Economics Bulletin

Volume 41, Issue 2

American Baseball Fans Do Not Influence Game Outcomes

Timothy E. Zimmer
University Of Indianapolis

Allison Snyder
University of Indianapolis

Lawrence Bukenya
University of Indianapolis

Abstract

The paper examines U. S. Major League Baseball (MLB) data to evaluate home-field advantage and the relationship between fan attendance and game outcomes. Due to the COVID-19 pandemic, fans were not allowed in stadiums during the 2020 MLB regular season. The results indicate the absence of fans did not statistically alter home-field advantage when controlling for quality of play.

There is no funding for this research
Citation: Timothy E. Zimmer and Allison Snyder and Lawrence Bukenya, (2021) "American Baseball Fans Do Not Influence Game Outcomes", *Economics Bulletin*, Vol. 41 No. 2 pp. 741-750.
Contact: Timothy E. Zimmer - zimmert@uindy.edu, Allison Snyder - snyderah@uindy.edu, Lawrence Bukenya - lawrence.bukenya@cummins.com.
Submitted: January 15, 2021. **Published:** April 09, 2021.

American Baseball Fans Do Not Influence Game Outcomes

1. Introduction

The relationship between fans and on-field play in sports is a source of anecdotal interest. Many baseball fans are proudly superstitious in their efforts to bring favor to a particular team. The use of "Rally Caps" and other techniques commonly seen at games have a long history that adds to its allure and enjoyment for many baseball fans. However, academic research or empirical study supporting the relationship between fans and on-field play remains more elusive.

Direct fan involvement at profession sports games is rare though examples exist (*Winning Time, Catching Hell*). The more common method of fan involvement is indirect. Teams incentivize fans to engage in communal celebration to inspire one team or distract another. The effectiveness of this fan behavior or even their presence to alter game outcomes is challenging to determine.

Since MLB stadiums remained empty of fans during the 2020 regular season due to the ongoing global pandemic, it naturally raised the question about the impact on the game. O'Connell (2020) pondered whether the lack of fans diminishes home-field advantage or alters the quality of play. A home team enjoys the encouragement of the fans and the lack of these fans might suppress the typical advantage. The virus restrictions of the 2020 baseball season provide an opportunity to empirically test whether fan presence affects outcome by influencing home-field advantage.

2. Literature Review

Professional sports are reliant on fan support for their financial well-being. Three major sources of revenue include venue attendance, team merchandising, and broadcast (Buraimo, 2008). A significant portion of research is devoted to the relationship between game outcomes and fan support. For example, Demmert (1973), Noll (1974), and Horowitz (1978) in seminal work established the relationship between attendance and a team's on-field success. They found that a high winning percentage fostered increased fan interest and attendance. Davis (2008) subsequently concluded that the positive impact of winning was fleeting and did not extend into future periods. This is interesting as Whitney (1988) found the prospect of postseason play increased fan anticipation and induced increased attendance. Alternatively, Zimmer (2018) argued that an abundance of team success desensitized fans and reduced attendance.

Research also links quality-of-play and attendance. For example, increased attendance resulted from home run hitting and quality pitching (Greenstein & Marcum, 1981; Marburger, 1997; Horowitz, 2007). Krautmann and Hadley (2006) also found a modest influence of parity-of-play on attendance.

Although significant research links winning and quality-of-play with fan attendance, a smaller subset reverses causality and focuses on how fan attendance affects winning and quality-of-play. Of this literature, a number of studies assess the relationship of fans and outcomes in English football (soccer) by assessing home-field advantage. Pollard (2006) indicates a strong home-field advantage and concludes that fans provide advantage. Neville et al. (1996) also finds attendance is a significant advantage on game outcomes. However, some doubt exists about the crowd hypothesis as provided by Pollard and Pollard (2005). Competing research suggests that fan supported home-field advantage is waning (Koyama and Reade, 2009). The evidence for the diminishing effects of home advantage is further supported with from the recent pandemic season of German professional soccer played without fans allowed in stadiums (Tilp and Thaller, 2020). This study is an extension of this line of inquiry with a focus on baseball, since the relationship between fans and play varies by sport.

