competitive balance of the league. Although other features of attending games such as stadium amenities are now viewed as important determinants of demand as well, the early focus was on needing to provide a competitive game to maintain and increase fan interest. Recognizing the importance of on-field competition, Rottenberg (1956) is credited with the earliest discussion of the uncertainty of outcome hypothesis, which posits that attendance will decline if the home team's chance of winning is too great. Empirical research, therefore, became focused on estimating demand for games/ticket sales, with a particular emphasis on the effect on-field competition or a league's overall level of competitive balance has on fan preferences and their willingness to pay to watch games.

Demmert (1973) and Noll (1974) provided the early empirical studies of season-long MLB attendance, while Siegfried and Eisenberg (1980) estimated season-long

attendance for minor league baseball teams. Hill, Madura, and Zuber (1982) provided the first analysis using game-level observations from the 1977 MLB season, and Knowles, Sherony, and Haupert (1992) used data from the 1988 season to test the uncertainty of outcome hypothesis. Using games from the 1996 season, Rascher (1999) confirmed the result of Knowles, Sherony, and Haupert that the uncertainty of outcome hypothesis held (i.e., attendance eventually decreases if the home team's probability of winning gets too large).[2] Most recently, Meehan, Nelson, and Richardson (2007) estimated attendance at MLB games during the 2000–2002 seasons and found that attendance was positively related to on-field competition, as measured by differences in win percentage, but only when the home team had a better record than the visiting team. When the visiting team had a better record than the home team, they found no relationship between attendance and on-field competition.

This article makes several contributions to the literature. First, the article includes and highlights the importance of several new explanatory variables that have heretofore been omitted from the literature. Second, several empirical models are estimated using various dependent variables and estimation techniques to allow for comparisons along these dimensions. Third, the model is reestimated for subsamples of the data. In so doing, the relationship between attendance and game characteristics is found to be fundamentally different when the data are split by month, market size, or each team's recent propensity to qualify for the playoffs.

Three general results stand out. First, attendance in September depends greatly on playoff position. Second, attendance is much more responsive to game characteristics— e.g., when the game is played, which team is visiting, whether there is a fireworks display following the game—for small market teams than for large market teams. Third, in contrast to the earlier empirical results of Knowles, Sherony, and Haupert (1992) for the 1988 season and Rascher (1999) for the 1996 season, attendance is expected to be higher for games during the 2007 season the greater the home team's chance of winning the game.

2. The Empirical Specification

Four issues must be addressed when estimating attendance at sporting events—the dependent variable, the estimation technique, the sample criteria, and the empirical specification. Previous empirical studies have differed in all of these aspects. This section discusses each of these four issues in turn.

The main studies of attendance at MLB games, Hill, Madura, and Zuber (1982), Knowles, Sherony, and Haupert (1992), Meehan, Nelson, and Richardson (2007), and Rascher (1999)—henceforth HMZ, KSH, MNR, and Rascher respectively—, all used game attendance (or ticket sales) as the dependent variable, as did Jennet (1984) when looking at Scottish league football matches. In contrast, Siegfried and Eisenberg (1980) used the natural log of season attendance for minor league baseball clubs, while Borland and Lye (1992) and Peel and Thomas (1988) used the natural

log of attendance at individual matches of Australian Rules football and English football, respectively. In an attempt to control for varying stadium capacities, Welki and Zlatoper (1999) estimated attendance at U.S. football games by calculating tickets bought and used as a percentage of stadium capacity. Although most of the results in this article use attendance (i.e., ticket sales) as the dependent variable, some log attendance regressions are also included.

The choice of estimation procedure is equally diverse in the literature. HMZ and KSH used ordinary least squares (OLS). MNR used a censored-normal regression (CNR)[3] to take into account that some games sell out,[4] while Rascher used CNR with home-team fixed-effects. To facilitate a comparison across techniques, section 4 presents results under OLS and CNR, both with and without fixed-effects.

In addition to allowing for comparisons based on the dependent variable and estimation technique, we improve on the literature in three ways. First, we include the most complete set of explanatory variables to be found in the literature. Second, we include interaction terms to allow for a more flexible functional form for games played in September. Third, we estimate our preferred model under three criteria important to baseball—by month, by market size, and by playoff propensity.[5] As team revenue is largely driven by ticket sales in baseball (Levin et al., 2000), it is important for the league to have teams sell tickets not only in large cities or in September or to fans of playoff-bound teams but also to sell tickets in small markets, throughout the summer, to fans of teams that are not bound for the playoffs.

Explaining attendance at baseball games is difficult in part because most explanatory variables enter a fan's preferences in several ways such as wanting to enjoy the overall experience and wanting to see the home team win, which are not necessarily the same objective. Also, the literature mostly resigns itself to explaining patterns in attendance rather than formally estimating a demand equation. This is largely for institutional reasons—not only is the supply of tickets fixed from game to game within a season, but the price of tickets is also set months in advance of fans seeing the quality of play of their home team. To provide a general structure for the empirical model estimated below, the explanatory variables are classified in five groups: time factors (T), fan interest variables (F), city characteristics (CC), playoff chances (PC), and the probability of winning (PW). The general model takes the form:

$$y_{i,g} = \beta_0 + \beta_1 T_{i,g} + \beta_2 F_{i,g} + \beta_3 CC_{i,g} + \beta_4 PC_{i,g} + \beta_5 PW_{i,g} + \mu_i + \varepsilon_{i,g}$$

where y is either game attendance or its log, T, F, CC, PC, and PW are vectors of explanatory variables, β_1, β_2, β_3, β_4, β_5 are vectors of parameters, μ is a vector of home-team fixed-effects,[6] ε is the error term, and i indexes the home team while g indexes the game.

Lastly, baseball games are played in series. On a ten-game homestand, for example, the Yankees might host the Red Sox for three games, then the Royals for three, and then the Orioles for four. Not only is the opponent the same for three or four games in a row, but each series tends to occur over the weekend (possibly starting on

Thursday or going through Monday) or on weekdays. Thus, the error terms, $\varepsilon_{i,g}$, are almost certainly correlated across the games within each series. We adjust all of the estimation procedures to produce standard errors that are robust to this type of clustering.

3. The Data

The data were collected for all 2,430 games (81 home games played by each of 30 MLB teams) for the 2007 season. Four cuts on the data were then made. First, all Toronto home games were removed due to missing data. Second, the second game of each of the five traditional doubleheaders that were played during the season was omitted as only one attendance figure is available for the day.[7] Third, the Cleveland Indians played three "homes games" against the Los Angeles Angels at Miller Park in Milwaukee as snow fell in Cleveland. These three games plus a make-up game played in Seattle but designated a home game for Cleveland were removed from the data set because of the change in venue. Fourth, as some variables rely on data from recent games, 144 games were also dropped from the analysis as the home team had not yet played at least ten games.[8] The result is a data set of 2,196 games played during the 2007 MLB season, with each of 29 teams contributing at least 73 and at most 78 home games.

Most of the data come from box score information reported by MLB. The box score for each game for the 2007 season is available online at www.mlb.com. As MLB records the number of purchased tickets rather than the number of tickets used (turnstile attendance), the dependent variable for the analysis is paid attendance (or its natural log), which we will refer to henceforth simply as attendance. Average attendance at MLB games for the 2007 season was 33,004, with the smallest crowd being 8,201 (the Pirates hosting the Astros on April 25) and the largest crowd being 56,438 (the Mets hosting the Yankees on May 20). We also used the "pregame" description of each game as reported daily in the *Chicago Tribune* to gather information on scheduled starting pitchers and the betting line. Descriptive statistics and the data source for all variables used in the analysis are reported in Table 1.

Time Factors

Ten dummy variables control for the day of the week and the time of the game. There is a separate dummy variable for each day, Monday through Friday, plus a variable for playing a day game during the week. Saturday and Sunday games, however, are each separated by time of day: 3.6% of all games were played on Saturday afternoon, 12.8% on Saturday night, 14.7% on Sunday afternoon, and 1.4% on Sunday night. Six more variables control for the month. Whereas all studies of baseball attendance recognize the importance of the schedule (time of day, day of the week,

Table 1. Descriptive Statistics for MLB Games during the 2007 Season

Dependent Variable	Mean	SD	Min	Max
Attendance[a]	33,004	11,131	8,201	56,438
Time Factors				
Played on Monday.	0.102	0.303	0	1
Played on Tuesday.	0.152	0.359	0	1
Played on Wednesday.	0.150	0.357	0	1
Played on Thursday.	0.111	0.314	0	1
Played on Friday.	0.161	0.368	0	1
Played during the day, Monday–Friday.	0.100	0.300	0	1
Played on Saturday afternoon.	0.036	0.185	0	1
Played on Saturday night.	0.128	0.334	0	1
Played on Sunday afternoon.	0.147	0.354	0	1
Played on Sunday night.	0.014	0.116	0	1
Played in April.	0.095	0.294	0	1
Played in May.	0.185	0.389	0	1
Played in June.	0.178	0.382	0	1
Played in July.	0.173	0.379	0	1
Played in August.	0.186	0.389	0	1
Played in September.	0.182	0.386	0	1
Public schools are on summer vacation.[b]	0.449	0.498	0	1
Played on Memorial Day, July 4th, or Labor Day.	0.018	0.134	0	1
Fan interest				
Game-time temperature (°F) for outdoor parks ($N = 1{,}744$).[a]	75.3	10.4	40	106
Rain fell during the game.[a]	0.021	0.144	0	1
The home team plays in a dome or retractable roof stadium.	0.206	0.404	0	1
The home team is in the American League.	0.447	0.497	0	1
A divisional game.	0.428	0.495	0	1
A divisional rivalry game.[c]	0.094	0.292	0	1
An interleague game.	0.111	0.314	0	1
An interleague rivalry game.	0.027	0.163	0	1
Barry Bonds has 753–755 career home runs.	0.009	0.093	0	1
The home team's pitcher is a fan favorite.[c]	0.139	0.346	0	1
Home runs per game hit by the home team.	0.961	0.219	0.200	1.583
Wins–losses of home pitcher less wins–losses of visiting pitcher.[d]	–0.078	4.822	–20	20
Number of games won by the home team in its last 10.	4.94	1.65	0	10
Age of stadium (in years).[e]	26	26	2	96
Stadium opened in 2000 or more recently.[e]	0.311	0.463	0	1
There is a giveaway/promotion at the stadium.[a]	0.309	0.462	0	1
A fireworks display followed the game.[a]	0.060	0.237	0	1
Average 2007 ticket price.[f]	22.81	6.979	13.79	47.74
Home team qualified for the 2006 playoffs.	0.279	0.448	0	1

