The Effects of Closed-Door Baseball Games

Nathan Hong[1] and Danilo Tadeo[#]

[1]Yongsan International School of Seoul, South Korea
[#]Advisor

ABSTRACT

With the ongoing trend of closed-door baseball games, we researched to discover the actual effect of the crowd during this isolated season. Taking advantage of the current data provided by three reliable sports-related websites, we obtained significant results comparing previous years' home team winning percentage data to this year's. Initially, we found that the overall home team winning percentage of 2020 (the crowd-less season) was lower than that of the previous year (with crowds). From this finding, additional analyses were conducted, each controlling for another different factor of home winning percentage. Analysis 1 found that holding for skill difference amongst teams, the matchup of teams with a similar skill level showed a positive effect of crowds, aiding the home team wins. Analysis 2 found that home-field advantage of home teams in 2020 had decreased compared to 2019, proving once again that crowd does have a positive effect on home winning percentage. Thus, it was ultimately determined through the multiple qualifications and subsequent tests that there was evidence of a positive correlation between crowds and increased home winning.

Introduction

Background

Hearing the cheers and boos of the crowd is an essential part of popular sports and is a necessary factor in winning games. With team sports usually divided into home and away, the role of the crowd is more prominent as it becomes an asset to the home team. Or does it? In the past and current year, countless sports, including baseball, have turned closed-door or crowd-less (Lee, 2020). With this unexpected situation came the perfect opportunity to compare the direct effect of crowd-less baseball and its subsequent impact on the outcomes of games. To observe such, data of home/away winning percentages by season and team were collected from Korea Baseball Organization (KBO, www.koreabaseball.com), a reliable data set because it is the largest league in Korea with the most games.

Objectives

This study aims to determine the objectively higher winning percentage across different years/seasons, consider winning percentages after controlling for skill differences throughout the season, and observe the fluctuations of winning percentage factoring in home-field advantage.

Terminology

1. Throughout this research, KBO stands for the Korea Baseball Organization and is composed of 10 teams referenced by their initials—NC, kt, DS, LG, KW, KIA, LD, SSG, HH.
2. Winning percentage (win%) is the total number of wins divided by the total number of games, excluding draws, of a particular team.

3. Home Team Winning Percentage is the same as Winning percentage, but only considering the subset of home games played by the team. Thus, 'home wins' divided by 'home games,' again, excluding draws.
4. Skill Difference is defined as the difference of two teams' cumulative weekly ranking (Home - Away). Thus, because rankings span from 1 to 10, differences range from 9 to -9.
5. Home-field Advantage is defined and calculated as "home winning percentage – overall winning percentage."

Methodology

The descriptive method of research was used in this study. The goal of descriptive analysis is to comprehensively summarize and apply meaning to the quality of gathered information. Thus, data was collected and organized to answer the questions concerning the status of the study.

Data Collection

For this study, the data on the 2019-2020 KBO Regular Season were collected from a total of three sources: (1) the game schedule, results, and winning percentage from the official KBO website shown in Figure 1; (2) the spectators by games were from the Professional Sports Data Portal shown in Figure 2; and (3) the weekly team rankings were from Naver Sports shown in Figure 3. For the data collection process, while the collection itself and the data cleaning were not that difficult, the biggest hurdle was combining and matching data from multiple different sources into one complete data set, as shown in Figure 4.