3. Methodology

The study uses empirical results from a fixed-effects model with six years of data for 30 MLB teams. The data include five seasons of pre-pandemic data (2015 through 2019 seasons) and the pandemic 2020 season (Baseball Almanac, 2020; Baseball Reference, 2020). Game level observations generate an unbalanced panel dataset owing to the unequal number of games across seasons because of the pandemic shortened 2020 season. Table 1 lists MLB franchises. The list of variables and summary statistics, excluding team binaries, are presented in Table 2.

Table 1

#	Team	#	Team
1	Arizona Diamondbacks	16	Milwaukee Brewers
2	Atlanta Braves	17	Minnesota Twins
3	Baltimore Orioles	18	New York Mets
4	Boston Red Sox	19	New York Yankees
5	Chicago Cubs	20	Oakland Athletics
6	Chicago White Sox	21	Philadelphia Phillies
7	Cincinnati Reds	22	Pittsburg Pirates
8	Cleveland	23	San Diego Padres
9	Colorado Rockies	24	San Francisco Giants
10	Detroit Tigers	25	Seattle Mariners
11	Houston Astros	26	St. Louis Cardinals
12	Kansas City Royals	27	Tampa Bay Rays
13	Los Angeles Angels	28	Texas Rangers
14	Los Angeles Dodgers	29	Toronto Blue Jays
15	Miami Marlins	30	Washington Nationals

Table 2

Summary Statistics Seasons 2015 to 2020
(Excluding Team Binaries)

	Dependent Variables	Notation	Obs.	Mean	Std. Dev.	Min.	Max.
1	Run Difference	$RunDiff_{g,i,s}$	26,118	0.007	4.480	-21	21

	Independent Variables		Obs.	Mean	Std. Dev.	Min.	Max.
1	Home Game Binary	$Home_{g,i,s}$	26,118	0.500	0.500	0	1
2	Home Game Binary Plus 2020 Season	$Home2020_{g,i,s}$	26,118	0.035	0.184	0	1
3	Win Percentage	$WinPct_{g,i,s}$	26,118	0.501	0.083	0.290	0.717
4	Opponents Win Percentage	$OWinPct_{g,i,s}$	26,118	0.500	0.083	0.290	0.717
5	Extra Innings	$ExInn_{g,i,s}$	26,118	0.158	0.786	-4	10
6	Day / Night Game Binary	$DayNight_{g,i,s}$	26,118	0.333	0.471	0	1
7	Runs Per Game	$RG_{g,i,s}$	26,118	4.546	0.469	3.610	5.820
8	Runs Allowed Per Game	$RAG_{g,i,s}$	26,118	4.536	0.555	3.240	6.060
9	Opponents Runs Per Game	$ORG_{g,i,s}$	26,118	4.540	0.476	3.540	5.820
10	Opponents Runs Allowed Per Game	$ORAG_{g,i,s}$	26,118	4.538	0.555	3.240	6.060

The COVID-19 pandemic witnessed sports leagues worldwide alter operations and adopt mitigation standards to help ensure the safety of both participants and spectators. MLB conducted the entire 2020 regular season without fans allowed at stadiums.

The 2020 MLB season was different in many respects beyond the absence of fans. A few examples included a late start to the season that may have altered preparation and player routines. The season was truncated and travel schedules condensed which may have affected player performance. The pandemic also affected the player availability and player rotations. The many differences between the 2020 season and prior seasons make simple comparisons across seasons an incomplete assessment method.

This study examines home-field advantage while controlling for team skill. If fans influence game outcomes, their absence should diminish home-field advantage. Performance measures control for atypical play resulting from the pandemic. Estimating home-field advantage in this manner isolates the influence of fan absence.

The fixed-effects model accounts for time variation in panel data. The base model is equation (1).

(1)

$$RunDiff_{g,i,s} = \beta_0 + \beta_1 Home_{g,i,s} + \beta_2 Home2020_{g,i,s} + \beta_3 WinPct_{g,i,s} + \beta_4 OWinPct_{g,i,s} + \beta_5 ExInn_{g,i,s} + \beta_6 DayNight_{g,i,s} + \beta_7 RG_{g,i,s} + \beta_8 RAG_{g,i,s} + \beta_9 ORG_{g,i,s} + \beta_{10} ORAG_{g,i,s} + \sum_{z=11}^{40} \beta_z (X_i) + \varepsilon_{g,i,s}$$

$$where\ g = game, i = team, s = season$$

The run difference is the margin between runs scored by home and opposing team as noted in equation (2). If the run difference is positive, the home team scored more runs in the game. Alternatively, if the run difference is negative, the opposing team scored more runs. Run difference ($RunDiff_{g,i,s}$) is the dependent variable.