(continued)

Table 1 (continued)

Dependent Variable	Mean	SD	Min	Max
Visiting team qualified for the 2006 playoffs.	0.265	0.442	0	1
Visiting team was the Boston Red Sox.	0.030	0.171	0	1
Visiting team was the New York Yankees.	0.031	0.173	0	1
Visiting team was the Chicago Cubs.	0.034	0.182	0	1
	Mean	SD	Min	Max
City Characteristics				
MSA 2006 population (in 1,000s).[g]	5,775	4,618	1,510	18,819
MSA 2005 per capita income (in US$1,000s).[g]	39.2	5.2	32.4	52.5
MSA 2005 unemployment rate.	6.9	2.4	3.8	15.3
MSA 2005 poverty rate.[g]	21.4	4.9	12.2	32.4
MSA 2006 percentage black population.[g]	14.7	7.4	4.5	31.3
MSA 2006 percentage Hispanic population.[g]	16.5	12.7	1	44
Number corporations hdqrtrd in MSA with 500+ emplys in 2008.[h]	531.3	354.8	147	1357
MSA is home to two MLB teams.	0.208	0.406	0	1
Game was not available on TV locally.[a]	0.022	0.148	0	1
Game was on national TV.[a]	0.065	0.246	0	1
Miles between stadiums (in 1,000s).[i]	1.233	0.851	0.010	4.884
Playoff Chances				
Home team's games back in the playoff race at the start of play.	5.51	5.76	0	26.5
Visiting team's games back in the playoff race at the start of play.	5.50	5.55	0	24.5
Home team was leading its division at the start of play.	0.216	0.412	0	1
Away team was leading its division at the start of play.	0.220	0.415	0	1
Home team was leading the wild card race at the start of play.	0.072	0.258	0	1
Away team was leading the wild card race at the start of play.	0.055	0.227	0	1
Home team was in playoff contention at the start of play.	0.327	0.470	0	1
Probability of Winning				
Implied probability the home team will win (from betting lines).[d]	0.543	0.085	0.290	0.750

Notes: The data set includes information on 2,196 home games played during the 2007 MLB season. See the text for more complete variable descriptions.

[a] Major League Baseball (www.mlb.com).

[b] Online school district calendars for each host city of a MLB team.

[c] Authors' assessment.

[d] Chigago Tribune.

[e] Ballparks of Baseball (www.ballparksofbaseball.com).

[f] Team Marketing Report (www.teammarketingreport.com).

[g] 2008 *U.S. Statistical Abstract* (www.census.gov/compendia/stataab/).

[h] infoUSA.com.

[i] mapquest.com.

and month), no previous study has allowed this amount of flexibility in the specification.

We also control for whether the largest public school district in the home team's metropolitan statistical area (MSA) was on summer vacation. Although there is little variation in when students start school in the fall (usually late August or early September), there is considerable variation in when the school year ends (from before Memorial Day to the last week of June). We also control for whether the game was played on Memorial Day, the Fourth of July, or Labor Day. These factors are important to take into account as families are more likely to attend games when children are not in school or when people are off work for a holiday.

Fan Interest Variables

As weather is likely to affect one's experience, the temperature (at outdoor stadiums), whether it rained during the game (zero for domed stadiums), and whether the stadium has a roof are all included in the specification. The average game-time temperature was 75.3°F, and measurable rain fell during 2.1% of games. More than one in every five games was played in domed stadiums (Minnesota, Tampa Bay) or in stadiums with a retractable roof (Arizona, Houston, Milwaukee, Seattle).

Our sample includes 16 National League home teams and 13 American League home teams. Each of the two leagues is further divided into three divisions—East, Central, and West. Almost 43% of games were played between two teams in the same division in the same league. We also subjectively assigned up to two division rivals for each team based on conversations with local sports reporters and sports economists.[9] Moreover, the assignment of division rivals was not constrained to be symmetric. For example, while the Orioles consider the Red Sox and Yankees to both be division rivals in part because of their proximity and in part because of their recent success, the Orioles are not considered to be a division rival of either the Red Sox or the Yankees as both of those teams fixate on each other and serve as each other's sole division rival. The assignment of division rivals is given in Table 2. Division rivalry games account for 9.4% of all games.

Current studies of attendance also need to account for interleague games. Whereas MNR included an interleague dummy variable, we include this plus another variable indicating the home team played an interleague rival. In particular, MLB schedules ten interleague rivalry match-ups each year based on geography: Baltimore vs. Washington, the Chicago White Sox vs. Cubs, Cleveland vs. Cincinnati, Kansas City vs. St. Louis, the Los Angeles Angels vs. Dodgers, Minnesota vs. Milwaukee, the New York Yankees vs. Mets, Oakland vs. San Francisco, Tampa Bay vs. Florida, and Texas vs. Houston. Interleague games account for 11.1% of all games played in our sample, while interleague rivalry games account for 2.7%.

The 2007 MLB season witnessed Barry Bonds' historic chase of Hank Aaron's career home run record of 755. Although Bonds' pursuit of the record garnered publicity throughout the year, the potential for tying and passing Aaron as Bonds

Table 2. Division Rivalries

Home Team	Division Rivals
Baltimore Orioles	Boston, New York
Boston Red Sox	New York
New York Yankees	Boston
Tampa Bay Devil Rays	None
Toronto Blue Jays	New York
Cleveland Indians	Detroit
Chicago White Sox	Cleveland, Detroit
Detroit Tigers	Cleveland
Kansas City Royals	None
Minnesota Twins	Cleveland
Los Angeles Angels of Anaheim	Oakland
Oakland Athletics	Los Angeles
Seattle Mariners	None
Texas Rangers	None
Atlanta Braves	New York, Florida
Florida Marlins	Atlanta
New York Mets	Atlanta, Philadelphia
Philadelphia Phillies	New York
Washington Nationals	None
Chicago Cubs	Milwaukee, St. Louis
Cincinnati Reds	None
Houston Astros	None
Milwaukee Brewers	Chicago
Pittsburgh Pirates	None
St. Louis Cardinals	Chicago
Arizona Diamondbacks	Colorado
Colorado Rockies	Arizona
Los Angeles Dodgers	San Francisco
San Diego Padres	None
San Francisco Giants	Los Angeles

approached 755 generated extraordinary interest and certainly brought fans to Giants games. We include a dummy variable for the 19 Giants games in which Bonds was within three home runs (753 to 755) of breaking the career home run record.

Whereas HMZ included whether the starting pitchers were minorities and Rascher included several dummy variables capturing the race of the pitchers, we talked to local sports reporters and sports economists to subjectively define pitchers to be ''fan favorites'' if the pitcher was likely to attract additional fans to home games simply because he was pitching. As we have defined them, the fan-favorite pitchers are Johnson and Webb (Arizona), Smoltz (Atlanta), Beckett and Matsuzaka (Boston), Zambrano (Chicago Cubs), Harang (Cincinnati), Carmona and Sabathia (Cleveland), Francis (Colorado), Willis (Florida), Oswalt (Houston), Penny

(Dodgers), Sheets (Milwaukee), Santana (Minnesota), Clemens and Pettitte (Yankees), Hamels (Philadelphia), Peavy (San Diego), and Hernandez (Seattle). Of these twenty pitchers, sixteen were well-known, productive pitchers before 2007. Carmona, Francis, Hamels, and Hernandez, however, were all at best average pitchers with not much major league experience through 2006. For each of these pitchers, 2007 was a breakout year. Although none of them would have been considered a fan favorite in April, they were all in the midst of a dominating season by June. The results are not overly sensitive to whether these four pitchers are considered favorites or not, but the point estimate on having a fan-favorite pitcher start a game is larger and more significant when these four pitchers are included in the definition.

The next three variables listed in Table 1 relate to team performance and are intended to measure the potential excitement of the game. At the start of play each day, the home team had previously hit an average of 0.96 home runs per game. To capture the quality of the pitching match up to date in 2007, the visiting pitcher's net wins was subtracted from the home pitcher's net wins. As expected, the average difference was close to zero. Lastly, the home team averaged 4.94 wins in its last 10 games.[10]

The next five fan interest variables capture the atmosphere of the stadium experience. The average stadium was 26 years old in 2007, but 31% of games were played in a stadium that opened as recently as 2000. Almost 31% of games were associated with a giveaway/promotion such as cap day or bobblehead day,[11] and 6% of games were followed by a fireworks display. The 2007 average ticket price as reported by *Team Marketing Report* was US$22.81.

Finally, attendance likely depends on which team is visiting (e.g., a playoff team from last season or a team with a superstar like Alex Rodriguez). Likewise, some teams are thought to have loyal fans that ''travel well'' while other teams might bring home-team fans to the stadium out of spite. The Cubs, Red Sox, and Yankees were the only road teams to average an attendance of at least 36,000 per game.[12] The fan experience variables, therefore, include whether the home or visiting team qualified for the 2006 playoffs and whether the Cubs, Red Sox, or Yankees were the visiting team.

City Characteristics

Unlike several of the fan interest variables, none of the city characteristic variables are unique to our analysis. The home team's city population, per capita income, unemployment rate, poverty rate, percentage population that is black, and percentage population that is Hispanic were taken at the MSA level from the *2008 Statistical Abstract of the United States*. Market size, as approximated by population, varies greatly across teams from 1.5 million (Milwaukee) to 19 million (New York). Per capita income also varies considerably, from US$32,400 in Phoenix to US$52,500 in San Francisco. There is similar variation in the other city variables as well.