Figure 1. Data from Naver Sports

	단체	리그	시즌	연도	월	일	요일	홈팀	원정팀	경기장	날씨	평균기온	관중수
1	KBO	정규리그	2020 시즌	2020	10	31	토	KIA	NC	광주-기아 챔피언스 필드	구름조금	20℃	2,056
2	KBO	정규리그	2020 시즌	2020	10	30	금	두산	키움히어로즈	서울 잠실 야구장	맑음	19.6℃	7,216
3	KBO	정규리그	2020 시즌	2020	10	30	금	한화	KT	대전한밭종합운동장	맑음	19.9℃	2,883
4	KBO	정규리그	2020 시즌	2020	10	30	금	삼성	NC	대구 삼성 라이온즈 파크	맑음	18.7℃	6,547
5	KBO	정규리그	2020 시즌	2020	10	30	금	롯데	KIA	부산 사직 야구장	맑음	21℃	3,223
6	KBO	정규리그	2020 시즌	2020	10	30	금	SK	LG	인천SSG 랜더스필드	맑음	17.2℃	6,023
7	KBO	정규리그	2020 시즌	2020	10	29	목	KIA	두산	광주-기아 챔피언스 필드	맑음	19℃	1,438
8	KBO	정규리그	2020 시즌	2020	10	29	목	한화	KT	대전한밭종합운동장	구름조금	17.4℃	1,803
9	KBO	정규리그	2020 시즌	2020	10	29	목	롯데	NC	부산 사직 야구장	구름조금	21.9℃	1,515
10	KBO	정규리그	2020 시즌	2020	10	28	수	LG	한화	서울 잠실 야구장	구름조금	18.2℃	6,775
11	KBO	정규리그	2020 시즌	2020	10	28	수	롯데	NC	부산 사직 야구장	구름조금	22.9℃	1,204
12	KBO	정규리그	2020 시즌	2020	10	28	수	KIA	KT	광주-기아 챔피언스 필드	구름조금	20.9℃	650
13	KBO	정규리그	2020 시즌	2020	10	27	화	KIA	KT	광주-기아 챔피언스 필드	구름조금	21.8℃	511
14	KBO	정규리그	2020 시즌	2020	10	27	화	NC	삼성	창원NC파크	구름조금	21.6℃	5,528

	K LEAGUE	KBO	KBL	WKBL	KOVO		합계	평균
누적 합계	0	7,614,325	0	0	0	전 종목	7,614,325	5,287

Figure 2. Data from Professional Sports Data Portal

순위	팀명	경기	승	패	무	승률	게임차	최근10경기	연속	홈	방문
1	NC	144	83	55	6	0.601	0	4승2무4패	1패	45-3-24	38-3-31
2	두산	144	79	61	4	0.564	5	6승0무4패	4승	41-2-29	38-2-32
3	KT	144	81	62	1	0.566	4.5	7승0무3패	1패	44-0-28	37-1-34
4	LG	144	79	61	4	0.564	5	5승1무4패	2패	39-3-30	40-1-31
5	키움	144	80	63	1	0.559	5.5	6승0무4패	1패	41-1-30	39-0-33
6	KIA	144	73	71	0	0.507	13	4승0무6패	1승	37-0-35	36-0-36
7	롯데	144	71	72	1	0.497	14.5	3승0무7패	1승	41-0-31	30-1-41
8	삼성	144	64	75	5	0.460	19.5	6승2무2패	1승	33-2-37	31-3-38
9	SK	144	51	92	1	0.357	34.5	5승0무5패	1승	28-0-44	23-1-48
10	한화	144	46	95	3	0.326	38.5	3승1무6패	1승	25-2-45	21-1-50