(2)

$$RunDiff_{g,i,s} = (Runs\ Scored\ by\ Team - Runs\ Scored\ by\ Opposing\ Team)_{g,i,s}$$

Independent variables include factors that influence run difference. The home game binary variable ($Home_{g,i,s}$) assesses the value of the home-field advantage. It is affirmative (or 1) when the team is hosting. The variable represents the home team advantage unattributed to controls including the quality and skills of teams involved. If the home-field advantage exists, the variable coefficient should be significant and positive. The home game binary plus 2020 season ($Home2020_{g,i,s}$) is affirmative (or 1) when the team is hosting a game during the 2020 Season. If the absence of fans is detrimental to home-field advantage, the variable should have a negative impact as it diminishes the standard home-field advantage effect. The variable coefficient should be significant and negative.

The fan theory being investigated (O'Connell, 2020) is that fans alter game outcomes after controlling for the ability of teams involved. If true, the coefficient of the variable ($Home_{g,i,s}$) should be significant and positive as home fans contribute to victory. The absence of fans in the 2020 season should be detrimental to this advantage causing the variable ($Home2020_{g,i,s}$) to be significant and negative. The lack of fans during the 2020 season should lessen the typical home-field advantage.

The model includes variables to control for the quality of the teams playing. These include the winning percentage of the team ($WinPct_{g,i,s}$) and the opponent team's winning percentage ($OWinPct_{g,i,s}$). The variables are end-of-season winning percentage applied to each game during the prior season. These variables represent proxies for the overall quality of a team and the opponent in each game.

Variables are included for games that exceed the standard duration and the time of day. The extra innings variable ($ExInn_{g,i,s}$) counts innings beyond regulation. The variable is zero for a standard game. The day/night binary variable ($DayNight_{gi,i,s}$) assesses start time to control for play variation due to playing under artificial lights. The variable is affirmative (or 1) for day games.

Team performance measures assess the ability of teams to score and prevent scoring. The four performance variables are end-of-season measures applied to each game during the prior season. The variables control for the overall offensive and defensive ability of a team and the opponent in each game.

The runs per game ($RG_{g,i,s}$) variable is the average runs scored per game for the team in a given season. The variable is a measure of the team's offensive proficiency to score runs. The runs allowed per game ($RAG_{g,i,s}$) variable is the average runs scored per game against the team in the given season. The variable is a measure of the team's defensive proficiency to prevent scored runs against their team. The opponents runs per game ($ORG_{g,i,s}$) variable is the average runs scored per game for the opponent team in a given season. The variable is a measure of the opponent team's offensive proficiency to score runs. The opponents runs allowed per game ($ORAG_{g,i,s}$) variable is the average runs scored per game against the opponent team in the given season. The variable is a measure of the opponent team's defensive proficiency to prevent scored runs against their team.

The relationship between a team and its fans is unique. Fan cheering and engagement differ by location and local custom. Stadium characteristics are unique to each team due to non-uniformity in stadium layout that may alter play or fan proximity and interaction with players, officials, and other game participants. Fan loyalty varies between markets and some teams maintain higher attendance as a result. Therefore, it possible that teams with loyal fans and full stadiums may be affected more by their absence during the pandemic (Fischer & Haucap, 1981). To assess these impacts and others, the model includes team binaries (X_i). The binaries capture team specific effects on home-field advantage before and during the pandemic. If identified team effects change during the pandemic, additional efforts can be taken to determine the source.

The model is estimated three times using complete data and subsets for robustness. The first (main or model #1) estimation includes all six seasons. The second estimation (model #2) includes the first five seasons (2015-2019) and represents the pre-pandemic period with fans in attendance. As the data excludes the 2020 season, the variable ($Home2020_{g,i,s}$) is omitted. The third estimation (model #3) only includes the last season (2020) and represents the pandemic period without fans in attendance. The variable ($Home2020_{g,i,s}$) is omitted as the subset only contains the 2020 season.