Because corporations support local teams by purchasing advertising, club seats, and luxury boxes, we include the number of corporations with at least 500 employees headquartered in each home team's MSA in 2008 as reported by infoUSA, Inc. We also control for two MLB teams existing in the same MSA, which occurs in Chicago, Los Angeles, and New York.

Whereas HMZ included a variable for when the game was broadcast on local TV (a rarity in 1977), we include separate variables for when the game is not available on local TV (a rarity in 2007) and when the game is aired nationally on FOX, ESPN, or ESPN2. Finally, following KSH, we include the distance between the home team's and visiting team's stadium (measured in 1,000s as reported by mapquest.com) to control for the ease with which fans of the visiting team can attend the game. The average distance between stadiums is 1,233 miles.

Playoff Chances

Several measures of playoff position are included for both the home and visiting team. Although HMZ and MNR included the number of games back each team was in its respective division, we include each team's games back in the playoff race (i.e., the lesser of games back in the divisional or wildcard races). We also include dummy variables for whether the home or visiting team is leading its division or is not leading its division but is leading its league's wild card race. Lastly, we include a dummy variable indicating whether the home team does not hold a playoff position but is "in playoff contention" at the start of play. To be in playoff contention April through August, the home team cannot be leading its division or wild card race and must be within 6 games of a playoff spot. In September, the home team must be within 6 games with 8 or more to play, within 4 games with 5 to 7 games to play, or with any mathematical possibility of making the playoffs with 4 or fewer games to play.[13]

Probability of Winning

The last variable reported in Table 1 is the probability that the home team will win the game. It has long been recognized that competition on the field is important for attracting fans to games. It is generally thought that fans prefer to see the home team win, but the question arises whether fans have a greater and greater taste for victory (and therefore a taste for competitive imbalance on the field) or whether fans prefer to see a competitive game (the uncertainty of outcome hypothesis). Following HMZ and Rascher, our empirical model includes the probability of the home team winning as a quadratic.

We computed the probability of the home team winning each game from the Glantz-Culver betting lines reported in the *Chicago Tribune*. Betting lines are listed as $-F$ and U, where the negative sign on F indicates the favored team. A typical line, for example, is –150 Chicago Cubs and 140 Milwaukee Brewers. In this case, one could wager US$1.50 on the Cubs and win US$1.00 plus the original US$1.50 if the Cubs beat the Brewers, or one could wager US$1.00 on the Brewers and win

US$1.40 plus the original US$1.00 if the Brewers beat the Cubs. A fair bet in this context arises only if $F = U$, which the bookmaker never allows.[14] As no commission is charged when betting on baseball and as the bookmaker is indifferent as to which side he makes his money on, it is natural to assume that the bookmaker believes the fair bet is $R = (F + U) / 2$. Using R as the fair line, the implied probability that the favored team wins is $R / (R + 100)$ while the implied probability the underdog wins is $100 / (R + 100)$. The probability of the home team winning averages 0.543 but ranges from 0.29 to 0.75 in our sample.

4. Regression Results

Table 3 reports the estimated coefficients from eight regressions—four each when using attendance and log attendance as the dependent variable. Each dependent variable is estimated using ordinary least squares (OLS), a Constrained Normal Regression (CNR) which censors attendance at the stadium's capacity, a fixed-effects model (FE) in which each home team receives its own fixed-effect, and a CNR-FE model. The absolute value of the t statistic for each estimated coefficient is reported in parentheses, where the standard errors have been corrected for the cluster problem due to games being played in series. As the stadium, league, MSA, ticket prices, and whether the home team made the 2006 playoffs do not vary for a home team, coefficients for these variables are not estimated for any of the fixed-effects models.

When using attendance or log attendance as the dependent variable, estimated coefficients are interpreted, respectively, as changes in attendance or percentage changes in attendance. Under OLS, for example, a Monday night game averages 5,184 fewer people in attendance (or 19.1% fewer when using the log specification) compared to an otherwise identical Sunday afternoon game.

In terms of the model specification, several features stand out. First, the signs and statistical significance of coefficient estimates are largely the same when using attendance or log attendance. Second, although the OLS estimates closely match the CNR estimates and the FE estimates closely match the CNR-FE estimates, there are frequently substantial differences between the OLS/CNR and FE/CNR-FE estimates. Which estimation procedure should be preferred is debatable. Intuitively, fixed-effects should matter. Drawing fans to Cardinals, Cubs, Red Sox, or Yankees games is simply different from drawing fans to Pirates, Rays, Royals, or Twins games. Similarly, stadium capacity limits attendance at some games. Thus, from the standpoint of how attendance should be modeled econometrically, CNR-FE is the best candidate. The remainder of the article, therefore, focuses exclusively on CNR-FE results when using attendance as the dependent variable to facilitate comparisons with HMZ, KSH, MNR, and Rascher, even though the fixed-effects model fails to identify effects from any unchanging home-team variable such as average ticket prices, stadium features, or city characteristics.

Table 3. Regression Results—Full Sample

Explanatory Variable / Regression Technique	Dep. Variable = Attendance				Dep. Variable = Ln (Attendance)			
	OLS	CNR	FE	CNR-FE	OLS	CNR	FE	CNR-FE
Time Factors:								
Played on Monday.	–5184	–5646	–5295	–5780	–0.191	–0.210	–0.194	–0.214
	(9.879)	(10.459)	(12.083)	(12.824)	(9.267)	(9.973)	(11.589)	(12.534)
Played on Tuesday.	–4595	–5040	–4538	–4997	–0.171	–0.190	–0.167	–0.186
	(9.940)	(10.604)	(12.039)	(12.983)	(9.718)	(10.465)	(11.945)	(13.030)
Played on Wednesday.	–4580	–5013	–4387	–4816	–0.170	–0.188	–0.161	–0.179
	(9.839)	(10.395)	(11.964)	(12.728)	(9.412)	(10.05)	(11.57)	(12.467)
Played on Thursday.	–4347	–4794	–3908	–4313	–0.163	–0.181	–0.146	–0.162
	(8.004)	(8.560)	(8.424)	(9.137)	(7.552)	(8.173)	(8.059)	(8.853)
Played on Friday.	–1419	–1675	–1159	–1434	–0.048	–0.058	–0.036	–0.048
	(4.373)	(4.735)	(3.787)	(4.423)	(4.003)	(4.452)	(3.292)	(4.051)
Day game.	1208	1400	668	727	0.043	0.050	0.022	0.024
	(2.380)	(2.711)	(1.570)	(1.685)	(2.048)	(2.390)	(1.274)	(1.389)
Saturday afternoon game.	871	1358	281	757	0.010	0.027	–0.009	0.008
	(1.498)	(2.092)	(0.572)	(1.390)	(0.497)	(1.154)	(0.511)	(0.400)
Saturday night game.	2523	2806	2644	2869	0.091	0.103	0.094	0.104
	(7.587)	(7.595)	(8.315)	(8.307)	(7.736)	(7.797)	(8.261)	(8.350)
Sunday night game.	–1062	–1319	–2096	–2503	–0.040	–0.052	–0.081	–0.099
	(1.141)	(1.309)	(2.450)	(2.751)	(1.338)	(1.543)	(2.888)	(3.250)
Played in April.	–5186	–5377	–2438	–2664	–0.186	–0.194	–0.075	–0.083
	(5.075)	(5.129)	(2.852)	(3.073)	(4.766)	(4.870)	(2.363)	(2.616)
Played in May.	–3226	–3376	–2005	–2238	–0.115	–0.122	–0.065	–0.074
	(3.960)	(4.025)	(3.103)	(3.414)	(3.739)	(3.872)	(2.705)	(3.083)
Played in June.	–1321	–1363	–1068	–1205	–0.041	–0.042	–0.027	–0.031
	(1.868)	(1.869)	(1.962)	(2.172)	(1.513)	(1.509)	(1.315)	(1.507)

(continued)

Table 3 (continued)

Explanatory Variable / Regression Technique	Dep. Variable = Attendance				Dep. Variable = Ln (Attendance)			
	OLS	CNR	FE	CNR-FE	OLS	CNR	FE	CNR-FE
Played in August.	387	477	365	337	0.008	0.011	0.009	0.008
	(0.652)	(0.752)	(0.788)	(0.705)	(0.370)	(0.462)	(0.540)	(0.457)
Played in September.	1667	1383	−1706	−2286	0.093	0.081	−0.051	−0.074
	(0.676)	(0.554)	(0.942)	(1.250)	(0.961)	(0.836)	(0.696)	(1.018)
Public schools are on summer vacation.	1467	1710	1460	1648	0.050	0.059	0.052	0.060
	(2.571)	(2.914)	(3.423)	(3.758)	(2.284)	(2.673)	(3.196)	(3.624)
Game was played on Memorial Day, July 4, or Labor Day.	3299	3239	3650	3725	0.115	0.114	0.130	0.135
	(3.095)	(2.973)	(3.976)	(3.994)	(2.966)	(2.879)	(3.926)	(3.983)
Fan Interest Variables:								
Game-time temperature (outdoor stadium).	366	436	513	579	0.012	0.015	0.018	0.021
	(2.087)	(2.450)	(3.549)	(3.939)	(1.767)	(2.139)	(3.093)	(3.495)
Game-time temperature squared (outdoor stadium).	−2	−3	−3	−4	−6.5E-05	−8.5E-05	−1.1E-04	−1.3E-04
	(1.741)	(2.129)	(3.261)	(3.702)	(1.379)	(1.781)	(2.844)	(3.293)
Rain during game.	97	−111	−607	−826	0.015	0.006	−0.021	−0.030
	(0.103)	(0.117)	(0.827)	(1.097)	(0.370)	(0.157)	(0.622)	(0.917)
Game was played in a dome or retractable roof stadium.	17417	19992			0.628	0.734		
	(2.727)	(3.105)			(2.406)	(2.787)		
Home team is in the American League.	160	138			0.041	0.040		
	(0.254)	(0.215)			(1.718)	(1.639)		
A divisional game.	−38	−48	66	36	−0.009	−0.010	0.002	0.001
	(0.079)	(0.097)	(0.173)	(0.094)	(0.443)	(0.491)	(0.149)	(0.054)
A divisional rivalry game.	1140	981	1112	977	0.026	0.021	0.034	0.029
	(1.452)	(1.217)	(1.662)	(1.446)	(0.938)	(0.722)	(1.421)	(1.201)
An interleague game.	1349	1514	1828	2006	0.042	0.048	0.057	0.064
	(1.628)	(1.795)	(2.876)	(3.124)	(1.438)	(1.642)	(2.599)	(2.945)
An interleague rivalry game.	2228	2339	1557	1607	0.064	0.070	0.057	0.059