Figure 3. Data from Korea Baseball Organization

월	일	홈	원정	주차 순위			점수		관객수	승패	무	시간	주차	주차
				홈	원정	차	홈	원정						
5	5	KIA	키움	1	1	0.0	2	11	0	L	FALSE	14:00	1	
5	5	kt	롯데	1	1	0.0	2	7	0	L	FALSE	14:00	1	
5	5	LG	두산	1	1	0.0	8	2	0	100	FALSE	14:00	1	
5	5	SSG	한화	1	1	0.0	0	3	0	L	FALSE	14:00	1	
5	5	삼성	NC	1	1	0.0	0	4	0	L	FALSE	14:00	1	
5	6	KIA	키움	1	1	0.0	2	3	0	L	FALSE	18:30	1	
5	6	kt	롯데	1	1	0.0	4	9	0	L	FALSE	18:30	1	
5	6	LG	두산	1	1	0.0	2	5	0	L	FALSE	18:30	1	
5	6	SSG	한화	1	1	0.0	5	2	0	100	FALSE	18:30	1	
5	6	삼성	NC	1	1	0.0	3	4	0	L	FALSE	18:30	1	
5	7	KIA	키움	1	1	0.0	8	5	0	100	FALSE	18:30	1	
5	7	kt	롯데	1	1	0.0	3	7	0	L	FALSE	18:30	1	
5	7	LG	두산	1	1	0.0	3	9	0	L	FALSE	18:30	1	
5	7	SSG	한화	1	1	0.0	4	8	0	L	FALSE	18:30	1	1
5	7	삼성	NC	1	1	0.0	2	8	0	L	FALSE	18:30	1	
5	8	NC	LG	1	1	0.0	13	5	0	100	FALSE	18:30	1	
5	8	두산	kt	1	1	0.0	3	12	0	L	FALSE	18:30	1	
5	8	롯데	SSG	1	1	0.0	9	8	0	100	FALSE	18:30	1	
5	8	삼성	KIA	1	1	0.0	5	0	0	100	FALSE	18:30	1	
5	8	키움	한화	1	1	0.0	5	3	0	100	FALSE	18:30	1	
5	9	삼성	KIA	1	1	0.0	14	2	0	100	FALSE	17:00	1	
5	9	키움	한화	1	1	0.0	5	3	0	100	FALSE	17:00	1	
5	10	NC	LG	1	1	0.0	8	10	0	L	FALSE	14:00	1	
5	10	두산	kt	1	1	0.0	13	12	0	100	FALSE	14:00	1	
5	10	롯데	SSG	1	1	0.0	4	0	0	100	FALSE	14:00	1	
5	10	삼성	KIA	1	1	0.0	3	12	0	L	FALSE	14:00	1	
5	10	키움	한화	1	1	0.0	6	3	0	100	FALSE	14:00	1	

Figure 4. Combined Final Dataset

Basic Analysis: Home Team Winning Percentage Comparison by Year

If most of the stadium spectators are fans of the home team, we compared the home team winning percentages of 2019 and 2020 and explored the relationship between spectators and the home team winning percentage using the complete dataset shown in Figure 4.
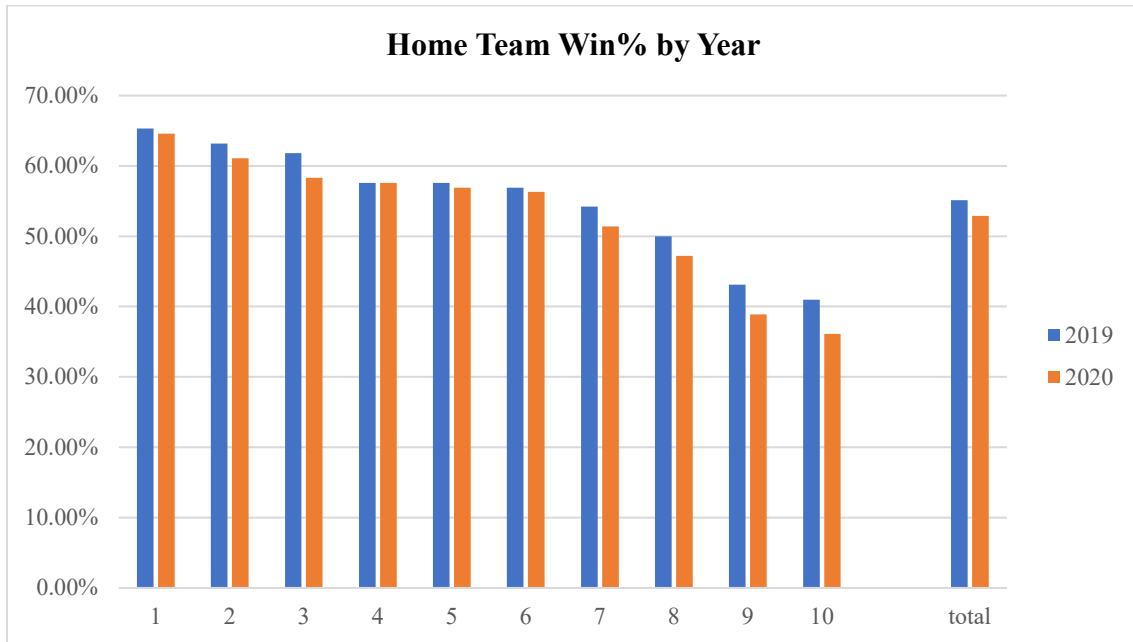