4. Results

The detailed model results are in Table 3.

Table 3

Fixed-Effects Results

		#1 Main Model 2015 to 2020 Seasons Run Difference ($RunDiff_{g,i,s}$)		#2 Subset Model 2015 to 2019 Seasons Run Difference ($RunDiff_{g,i,s}$)		#3 Subset Model 2020 Season Run Difference ($RunDiff_{g,i,s}$)	
	Independent Variables	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
1	Home Game Binary ($Home_{g,i,s}$)	0.261 **	0.056	0.261 **	0.056	0.427 *	0.214
2	Home Game Plus 2020 Season ($Home2020_{g,i,s}$)	0.165	0.212	Omitted		Omitted	
3	Win Percentage ($WinPct_{g,i,s}$)	0.093	0.566	0.086	0.573	1.104	24.449
4	Opponents Win Percentage ($OWinPct_{g,i,s}$)	-0.088	0.522	-0.091	0.530	-0.497	3.179
5	Extra Innings ($ExInn_{g,i,s}$)	-0.002	0.035	-0.002	0.036	-0.001	0.141
6	Day/Night Game Binary ($DayNight_{g,i,s}$)	-0.002	0.057	-0.002	0.059	0.009	0.237
7	Runs Per Game ($RG_{g,i,s}$)	0.909 **	0.100	0.894 **	0.114	Omitted	
8	Runs Allowed Per Game ($RAG_{g,i,s}$)	-0.900 **	0.090	-0.916 **	0.097	Omitted	
9	Opponents Runs Per Game ($ORG_{g,i,s}$)	-0.924 **	0.076	-0.922 **	0.080	-0.983 **	0.364
10	Opponents Runs Allowed Per Game ($ORAG_{g,i,s}$)	0.898 **	0.073	0.910 **	0.077	0.748 **	0.403
11	Arizona Diamondbacks	0.068	0.206	0.057	0.215	-0.024	1.006
12	Atlanta Braves	-0.013	0.225	0.003	0.233	0.701	3.334
13	Baltimore Orioles	0.147	0.212	0.160	0.220	-0.118	1.083
14	Boston Red Sox	0.080	0.224	0.100	0.239	-0.596	1.421
15	Chicago Cubs	0.039	0.221	0.044	0.231	0.112	2.971
16	Chicago White Sox	0.019	0.209	0.015	0.216	0.771	3.349
17	Cincinnati Reds	0.001	0.209	0.101	0.217	-0.246	1.816
18	Cleveland	-0.036	0.222	-0.049	0.231	0.454	3.346
19	Colorado Rockies	0.073	0.222	0.076	0.236	-0.854	0.843
20	Detroit Tigers	0.009	0.211	-0.003	0.218	-0.873	1.522
21	Houston Astros	0.071	0.225	0.084	0.238	-0.071	1.005
22	Kansas City Royals	0.013	0.209	-0.001	0.217	-0.220	0.878
23	Los Angeles Angels	0.118	0.211	0.127	0.219	0.395	0.848
24	Los Angeles Dodgers	0.028	0.227	0.020	0.235	1.862	6.577
25	Miami Marlins	0.028	0.210	0.007	0.217	-0.377	1.787
26	Milwaukee Brewers	0.065	0.211	0.068	0.219	-0.373	1.128
27	Minnesota Twins	-0.024	0.215	-0.199	0.226	0.655	3.752
28	New York Mets	0.011	0.212	0.003	0.219	-0.153	0.833
29	New York Yankees	0.097	0.222	0.106	0.234	0.687	2.550
30	Oakland Athletics	0.065	0.212	0.091	0.221	0.260	3.742
31	Philadelphia Phillies	Omitted		Omitted		Omitted	
32	Pittsburg Pirates	0.059	0.211	0.057	0.218	-1.068	3.350
33	San Diego Padres	0.065	0.210	0.056	0.217	1.283	4.106
34	San Francisco Giants	0.077	0.213	0.055	0.220	0.286	1.007
35	Seattle Mariners	0.091	0.210	0.096	0.218	-0.686	0.733
36	St. Louis Cardinals	0.045	0.217	0.047	0.226	-0.027	1.821
37	Tampa Bay Rays	0.108	0.217	0.106	0.224	0.771	5.361
38	Texas Rangers	0.074	0.214	0.097	0.226	-1.315	2.159
39	Toronto Blue Jays	0.110	0.213	0.115	0.222	-0.068	2.158
40	Washington Nationals	-0.034	0.221	-0.040	0.232	Omitted	
	Intercept	-0.111		-0.025		0.559	
		F (39 , 26073) = 39.32 *Prob > F 0.0000* *Observations 26,118*		*F (38 , 24251) = 37.20* *Prob > F 0.0000* *Observations 24,294*		*F (35 , 1788) = 3.4* *Prob > F 0.0000* *Observations 1,824*	