(continued)

Table 3 (continued)

Explanatory Variable / Regression Technique	Dep. Variable = Attendance				Dep. Variable = Ln (Attendance)			
	OLS	CNR	FE	CNR-FE	OLS	CNR	FE	CNR-FE
	(1.289)	(1.237)	(1.407)	(1.310)	(0.964)	(0.953)	(1.278)	(1.210)
Barry Bonds had between 753 and 755 home runs.	6355	12921	4777	10102	0.175	0.419	0.120	0.314
	(3.368)	(3.782)	(4.362)	(4.245)	(2.820)	(3.494)	(3.360)	(3.882)
Home team's starting pitcher is a fan favorite.	519	662	324	427	0.020	0.027	0.007	0.011
	(1.375)	(1.674)	(0.989)	(1.249)	(1.483)	(1.852)	(0.550)	(0.883)
Home runs the home team hits per game on average.	–7339	–7328	1594	2090	–0.354	–0.353	0.024	0.045
	(6.747)	(6.640)	(1.036)	(1.307)	(7.602)	(7.590)	(0.400)	(0.739)
Wins–losses of home starting pitcher less wins–losses of visiting pitcher.	–51	–44	–13	–8	–0.002	–0.002	–0.001	–0.001
	(1.718)	(1.412)	(0.533)	(0.297)	(1.977)	(1.652)	(1.220)	(0.923)
Number of games the home team has won in its last 10 games.	41	41	114	99	0.002	0.002	0.005	0.004
	(0.332)	(0.321)	(1.147)	(0.984)	(0.490)	(0.468)	(1.317)	(1.132)
Age of stadium (in years).	–430	–458			–0.025	–0.026		
	(6.660)	(6.939)			(9.044)	(9.377)		
Age of stadium squared.	4	4			2.1E-04	2.3E-04		
	(5.596)	(6.112)			(8.138)	(8.660)		
Home team plays in a stadium built in 2000 or more recently.	2256	3231			0.032	0.073		
	(2.774)	(3.979)			(0.990)	(2.303)		
There is a giveaway/promotion at the stadium.	1721	1801	1796	1784	0.067	0.071	0.074	0.075
	(5.545)	(5.479)	(6.763)	(6.386)	(5.744)	(5.742)	(7.625)	(7.299)
There is a fireworks display at the stadium following the game.	5022	5780	4108	4777	0.184	0.214	0.143	0.170
	(7.245)	(7.890)	(6.740)	(7.680)	(7.529)	(8.245)	(6.558)	(7.659)
Average ticket price in 2007, inflation adjusted.	355	311			0.007	0.006		
	(6.032)	(5.156)			(3.257)	(2.471)		
Home team qualified for the 2006 season playoffs.	4593	4681			0.156	0.159		
	(7.354)	(7.108)			(6.822)	(6.647)		

(continued)

Table 3 (continued)

Explanatory Variable / Regression Technique	Dep. Variable = Attendance				Dep. Variable = Ln (Attendance)			
	OLS	CNR	FE	CNR-FE	OLS	CNR	FE	CNR-FE
Visiting team qualified for the 2006 season playoffs.	1123	1266	1276	1392	0.043	0.049	0.050	0.055
	(2.463)	(2.706)	(3.669)	(3.921)	(2.439)	(2.716)	(3.782)	(4.090)
Visiting team was the Red Sox.	7554	8549	7983	8748	0.250	0.292	0.278	0.312
	(5.414)	(5.908)	(6.752)	(7.212)	(4.869)	(5.546)	(5.994)	(6.568)
Visiting team was the Yankees.	7642	8471	8188	9355	0.263	0.297	0.291	0.339
	(4.825)	(4.851)	(6.020)	(6.848)	(4.593)	(4.675)	(5.447)	(6.347)
Visiting team was the Cubs.	4034	3974	3843	3833	0.150	0.148	0.145	0.145
	(3.135)	(3.033)	(4.502)	(4.540)	(3.292)	(3.166)	(5.083)	(5.141)
City Characteristics:								
MSA 2006 population in 1,000s.	1	1			3.1E-05	3.1E-05		
	(8.138)	(7.881)			(6.995)	(6.792)		
MSA 2005 per capita income in US$1,000s.	373	409			0.021	0.023		
	(5.456)	(5.812)			(7.228)	(7.729)		
MSA 2005 unemployment rate.	–665	–549			–0.033	–0.028		
	(4.043)	(3.317)			(5.131)	(4.425)		
MSA 2005 poverty rate.	190	116			0.014	0.011		
	(2.089)	(1.256)			(3.536)	(2.760)		
MSA 2006 percentage black population.	221	254			0.008	0.009		
	(4.095)	(4.658)			(3.964)	(4.603)		
MSA 2006 percentage Hispanic population.	29	37			0.002	0.002		
	(1.153)	(1.370)			(2.231)	(2.465)		
Corporations headquartered in MSA.	–6	–6			–1.0E–04	–8.2E-05		
	(6.830)	(6.052)			(3.033)	(2.388)		
MSA is home to two MLB teams.	8750	8826			0.344	0.346		
	(6.770)	(6.584)			(7.271)	(7.043)		
Game is not on local TV.	–1838	–2302	–400	–607	–0.067	–0.086	0.004	–0.004

(continued)

Table 3 (continued)

Explanatory Variable / Regression Technique	Dep. Variable = Attendance				Dep. Variable = Ln (Attendance)			
	OLS	CNR	FE	CNR-FE	OLS	CNR	FE	CNR-FE
	(2.009)	(2.506)	(0.533)	(0.803)	(1.774)	(2.294)	(0.130)	(0.118)
Game is on national TV.	–437	–140	–509	–282	–0.026	–0.014	–0.027	–0.018
	(0.764)	(0.220)	(1.004)	(0.501)	(1.351)	(0.649)	(1.526)	(0.914)
Miles between stadiums in 1,000s.	–128	–325	–709	–980	–0.008	–0.017	–0.010	–0.022
	(0.156)	(0.396)	(1.219)	(1.670)	(0.249)	(0.542)	(0.426)	(0.914)
Miles between stadiums squared.	–60	–28	135	197	–0.003	–0.002	0.001	0.003
	(0.258)	(0.121)	(0.932)	(1.375)	(0.335)	(0.181)	(0.113)	(0.551)
Playoff Chances:								
Games back the home team was in the division or wildcard.	–270	–279	26	8	–0.005	–0.006	0.006	0.006
	(2.843)	(2.904)	(0.335)	(0.100)	(1.244)	(1.319)	(1.741)	(1.613)
Home team's games back × Sept. game.	–118	–114	20	34	–0.008	–0.007	–0.001	–0.001
	(0.752)	(0.737)	(0.162)	(0.280)	(1.136)	(1.153)	(0.271)	(0.183)
Games back the visiting team was in the division or wildcard.	–71	–74	–21	–18	–0.003	–0.003	–0.001	–0.001
	(1.334)	(1.365)	(0.468)	(0.405)	(1.231)	(1.274)	(0.526)	(0.443)
Visiting team's games back × September game.	66	95	–18	6	0.002	0.003	–0.002	–0.001
	(0.783)	(1.096)	(0.256)	(0.081)	(0.529)	(0.897)	(0.706)	(0.342)
Home team was leading its division.	3665	3937	379	514	0.191	0.203	0.042	0.048
	(3.189)	(3.375)	(0.411)	(0.554)	(3.720)	(3.971)	(1.071)	(1.237)
Home team was leading its division × September game.	–645	–313	4413	4850	–0.061	–0.048	0.153	0.170
	(0.263)	(0.128)	(2.303)	(2.537)	(0.639)	(0.509)	(2.004)	(2.267)
Visiting team was leading its division.	661	629	937	975	0.027	0.025	0.032	0.033
	(1.117)	(1.040)	(2.042)	(2.107)	(1.173)	(1.098)	(1.759)	(1.847)
Visiting team was leading its division × September game.	–610	–350	–494	–248	–0.021	–0.011	–0.014	–0.004
	(0.355)	(0.204)	(0.364)	(0.185)	(0.327)	(0.174)	(0.261)	(0.080)

(continued)

Table 3 (continued)