Figure 5. Home Team Winning Percentage by Year

The x-axis in Figure 5 is displayed in regular-season rankings to compare the home team winning percentage in equally ranking teams in 2019 and 2020. The general home team winning percentage by season is calculated and compared. Keeping in mind that in 2019, all games (720) had spectators present, while in 2020, 80 percent of matches (577 of 720) were closed-door, all rankings' respective winning percentages had fallen in 2020. Thus, it can be factually concluded from the graph that the 2020 season's home team's winning percentage is lower than that of 2019. In Figure 6, the home team's winning percentage showed an upward trend from 2016 to 2019 but fell significantly in 2020.
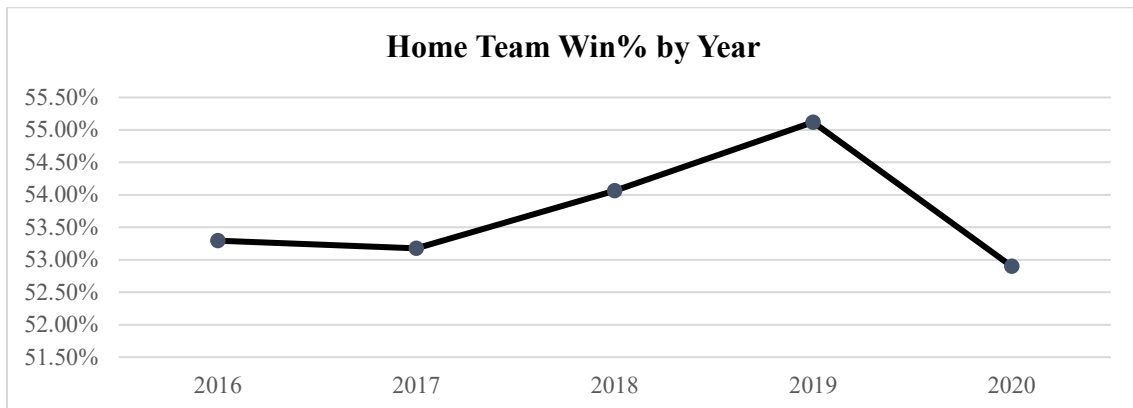
Figure 6. Home Team Winning Percentage Trend

Analysis 1: Skill Difference between Teams

The basic analysis in the previous section confirmed that the overall home winning percentage in 2020 fell but did not consider the difference in performance amongst the teams. Because there is a possibility that compared to 2019, the 2020 season had more games that pit stronger teams against weaker ones, the basic analysis on its own is insufficient. These instances are addressed through the unit of skill difference. Skill difference is defined as the difference of two teams' cumulative weekly ranking (Home - Away) and interpreted as the difficulty of a home team to win its game. Since skill difference amongst teams changes throughout the regular season, a cumulative weekly team ranking chart was referenced for each game in question, and skill difference indicators were created accordingly. Because rankings span from 1 to 10, differences range from 9 to -9, with a more excellent absolute value representing a skewed matchup and a lower number being a more favorable matchup for the home team. Finally, the home team winning percentages per skill difference were grouped, calculated, and compared.
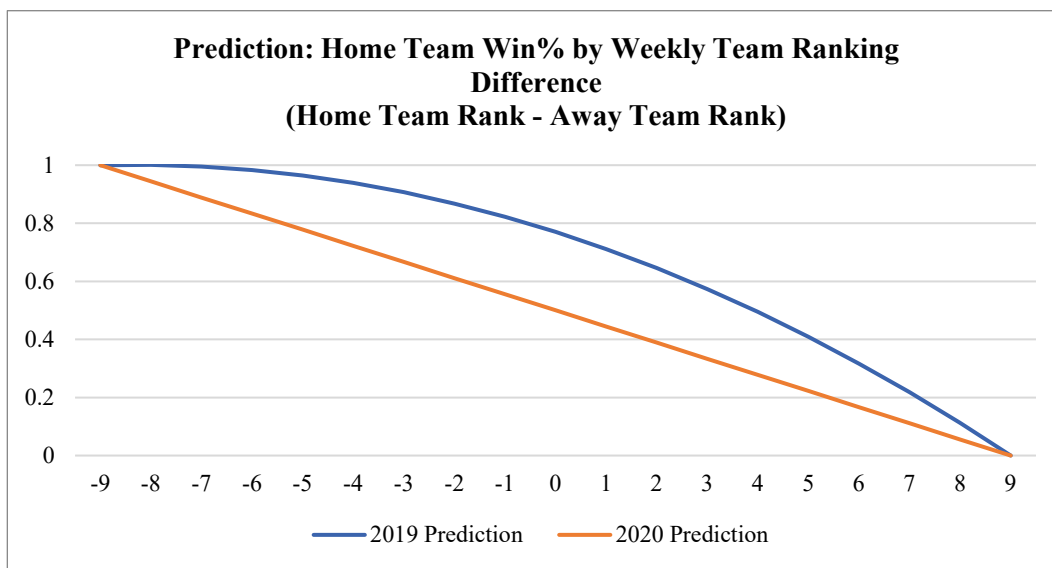