* Significant at 5%, ** Significant at 1%

The models indicate home-field advantage. After controlling for factors including quality of play performance measures, the coefficient for the home game binary ($Home_{g,i,s}$) is positive and significant in all models. As indicated in the main model (model #1), the home team enjoys an advantage of approximately 0.26 runs per game. The variable ($Home2020_{g,i,s}$) is positive but not significant. If the fan theory is correct, the lack of fans should be detrimental to the home team causing the coefficient to be significant and negative. This is not the case and the absence of fans in the 2020 season has no effect.

The influence of home-field advantage is shown to be consistent in the pre-pandemic estimation (model #2) using the subset of seasons (2015 to 2019). The coefficient for the home game binary ($Home_{g,i,s}$) is positive (0.26 runs) and significant. Interestingly, in the pandemic estimation (model #3) using the 2020 season, home-field advantage increases. The coefficient for the home game binary ($Home_{g,i,s}$) is positive (0.43 runs) and significant. This indicates that the home-field advantage increases without fans, which would be the opposite of fan theory. This would explain the positive coefficient of the variable ($Home2020_{g,i,s}$) in the main model (model #1). However, the increase is not statistically significant.

Variables are included to control for overall team quality. The winning percentage of the team ($WinPct_{g,i,s}$) and the opponent team's winning percentage of the ($OWinPct_{g,i,s}$) are used to assess team quality. The coefficients are not significant in any models perhaps due to correlation with the other performance measures.

Control variables for extra innings and day versus night games are not significant in any model. The binary variable for extra innings ($ExInn_{g,i,s}$) is not an influence on run difference. Likewise, the binary for day versus night games ($DayNight_{gi,i,s}$) is not significant.

The four performance measures are all statistically significant. The runs per game variable ($RG_{g,i,s}$) that measures the team's offensive proficiency. The variable coefficient is positive and significant. A team's ability to score runs increases the run difference. The runs allowed per game variable ($RAG_{g,i,s}$) measures the team's defensive proficiency. The variable coefficient is negative and significant. A team's inability to prevent runs scored against them decreases the run difference. The opponents runs per game variable ($ORG_{g,i,s}$) measures the opponent team's offensive proficiency. The variable coefficient is negative and significant. The ability of the opponent team to score runs decreases the run difference. The opponents runs allowed per game variable ($ORAG_{g,i,s}$) measures the opponent team's defensive proficiency. The variable coefficient is positive and significant. The inability of the opponent to prevent runs increases the run difference.

The team binary variables are all insignificant. No team specific effects were identified either prior to or during the pandemic. There is no evidence that a team gained extra home-field advantage relative to other teams prior to the pandemic or during the pandemic from a team specific source. Stadium layout, fan loyalty or behavior does not appear to influence the home-field of one team in excess of others.

5. Discussion

MLB franchises desire the monetary support of fans through event attendance. Fans attend games and cheer partially in the hope of encouraging a team to victory. What remains unclear is whether the presence of fans and by extension their activity is entertainment or actually creates an advantage. This study addresses the issue by using fixed-effect regression models to examine the influence of attendance on baseball outcomes. Specifically, this study examines the change in home-field advantage due to fan absence in the 2020 regular season while controlling for factors including quality of play. Results indicated that home-field advantage is statistically significant regardless of fan attendance. In recent seasons, the home team enjoys an advantage of approximately 0.26 runs per game given the quality and performance of the team and its opponent. Importantly, the advantage did not change because of the 2020 season. Thus, the lack of fans did not influence game results.