Explanatory Variable / Regression Technique	Dep. Variable = Attendance				Dep. Variable = Ln (Attendance)			
	OLS	CNR	FE	CNR-FE	OLS	CNR	FE	CNR-FE
Home team was leading the wild card race.	2946	2956	692	455	0.173	0.176	0.063	0.055
	(2.251)	(2.219)	(0.660)	(0.432)	(3.135)	(3.195)	(1.479)	(1.319)
Home team was leading wild card × September game.	–5493	–5978	2511	2744	–0.231	–0.254	0.097	0.104
	(2.057)	(2.164)	(1.253)	(1.356)	(2.292)	(2.470)	(1.229)	(1.323)
Visiting team was leading the wild card race.	1563	1698	1516	1618	0.056	0.062	0.046	0.051
	(1.987)	(2.013)	(2.423)	(2.427)	(1.730)	(1.840)	(1.875)	(1.986)
Visiting team was leading wild card × September game.	389	100	–1719	–2218	0.041	0.029	–0.028	–0.049
	(0.150)	(0.036)	(0.772)	(0.945)	(0.441)	(0.289)	(0.336)	(0.559)
Home team started the day in playoff contention.	1767	1894	720	670	0.085	0.091	0.047	0.045
	(1.996)	(2.118)	(1.152)	(1.068)	(2.184)	(2.346)	(1.670)	(1.634)
Home team was in playoff contention × September game.	–325	–593	–287	–364	0.002	–0.010	0.017	0.013
	(0.181)	(0.334)	(0.213)	(0.274)	(0.023)	(0.143)	(0.314)	(0.252)
Probability of Winning:								
Probability home team will win.	–35792	–29854	–48770	–47523	–1.636	–1.409	–1.982	–1.933
	(1.686)	(1.369)	(2.810)	(2.705)	(1.993)	(1.679)	(2.940)	(2.850)
Probability home team will win squared.	35953	29882	46738	45006	1.600	1.368	1.926	1.857
	(1.810)	(1.461)	(2.854)	(2.702)	(2.103)	(1.755)	(3.061)	(2.919)
Constant	7408	3393	46304	46537	9.363	9.196	10.879	10.889
	(0.783)	(0.351)	(8.754)	(8.663)	(25.047)	(24.154)	(52.853)	(52.424)
Number of observations.	2,196	2,196	2,196	2,196	2,196	2,196	2,196	2,196
Adjusted R-squared / log likelihood.	0.7577	–20567	0.8317	–20172	0.7149	–111.6	0.8055	–529.9

Note: See text for variable definitions and sources. The absolute value of the *t* statistic is reported in parentheses beneath the estimated coefficient. Standard *t* statistic thresholds of 2.565, 1.960, and 1.645 indicating statistical significance at the 1%, 5%, and 10% levels, respectively, apply. Standard errors have been corrected for the cluster correlation problem of having MLB games scheduled in series.

Sensitivity Analysis

In addition to focusing on censored normal regression with fixed-effects, several econometric issues, most notably multicollinearity, must be investigated. The first improvement to be made to the reduced form empirical model is to eliminate insignificant variables from the analysis. Using the CNR-FE results listed in column 4 of Table 3 as a baseline, variables with a *t* statistic less than 1.0 in absolute value were eliminated from the model. This process was then iterated one more time. As a result, we omit the August dummy variable (so month variables are compared to July and August), whether it rained during the game, whether it is a division game, home runs hit per game, the difference in net wins between starting pitchers, the number of games the home team has won in its last ten, whether the game is unavailable on local television or is being broadcast nationally, the number of games back the home team or visiting team is in the playoff race, and whether the visiting team is leading its division or wildcard race in September. Although the number of games back the home team is in the playoff race is omitted from the analysis, we keep whether the home team is leading its division, leading the wild card race, or is currently in contention for a playoff position to investigate the relationship between playoff positioning and attendance.

With the empirical model restricted this way, we then tested for auto-correlation and found none.[15] We also carried out a variance inflation factor (VIF) analysis. All of the variables remaining in the analysis were associated with a VIF under the standard baseline of 10 (and most were under 3) except for the three squared variables—temperature, distance, and probability of winning. The final empirical model, therefore, also omits the squared term on temperature and distance from the analysis. We keep the squared term on the probability of winning, however, because an important contribution of the paper is to compare our empirical results on the uncertainty of outcome hypothesis to what others have found.

Lastly, White and Breusch-Pagan tests for heteroskedasticity were carried out. Both tests found the presence of heteroskedasticity in the full data set but not when the sample was restricted by month or by team. When the heteroskedasticity stems from team effects, one solution is to force the variance-covariance matrix to be block-diagonal by team. This is imposed automatically, however, when the standard errors are made robust when clustering by series. Less can be done when the variance of error terms changes over time. To this end, the results in Table 4 that were estimated by month (columns 2–4) are probably better than the results for the entire sample (column 1).[16]

General Results

Several of the CNR-FE estimates using the entire sample of games (column 1 of Table 4) are worth highlighting. Monday through Thursday games draw significantly fewer fans than Saturday or Sunday afternoon games. Attendance is expected to be 2,080 less in September compared to July and August, 1,528 more when kids

Table 4. Censored-Normal Regression With Fixed-Effects—By Month, City Size, and Playoff Propensity

Explanatory Variable / Regression Technique	All	Month			Market Size			Playoff Experience		
		April/May	June/Aug	Sept.	Small	Med.	Large	0 Years	1–2 Years	3–7 Years
Time Factors:										
Played on Monday.	–5805	–7282	–5060	–5724	–6773	–5717	–3954	–7307	–5856	–4486
	(13.004)	(9.359)	(8.833)	(5.711)	(9.611)	(8.170)	(5.899)	(9.604)	(7.601)	(7.165)
Played on Tuesday.	–5056	–7020	–3835	–5599	–5888	–4755	–3616	–6031	–5634	–3714
	(13.064)	(10.447)	(7.629)	(6.779)	(9.846)	(7.582)	(6.049)	(9.533)	(8.128)	(6.642)
Played on Wednesday.	–4945	–6525	–3864	–5646	–6114	–3959	–4119	–5878	–5724	–3343
	(13.205)	(9.344)	(7.242)	(6.688)	(9.860)	(6.117)	(6.589)	(9.053)	(7.923)	(5.677)
Played on Thursday.	–4373	–5558	–3547	–4373	–5717	–3601	–2433	–5191	–4713	–3212
	(9.316)	(6.970)	(6.169)	(4.424)	(8.219)	(5.092)	(3.347)	(7.137)	(5.939)	(4.824)
Played on Friday.	–1458	–1087	–1445	–1027	–1064	–1935	–412	–2348	–223	–1979
	(4.451)	(1.572)	(2.852)	(1.206)	(1.758)	(3.073)	(0.636)	(3.640)	(0.326)	(3.339)
Day game played Monday–Friday.	644	911	1017	–645	1360	1717	–1573	547	762	386
	(1.554)	(1.356)	(1.971)	(0.640)	(2.129)	(2.606)	(2.777)	(0.812)	(1.143)	(0.644)
Saturday afternoon game.	500	278	1467	–1151	–1041	2271	653	4659	1354	–45
	(1.010)	(0.250)	(1.597)	(0.951)	(0.763)	(1.928)	(0.912)	(2.034)	(1.317)	(0.058)
Saturday night game.	2833	2118	3710	2613	3090	3470	954	2834	3305	2264
	(8.177)	(3.029)	(7.084)	(3.141)	(5.239)	(5.405)	(1.300)	(4.484)	(4.587)	(3.683)
Sunday night game.	–3008	–4274	–1837	–2964	–1661	–2391	293	–4180	824	–1855
	(3.334)	(2.229)	(1.607)	(1.404)	(0.757)	(1.919)	(0.218)	(2.313)	(0.449)	(1.617)
Played in April.	–3386	–436			–3375	–3887	–2623	–3435	–4600	–2120
	(4.760)	(0.947)			(3.909)	(4.576)	(2.613)	(3.484)	(4.977)	(2.687)
Played in May.	–2548				–1982	–3115	–2222	–2134	–4024	–1459
	(4.733)				(2.869)	(4.565)	(2.420)	(2.852)	(5.051)	(2.318)
Played in June.	–1276		–1724		–1343	–1576	–1033	–511	–2404	–994
	(2.615)		(4.335)		(2.238)	(2.533)	(1.311)	(0.758)	(3.688)	(1.695)

(continued)

Table 4 (continued)

Explanatory Variable / Regression Technique	All	Month			Market Size			Playoff Experience		
		April/May	June/Aug	Sept.	Small	Med.	Large	0 Years	1–2 Years	3–7 Years
Played in September.	–2080				–2395	–2114	–1429	–1562	–834	–2382
	(3.124)				(2.986)	(2.700)	(1.178)	(1.942)	(0.750)	(3.059)
Public schools are on summer vacation.	1528	934	2064	–1810	2830	609	919	2604	677	1341
	(3.676)	(0.663)	(5.674)	(1.519)	(4.907)	(1.123)	(1.205)	(4.226)	(1.083)	(2.656)
Played on Memorial Day, July 4, or Labor Day.	3733	5141	4122	3506	3863	3699	3791	6564	1859	1345
	(4.002)	(3.382)	(3.256)	(2.097)	(3.066)	(2.813)	(2.716)	(5.294)	(0.975)	(1.185)
Fan Interest Variables:										
Game-time temperature (outdoor stadium).	37	128	1	31	26	48	25	47	38	–8
	(2.039)	(5.071)	(0.037)	(0.764)	(1.022)	(1.880)	(1.062)	(1.877)	(1.299)	(0.360)
A divisional rivalry game.	1082	1393	1396	–348	3688	–533	1614	3709	1564	–1089
	(1.633)	(1.865)	(2.355)	(0.440)	(4.548)	(0.835)	(2.814)	(3.597)	(2.609)	(1.806)
An interleague game.	1845	–799	2854		1589	1520	1174	687	2481	1656
	(2.946)	(0.618)	(5.685)		(2.113)	(2.084)	(1.416)	(0.848)	(3.160)	(2.266)
An interleague rivalry game.	1997	4282	1736		2445	3357	4182	4036	168	2149
	(1.672)	(2.722)	(1.833)		(2.036)	(2.378)	(3.281)	(3.198)	(0.117)	(1.848)
Barry Bonds had between 753 and 755 home runs.	10156		9377		7723		7053		7758	
	(4.380)		(5.120)		(2.532)		(3.378)		(3.386)	
Home team's starting pitcher is a fan favorite.	419	1404	513	–435	808	461	511	1411	–13	800
	(1.232)	(2.258)	(1.136)	(0.544)	(1.413)	(0.832)	(0.933)	(1.908)	(0.024)	(1.623)
Giveaway/promotion game.	1798	2498	1252	1997	2248	1304	669	2669	800	1385
	(6.444)	(5.530)	(3.885)	(3.457)	(5.580)	(3.057)	(1.811)	(6.158)	(1.727)	(3.811)
Fireworks display at the stadium.	4812	4070	5398	2975	6555	4833	389	8139	2502	3276
	(7.766)	(4.071)	(8.202)	(2.884)	(8.150)	(5.953)	(0.462)	(9.382)	(3.045)	(3.949)
Visiting team qualified for the 2006 season playoffs.	1390	2218	1290	1142	2063	1808	–1104	1241	1043	1769
	(3.952)	(4.593)	(3.700)	(1.818)	(4.820)	(4.276)	(2.627)	(2.782)	(2.362)	(4.337)