Figure 7. Prediction: Home Team Winning Percentage by Cumulative Weekly Team Ranking Difference
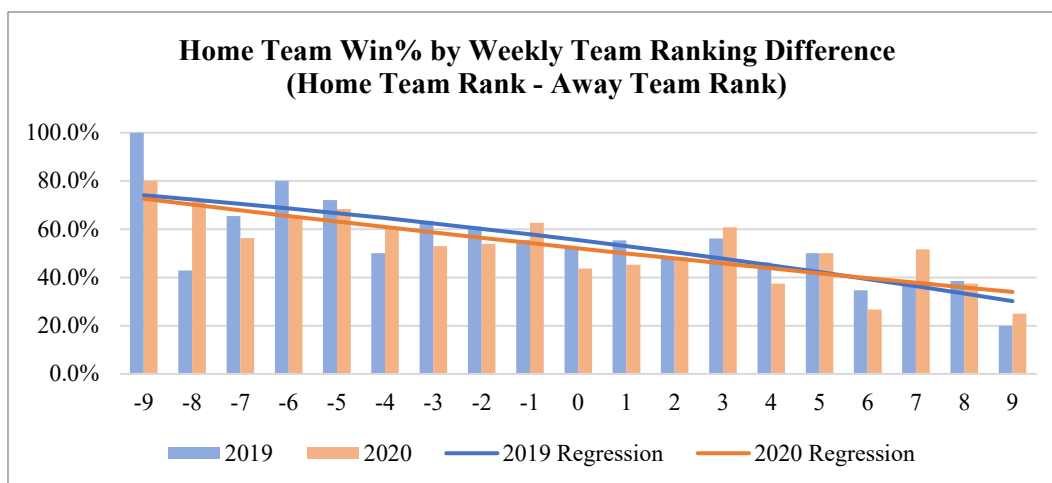

Figure 8. Home Team Winning Percentage by Cumulative Weekly Team Ranking Difference

Analysis 2: Home-field Advantage

The basic analysis comparing the home winning percentage by year showed that the home team's winning rate in 2020 fell. However, when observing ratios for each team, it was only the home team's winning percentage that decreased, or both the home and away teams' winning percentage dropped in 2020. To this end, the difference of home winning percentage and overall (home + away) winning percentage were taken to investigate the effect of home-field advantage.