The study finds no direct evidence that fan presence affects baseball game outcomes. The experience of going to a baseball game is enjoyable. The results liberate fans to appreciate the game free from the burden of concern that their presence or action influences the team they support. Relax and enjoy the baseball game in whatever fashion is most appealing because it will does not influence the outcome.

The results of this study are specific to baseball and extrapolating them to other sports is problematic. The interaction and influence of fans on game participants differ by sport. It is also challenging to replicate this analysis as other sports may have utilized different mitigation strategies to accommodate the circumstances of the 2020 season. Independent analysis of each sport should be undertaken as results likely differ. As noted in the paper, work exists on varying sports with differing results. Additional research could investigate the underlying source of these differences.

6. References

Baseball Almanac (2020) Team Statistics Retrieved from http://www.baseball-almanac.com on October 1, 2020.

Baseball Reference (2020) Team Statistics Retrieved from http://www.baseball-reference.com on October 1, 2020.

Buraimo, B. (2008) "Stadium attendance and television audience demand in English League Football" *Managerial and Decision Economics* 29, 513-523.

Catching Hell. Directed by Alex Kibney (2011) ESPN, 30 for 30 Production.

Davis, M.C. (2008) "The interaction between baseball attendance and winning percentage: A VAR analysis" *International Journal of Sport Finance* 3(1) 58-73.

Demmert, H.G. (1973) *The economics of professional team sports.* Lexington Books, Lexington, Mass.

Fischer, K. and Haucap J. (2020): Does Crowd Support Drive the Home Advantage in Professional Soccer? Evidence from German Ghost Games during the COVID-19 Pandemic, DICE Discussion Paper.

Greenstein, T.N., and Marcum, J.P. (1981) "Factors affecting attendance of Major League Baseball: I. Team performance" *Review of Sport and Leisure* 6, 21-34.

Horowitz, I. (1978) "Market entrenchment and the sports broadcasting act" *American Behavioral Scientist* 21, 415-430.

Horowtiz, I. (2007) "If you play well they will come: And vice versa: Bidirectional causality in Major League Baseball" *Managerial and Decision Economics* 28 (2, Mar.), 93-105.

Koyama, M. and Reade, J.J. (2009) "Playing Like the Home Team: An Economic Investigation into Home Advantage in Football" *International Journal of Sport Finance* 4(1) 16-41.

Krautmann, A., and Hadley L. (2006) "Dynasties versus Pennant races: competitive balance in Major League Baseball" *Managerial and Decision Economics* 27, 287-292.

Marburger, D.R. (1997) "Optimal ticket pricing for performance goods" *Managerial and Decision Economics* 18, 375-382.

Neville, A.M., Newel, S.M., and Gale, S. (1996) "Factors Associated with Home Advantage in English and Scottish Soccer Matches" *Journal of Sports Science* 14(2), 181-186.

Noll, R.G. (1974) *Government and the Sports Business*. Noll, R.G. (ed.). The Brookings Institution: Washington, D.C. 115-157.

O'Connell, R. (2020) "Baseball in empty stadiums is weird. How will it affect outcomes?" The New York Times (31, Jul.). Retrieved from http://www.nytimes.com

Pollard, R. (2006) "Worldwide Regional Variations in Home Advantage in Association Football" *Journal of Sports Sciences* 24(3), 969-973.

Pollard, R. and Pollard G. (2005) "Long-Term Trends in Home Advantage in Association Football" Journal of Sports Sciences 24(3), 231-240.

Tilp, M. and Thaller S. (2020) "Covid-19 has Turned Home Advantage in Home Disadvantage in the German Soccer Bundesliga" doi: https://doi.org/10.3389/fspor.2020.593499

Whitney, J.D. (1988) "Winning games versus winning championships: The economics of fan interest and team performance" *Economic Inquiry* 26(4), 703-724.

Winning Time: Reggie Miller vs. The New York Knicks. Directed by Dan Klores. (2010) ESPN, 30 for 30 Production.

Zimmer, T.E. (2018) “The negative influence of prior World Series victory on attendance: Winning and increased fan apathy” *Journal of Global Sport Management* DOI: 0.1080/24704067.2018.1442237.