(continued)

Table 4 (continued)

Explanatory Variable / Regression Technique	All	Month			Market Size			Playoff Experience		
		April/May	June/Aug	Sept.	Small	Med.	Large	0 Years	1–2 Years	3–7 Years
Visiting team was the Red Sox.	8725	6328	9511	10585	10211	11738	3182	11108	7343	7583
	(7.280)	(4.452)	(11.509)	(5.080)	(10.269)	(8.262)	(3.223)	(9.022)	(5.891)	(7.895)
Visiting team was the Yankees.	8770	6584	10643	8103	12465	4328	3434	13646	6326	3400
	(6.294)	(5.317)	(11.179)	(3.864)	(12.268)	(3.228)	(3.021)	(12.596)	(4.872)	(2.906)
Visiting team was the Cubs.	3936	1163	5123	3454	2967	5020	−2138	5392	3227	2739
	(4.700)	(1.069)	(6.550)	(2.610)	(3.216)	(5.515)	(1.785)	(5.947)	(2.791)	(3.104)
City Characteristics:										
Miles between stadiums in 1,000s.	−364	−193	−482	−1096	−567	−155	−262	−309	−564	−17
	(1.876)	(0.716)	(2.521)	(2.936)	(2.358)	(0.655)	(1.126)	(1.178)	(2.322)	(0.076)
Playoff Chances:										
Home team was leading its division.	993	−903	2038	3492	1127	2018	349	1502	3415	−3131
	(1.394)	(0.838)	(2.135)	(1.428)	(1.198)	(2.004)	(0.436)	(1.178)	(3.502)	(3.102)
Home team was leading its division × September game.	4353				8458	3992	1569		4503	1248
	(3.637)				(5.078)	(3.505)	(1.408)		(3.578)	(1.257)
Visiting team was leading its division.	1158	743	861	1663	1090	1527	−1001	1415	1241	999
	(2.959)	(1.429)	(2.185)	(2.248)	(2.324)	(3.188)	(1.997)	(2.837)	(2.455)	(2.211)
Home team was leading the wildcard race.	869	−620	1414	−1159	1749	1002	−671	565	3460	606
	(0.969)	(0.499)	(1.538)	(0.368)	(1.732)	(0.942)	(0.720)	(0.249)	(3.493)	(0.496)
Home team was leading wildcard × September game.	2512				609	3113	3210	7090	−2134	4046
	(2.014)				(0.331)	(0.917)	(1.928)	(1.364)	(1.151)	(2.282)
Visiting team was leading its wildcard race.	1323	380	1966	210	727	729	1586	816	2719	504
	(1.901)	(0.421)	(2.979)	(0.141)	(0.947)	(1.016)	(1.436)	(1.163)	(2.858)	(0.607)
Home team started the day in playoff contention.	800	303	−375	−2363	731	1186	553	1129	2971	18
	(1.628)	(0.458)	(0.638)	(2.485)	(1.264)	(1.885)	(0.869)	(1.706)	(3.517)	(0.036)

(continued)

Table 4 (continued)

Explanatory Variable / Regression Technique	All	Month			Market Size			Playoff Experience		
		April/May	June/Aug	Sept.	Small	Med.	Large	0 Years	1–2 Years	3–7 Years
Home team was in playoff contention × September game.	–490				188	–1258	1905	–509	–1175	–605
	(0.605)				(0.173)	(1.140)	(1.106)	(0.453)	(0.815)	(0.527)
Probability of Winning:										
Probability home team will win.	–45918	–35874	–51567	–39238	–9476	–44987	–15144	–7661	6529	–36957
	(2.615)	(1.271)	(2.745)	(1.409)	(0.422)	(1.972)	(0.550)	(0.310)	(0.231)	(1.543)
Probability home team will win squared.	42814	29959	48592	38415	9492	39675	11516	4688	–4863	36113
	(2.575)	(1.146)	(2.767)	(1.464)	(0.440)	(1.849)	(0.473)	(0.191)	(0.188)	(1.671)
Constant	48899	42947	50179	49931	38465	51458	46540	38513	12911	47351
	(10.328)	(5.522)	(9.551)	(6.561)	(6.437)	(8.298)	(5.662)	(6.058)	(1.582)	(7.080)
Number of observations.	2,196	616	1,180	400	904	835	457	756	761	679
Log likelihood.	–20,179	–5,823	–10,431	–3,743	–8,658	–7,269	–4,074	–7,098	–6,900	–6,020

Note: See text for variable definitions and sources. The absolute value of the *t* statistic is reported in parentheses beneath the estimated coefficient. Standard *t* statistic thresholds of 2.565, 1.960, and 1.645 indicating statistical significance at the 1%, 5%, and 10% levels, respectively, apply. Standard errors have been corrected for the cluster correlation problem of having MLB games scheduled in series.

are on summer vacation, and 3,733 more on holidays. Attendance increases slightly with temperature though an increase in 20° is predicted to increase attendance by only 740. Divisional rivalry games attract, on average, 1,082 more fans per game (p-value = .103). Interleague games draw 1,845 more fans to the ballpark, with interleague rivalry games drawing an additional 1,997 fans on top of that. Comparatively, MNR find that interleague games attract about 2,400 additional fans without accounting for interleague rivalry games. Fans also turned out to watch Barry Bonds challenge the home run record. Interestingly, attendance is not statistically sensitive to whether the home team starts a fan-favorite pitcher. Whereas HMZ found attendance to be 10,000 more at games with a giveaway or promotion, our results more closely match those of McDonald and Rascher (2000) in that giveaway dates and fireworks are associated with an additional 1,798 and 4,812 fans, respectively. Average attendance is 1,390 higher when the visiting team made the previous year's playoffs. And attendance is dramatically higher when the Red Sox (8,725 higher) or Yankees (8,770 higher) are the visiting team but is only slightly higher (3,936 higher) when the Cubs are the visiting team.

The estimated coefficient on the only remaining city characteristic shows that attendance falls by 364 for every 1,000 miles between the home team's and visiting team's stadiums. In comparison, KSH found that attendance decreased by about 1,850 for every 1,000 miles between stadiums.

The general result from the playoff chances variables is that attendance behavior changes in September. Attendance, April through August, is not related to whether the home team is leading its division, leading its wild card race, or is in playoff contention.[17] When September arrives, however, attendance is expected to increase by 4,353 when the home team is leading its division and by 2,512 when it is leading its wild card race. Home teams in playoff contention but not leading their division or wild card race, however, experience no increase in attendance in September (in fact, the point estimate is negative). Taken together, these results suggest that teams draw fans to the ballpark for the first 5 months of the seasons for reasons other than playoff chances or standings (e.g., to experience a new stadium, to watch a fireworks display, to see the Red Sox or Yankees, etc.). Attendance then responds to playoff chances in September, and dramatically so.

Finally, we consider the probability of the home team winning. The main purpose of KSH was to test the uncertainty of outcome hypothesis by including the probability of a home team win as a quadratic. Using data from the 1988 season, they estimated a positive coefficient on the linear term and a negative coefficient on the squared term, implying that attendance increased as the probability the home team would win increased up to 0.6, after which point attendance decreased. Using data from the 1996 season, Rascher similarly found that attendance peaked when the home team's probability of winning reached about 0.67.

Our results are strikingly different from KSH and Rascher, and do not support the uncertainty of outcome hypothesis. Our estimates consistently reveal a negative coefficient on the linear term and a positive term on the squared term, whether

measuring attendance in levels or in logs. Using the CNR-FE results, the estimated coefficients suggest that attendance falls until the probability the home team wins reaches 0.54, after which attendance increases. Our U-shaped pattern in attendance (in contrast to the inverted-U pattern suggested by the uncertainty of outcome hypothesis) is not sensitive to the inclusion or exclusion of any particular variable and even persists when attendance is regressed solely on the probability of winning and its square.[18]

That these results stand in stark contrast to previous findings raises an important puzzle to be solved by future research. Has a fundamental shift in fan preferences occurred? If so, when did it occur, and why? The CNR-FE results imply that 1,910 more fans attend, on average, when the probability the home team wins increases from 0.54 to 0.75. It remains unclear, however, what would happen if a home team's probability of winning were substantially more than 0.75. At a probability of 0.9 or 0.95, for example, it is possible that attendance would fall off sharply.

Results by Different Sample Criteria

Three criteria were chosen to reestimate the model on subsamples of the data in ways that might be important for MLB. The first criterion used to cut the data is months, as the process determining attendance in the early part of the season (April and May) is potentially different from the summer months (June, July, and August) which is potentially different from the last month of the season (September). The second criterion is market size. Many fans and owners, primarily in small markets, maintain that small market teams find it difficult if not impossible to compete with large market teams that receive substantially more local revenue from TV and radio contracts, merchandise sales, advertising, and ticket sales. Twelve teams play in MSAs with a 2006 population under 3.3 million (Baltimore, Cincinnati, Cleveland, Colorado, Kansas City, Milwaukee, Minnesota, Pittsburgh, San Diego, Seattle, St. Louis, Tampa Bay); 11 teams play in MSAs with populations between 4 and 6 million (Arizona, Atlanta, Boston, Detroit, Florida, Houston, Oakland, Philadelphia, San Francisco, Texas, Washington); and the remaining six teams play in New York, Los Angeles, and Chicago. The third criterion is playoff propensity. At issue is whether the fundamentals driving attendance depend on recent playoff performance. The three subsamples include the ten teams that have not made the playoffs in the seven years preceding the 2007 season (Baltimore, Cincinnati, Colorado, Kansas City, Milwaukee, Philadelphia, Pittsburgh, Tampa Bay, Texas, Washington/Montreal), the ten teams that qualified once or twice for the playoffs during that span (Arizona, Chicago Cubs, Cleveland, Chicago White Sox, Detroit, Florida, New York Mets, Los Angeles Dodgers, San Diego, Seattle), and the nine teams that qualified for the playoffs three or more times over the previous seven years (Atlanta, Boston, Houston, Los Angeles Angels, Minnesota, New York Yankees, Oakland, San Francisco, St. Louis).

There are several interesting coefficient estimates to note when the sample is split by time. Hosting a division rival increases attendance by about 1,400 from April through August, but there is no statistical impact on attendance from hosting a division rival in September. Interleague contests have no effect on attendance in April and May, unless it is an interleague rivalry game, in which case the net effect is statistically not different across time periods (p-value = .6245) with 3,483 more fans attending such a game in April and May and 4,590 more fans attending such a game during the summer. Giveaways and fireworks displays are both associated with greater attendance throughout the season, regardless of month.[19] Visiting teams that appeared in the 2006 playoffs helped sell tickets in April and May (2,218 greater attendance), but not as much later in the season (only 1,290 greater attendance June through August, and 1,142 more in September). In contrast, the negative effect of distance between stadiums is insignificant early on in the season but steadily grows to where attendance is expected to be almost 1,100 less for every 1,000 miles between stadiums by the end of the season. Leading one's division is insignificantly related to attendance during April and May but is associated with 2,038 greater attendance on average during the summer months and 3,492 (p-value of .154) greater attendance in September.[20] This suggests that money spent to improve a team, e.g., by signing high-priced free agents, pays off only if the team eventually leads its division or at worst is in playoff contention come September. Finally, all three subsamples find a U-shaped relationship between the probability the home team wins and attendance. The quadratic reaches its minimum at 0.60 in April and May, 0.53 during the summer months, and 0.51 in September.

Columns 5–7 of Table 4 report the regression results when the data are separated by market size. A small market team, for example, is expected to draw 6,773 fewer fans on a Monday night than it draws on Sunday afternoon, while a medium market team is expected to draw 5,717 fewer on Monday night and a large market team is expected to draw only 3,954 fewer. Small market teams do similarly poorly relative to medium or large market teams on Tuesday, Wednesday, and Thursday nights as well. Likewise, attendance falls off substantially more in September for small market teams than it does for large market teams. Attendance also responds more to children being on summer vacation or hosting a division rival in small markets than in medium or large markets. Interestingly, there is essentially no difference across city size in the effect of playing on a holiday or playing an interleague or interleague rivalry game.[21] Promotions and fireworks displays are also associated with greater increases in attendance in small markets than in medium or large markets. Small market fans also turn out to see the Red Sox, Yankees, and Cubs, much more so than the fans of large market teams. And while leading one's division fails to attract fans April through August for small and large markets (there is a statistically positive effect for medium markets but it is not statistically different from the point estimates for small or large market teams), small market fans show up in September when their team is leading the division (8,458 more fans) more so than in medium sized cities (3,992 more) and much more so than in large cities (a statistically insignificant 1,569 more fans).[22]

Overall, the main result when the data are split by market size is that attendance for small market teams is much more sensitive to factors surrounding the game and the ballpark experience than is attendance for large market teams. Whether this is because teams in large markets have more season ticket holders or whether this is simply because large market teams have more people to sell tickets too regardless of the game's characteristics, the implication is clear: the financial viability of small market teams requires paying attention to game characteristics to increase demand for tickets. The marketing departments of these teams must capitalize on their scheduling opportunities and help stimulate fan interest in attending games by providing an enticing environment. In fact, MLB has recently witnessed an increase in both of these strategies. In the last decade, almost every MLB team has initiated variable ticket pricing whereby tickets for "premium" games (depending on which team is visiting and/or the day-of-the week) cost more. In addition to charging premium prices for some games, the experience of attending a game in a newly constructed, state-of-the-art facility is being marketed by teams playing in such stadiums which all offer high quality food and beverage options along with an array of entertainment opportunities and amenities (e.g., TV lounges, kid areas, premium parking, luxury boxes, between inning contests, etc.) in addition to watching a baseball game.

The last three columns of Table 4 suggest that attendance is largely comparable between small market teams and teams that have failed to make the playoffs recently and between large market teams and teams that have made the playoffs recently. This is not surprising as there is a lot of overlap between small market teams and nonplayoff teams and between large market teams and playoff teams; still, there is not a one-to-one mapping between groups (e.g., only seven of the twelve small market teams never qualified for the playoffs), so comparing the results is informative. Consider, for example, that starting a fan-favorite pitcher has a significant effect on attendance for nonplayoff teams but no effect for playoff teams. Likewise, giveaways and fireworks displays, and having the Red Sox, Yankees, or Cubs visiting are all much more important for nonplayoff teams than for playoff teams.

The most interesting results in Table 4 regarding recent playoff history concerns fan reaction to playoff position for occasional (column 9) and perennial (column 10) playoff teams. Specifically, when leading one's division April through August, attendance increases by 3,415 per game on average for teams that occasionally made the playoffs in recent years but is actually expected to decrease by 3,131 for perennial playoff teams. And in September, attendance increases even more for occasional playoff teams but there is no statistically significant increase in attendance when a perennial playoff team is leading its division in September. Occasional playoff teams also experience an increase in attendance when leading their wild card race or are in playoff contention April through August (through there is no additional September effect) while there is no such effect for perennial playoff teams.[23] In short, it appears as if fans of occasional playoff teams are enthusiastic and are "on the bandwagon" from the start of the season, whereas fans of perennial playoff

teams may expect their team to be in the playoffs come October and therefore do not react to playoff position as much throughout the season.

5. Conclusion

Using 2,196 games during the 2007 MLB season, we explored the relationship between game attendance (i.e., tickets purchased) and a variety of game characteristics, including the schedule, fan preferences, being in playoff contention, and the chances that the home team will win the game. In this concluding section, we summarize the major findings.

First, the sign and statistical significance of estimated coefficients do not depend largely on whether the dependent variable is paid attendance or its natural log. The estimated coefficients, however, are sensitive to the choice of estimation technique. Whereas few meaningful differences present themselves when choosing between OLS or a censored normal regression, large differences arise between omitting or including fixed-effects. Because some games sell out and the fundamentals determining attendance likely vary in team-specific ways, attention is concentrated on the results from using a censored normal regression with home-team fixed-effects.

Next, the empirical specification is more complete than in any other article, and it is worth restating some of the original findings. Average attendance increases by roughly 1,500 when public schools are on vacation, by 3,700 on major holidays, and by 8,700 when the Red Sox or Yankees are in town. Whereas Meehan et al. (2007) found that attendance increases by about 2,400 for interleague games, we isolate the difference between interleague games and interleague rivalry games, such as the Cubs versus the White Sox, and find that paid attendance is expected to increase by roughly 1,850 at interleague games and by almost an additional 2,000 at interleague rivalry games. Average attendance also increases by almost 1,800 when there is a promotion or giveaway and by about 4,800 when a fireworks display follows the game. And attendance generally falls off in September unless the home team is leading its division or league's wild card race or is in playoff contention.

Another original and important finding of the paper is that the factors affecting attendance present themselves differently for small market teams than for large market teams as well as for perennial nonplayoff teams than for occasional or frequent playoff teams. Specifically, many of the effects noted above—higher attendance on holidays, on promoted days (giveaways or fireworks), when children are on summer vacation, when the Red Sox or Yankees are in town, and when the home team is contending for a playoff position—are all much more pronounced for small market teams than for large market teams. This pattern might stem simply from the ease with which large market teams sell tickets as they have many more people to sell tickets to. In 2007, for example, the smallest attendance at Fenway Park (Boston) was 34,424, at Yankee Stadium was 38,438, and at Busch Stadium (in St. Louis—a small market but frequent playoff team) was 42,029. In contrast, the

smallest attendance at Miller Park (Milwaukee) was 15,602, at RFK (Washington) was 15,611, and at Citizens Bank Park (Philadelphia—a medium sized town but a non-playoff team since 2000) was 23,526. This confirms that MLB, in its attempt to expand its fan base and to increase attendance, needs to approach its economic problems differently depending on which teams and cities are being considered. It also suggests that achieving the *Blue Ribbon Panel Report's* (Levin et al., 2000) definition of competitive balance—"every well-run club has a *regularly recurring reasonable hope of reaching postseason play*" (p. 5, original emphasis) —may have an important effect on attendance for the teams that otherwise rarely compete for playoff spots.

We conclude with a brief discussion of the uncertainty of outcome hypothesis. Intuition suggests that most fans want to see the home team win. If so, one should expect more fans to attend a game the more likely a home team win is. The uncertainty of outcome hypothesis contends that this behavior exists, but only up to a point as fans also want to see a competitive game. If correct, increasing the home team's probability of winning should be associated with an increase in paid attendance until the probability of a home win gets high enough that fans lose interest and attendance starts decreasing. Knowles et al. (1992) and Rascher (1999) both find evidence in support of the uncertainty of outcome hypothesis. Our results from the 2007 season, however, do not support the uncertainty of outcome hypothesis in that expected attendance continues to increase as the probability of a home team win increases. Simply put, fans in 2007 prefer to see the home team win, and demand to watch a game is greater the more likely it is that the home team will win. To illustrate, the third best drawing National League road team in 2007 was the Pittsburgh Pirates even though the Pirates were underdogs in all but two road games and were given at best a 40% chance of winning in more than half of their road games. This fundamental change in fan behavior or preferences is an interesting economic and social phenomenon to be explained with future research. It is also important for baseball to consider as it continues to discuss issues of competitive balance and tries to maintain and expand baseball's popularity.

Notes

1. Among others, Fort (2006), Fort and Maxcy (2003), Rottenberg (1956), Sanderson and Siegfried (2003), and Zimbalist (2002, 2003a) discuss varying aspects of competitive balance. Neale (1964) was the first to point out that the competitive nature of professional sports leagues is enigmatic in that within any given league there is a select group of team owners that restrict entry and aim to maximize profits but whose profits are directly tied to on-field competition.
2. Borland and Lye (1992), Jennet (1984), and Peel and Thomas (1988) all found evidence supporting the uncertainty of outcome hypothesis for other sports. See Cairns, Jennet, and Sloane (1986) for a review of the literature.
3. Borland and Lye (1992) also employed fixed-effects in their study of Australian Rules football, while Welki and Zlatoper (1994, 1999) used a tobit regression and a tobit regression with fixed-effects respectively.

4. Sellouts, defined as any game with attendance at or exceeding the posted stadium capacity for a baseball game, comprise 4.8% of the MNR sample and 7.5% of our sample.
5. Surprisingly, the literature has devoted little attention to whether the fundamentals driving attendance are constant throughout the season or across teams. Only Rascher (1999) estimated attendance over a shorter time period than a complete season.
6. The fixed-effects are necessarily omitted in the non fixed-effects models.
7. The results are not sensitive to which of the two games was used or to excluding doubleheaders from the analysis.
8. Rascher (1999) and others also omit the first 10 games of the season in order to include each team's number of wins in its last ten games. The 144 games represent just over 36% of all of the games played in April of 2007 and include all 29 opening days. Although the complete set of empirical results under-uses April data, the empirical section includes separate results for April/May versus June through August versus September games.
9. We emailed a beat reporter for a major local newspaper for each of the teams and asked them to comment on which team they thought the fans considered to be a division rival and which pitcher the fans would consider a local favorite. We only received 11 responses. Using these responses, our own opinions, and the opinions of several sports economists, we ultimately made subjective assignments as to which teams are division rivals and which pitchers are local fan favorites.
10. Because of the 162 game schedule in MLB, winning and losing streaks are a part of the game. How teams (and managers) respond to streaks is important for team performance and is potentially important for attendance (Fort and Rosenman, 1998, 1999). In another specification not reported here, we included the current winning and losing streaks of the home and visiting team. None of the coefficients were statistically significant and, at least for the home team, the streak variables are highly correlated with the number of wins in the home team's last ten games. Consequently, we have omitted streak variables from our current specification.
11. Whereas Hill et al. (1982) included all events and giveaways, we have limited promotions/giveaways to the more traditional giveaways such as bat day and fan appreciation night. Following McDonald and Rascher (2000), we do not regard pre-game festivities, Sunday Kids' Days, or Seniors' Days as promotional days.
12. The San Francisco Giants averaged an attendance of 35,954 as the road team, but much of this came when Barry Bonds approached the home run record, which is controlled for elsewhere.
13. The results are not sensitive to the specific definition of being in playoff contention as long as the definition is not so broad as to include all teams for most of the year or to include hardly any teams in September.
14. If both lines are listed at 105, the bookmaker believes each team has an equal chance of winning. Naturally we set the probability of the home team winning at 0.5 in this case.
15. The Durbin-Watson statistic was calculated using the error terms from the CNR-FE model and showed no autocorrelation with a value of 1.97. The variance inflation factor analysis was carried out on an OLS model rather than a CNR model.

16. Another potential solution to the heteroskedasticity problem is to weight the regression by stadium capacity assuming the source of the heteroskedasticity is coming from stadium size. There is, however, no evidence of this in our sample as only 7.5% of games sell out, and this happens in small and large parks. In any case, the results from weighting by stadium capacity are not substantially different from those reported in table 4.
17. A specification was estimated in which games back, leading one's division, leading the wildcard race, or being in playoff contention all enter the specification separately by month. None of the coefficient estimates in any month prior to September are statistically significant except for the home team leading the wildcard race in August.
18. Including the probability of winning as a cubic or quartic yields almost identical results. Fort and Quirk (1995) suggested a way to use the win percentages of both teams to approximate the probability that the home team will win. Our results are unchanged when we replace the probability variable implied by betting lines with Fort and Quirk's suggestion.
19. The only statistically significant differences across regressions are giveaways between April/May and the summer months (p-value $= 0.0378$) and fireworks displays between the summer months and September (p-value $= 0.0960$).
20. Although the only statistically significant difference across regressions is between the summer months and April/May, the point estimates are substantial and monotonic across time.
21. Interleague rivalry games are expected to have 4,034 greater attendance in small markets, 4,877 greater attendance in medium markets, and 5,356 greater attendance in large markets; however, none of these differences are statistically significant at the 10% level.
22. The differences across regressions in the estimated coefficient on leading one's division in September has a p-value of 0.0964 between April/May and the summer months and 0.0025 between April/May and September.
23. Oddly, the only boost to attendance for perennial playoff teams in terms of playoff position comes from leading one's wild card race in September.

Authors' Note

We thank Woody Eckard, Rod Fort, Mike Haupert, Jim Meehan, Dan Rascher, John Siegfried, and Andy Zimbalist for comments on an earlier draft. Comments from two anonymous referees also improved the paper. Financial assistance was provided by the Richter Apprentice Scholarship Program at Lake Forest College.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interests with respect to their authorship or the publication of this article.

Funding

Financial assistance was provided by the Richter Apprentice Scholarship Program at Lake Forest College.

References

Borland, J., & Lye, J. (1992). Attendance at Australian Rules football: a panel study. *Applied Economics*, *24*, 1053-1058.

Butler, M. R. (1995). Competitive balance in Major League Baseball. *The American Economist*, *39*, 46-52.

Cairns, J., Jennet, N., & Sloane, P. J. (1986). The economics of professional team sports: A survey of theory and evidence. *Journal of Economic Studies*, *13*, 3-80.

Chicago Tribune, (2007). Daily issues April 1 through September 30.

Demmert, A. G. (1973). *The Economics of Professional Team Sports.* Lexington, MA: D.C. Heath.

Depken, C. A. (1999). Free-agency and the competitiveness of Major League Baseball. *Review of Industrial Organization*, *14*, 205-217.

Eckard, E. W. (2001a). Baseball's Blue Ribbon Report: Solutions in search of a problem. *Journal of Sports Economics*, *2*, 213-227.

Eckard, E. W. (2001b). Free agency, competitive balance, and diminishing returns to pennant contention. *Economic Inquiry*, *39*, 430-443.

El-Hodiri, M., & Quirk, J. (1971). An economic model of a professional sports league. *Journal of Political Economy*, *79*, 1302-1319.

Fort, R. (2006). Competitive balance in North American professional sports. In J. Fizel (Ed.), *Handbook of Sports Economics Research* (pp. 190-206). Armonk, NY: M.E. Sharpe, Inc.

Fort, R., & Maxcy, J. (2003). Comment on "Competitive balance in sports leagues: an introduction." *Journal of Sports Economics*, *4*, 154-160.

Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature*, *33*, 1265-1299.

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

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.

Hadley, L., Ciecka, J., & Krautmann, A. C. (2005). Competitive balance in the aftermath of the 1995 players' strike. *Journal of Sports Economics*, *6*, 379-389.

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

Horowitz, I. (1997). The increasing competitive balance in Major League Baseball. *Review of Industrial Organization*, *12*, 373-387.

Humphreys, B. R. (2002). Alternative measures of competitive balance in sports leagues. *Journal of Sports Economics*, *3*, 133-148.

Jennet, N. (1984). Attendances, uncertainty of outcome and policy in Scottish league football. *Scottish Journal of Political Economy*, *31*, 176-198.

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

Levin, R. C., Mitchell, G. J., Volcker, P. A., & Will, G. F. (2000). *The Report of the Independent Members of the Commissioner's Blue Ribbon Panel on Baseball Economics.* New York: Major League Baseball.

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.

Meehan, Jr. J. W., 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.

Noll, R. G. (1974). Attendance and price setting. In R. G. Noll (Ed.), *Government and the Sports Business* (pp. 115-157). Washington, DC: The Brookings Institution.

Peel, D., & Thomas, D. (1988). Outcome uncertainty and the demand for football: an analysis of match attendances in the English football league. *Scottish Journal of Political Economy*, *35*, 242-249.

Rascher, D. (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.

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

Sanderson, A. R., & Siegfried, J. J. (2003). Thinking about competitive balance. *Journal of Sports Economics*, *4*, 255-279.

Schmidt, M. B. (2001). Competition in Major League Baseball: the impact of expansion. *Applied Economics Letters*, *8*, 21-26.

Schmidt, M. B., & Berri, D. J. (2001). Competitive balance and attendance: the case of Major League Baseball. *Journal of Sports Economics*, *2*, 145-167.

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

Siegfried, J. J., & Eisenberg, J. D. (1980). The demand for minor league baseball. *Atlantic Economic Journal*, *8*, 59-69.

Team Marketing Report (2007). *Fan Cost Index* [MLB]. Retrieved on June 22, 2008 from at www.teammarketing.com.

U.S. Census Bureau (2007). *Statistical Abstract of the United States: 2008* (127th ed). Washington, DC: Author.

Utt, J., & Fort, R. (2002). Pitfalls to measuring competitive balance with Gini coefficients. *Journal of Sports Economics*, *3*, 367-373.

Welki, A. M., & Zlatoper, T. J. (1994). U.S. professional football: the demand for game-day attendance in 1991. *Managerial and Decision Economics*, *15*, 489-495.

Welki, A. M., & Zlatoper, T. J. (1999). U.S. professional football game-day attendance. *Atlantic Economic Journal*, *27*, 285-298.

Zimbalist, A. S. (2002). Competitive balance in sports leagues: an introduction. *Journal of Sports Economics*, *3*, 111-121.

Zimbalist, A. S. (2003a). Competitive balance conundrums. *Journal of Sports Economics*, *4*, 161-163.

Zimbalist, A. S. (2003b). *May the Best Team Win: Baseball Economics and Public Policy*. Washington, DC: The Brookings Institution Press.