Home-field advantage was the advantage of a familiar playing field for the home team and the positive effect of fervent fans cheering on the same home team. Given that the stadium itself has not changed from 2019 and 2020, it was determined that examining the home-field advantage is essentially equivalent to studying the spectator's effect. Thus, analysis 2 allows the possibility to isolate the direct impact of the crowd.
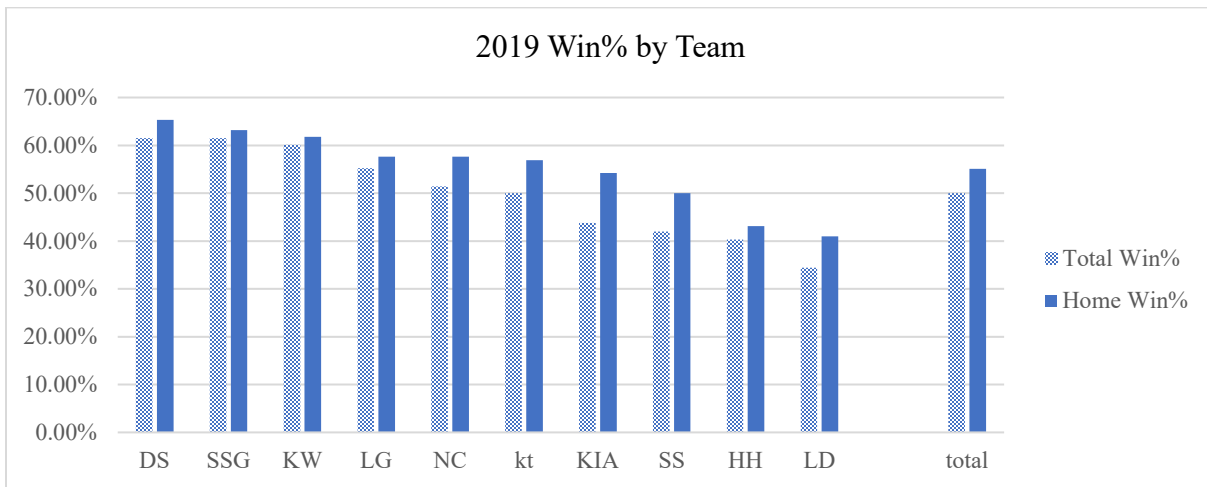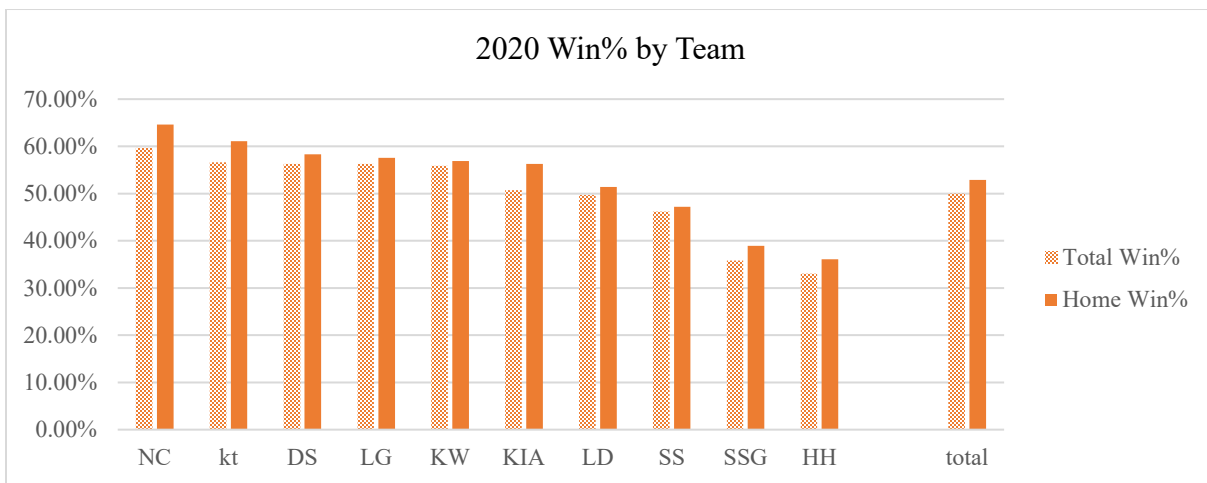

Figure 9. 2019 Winning Percentage by Team

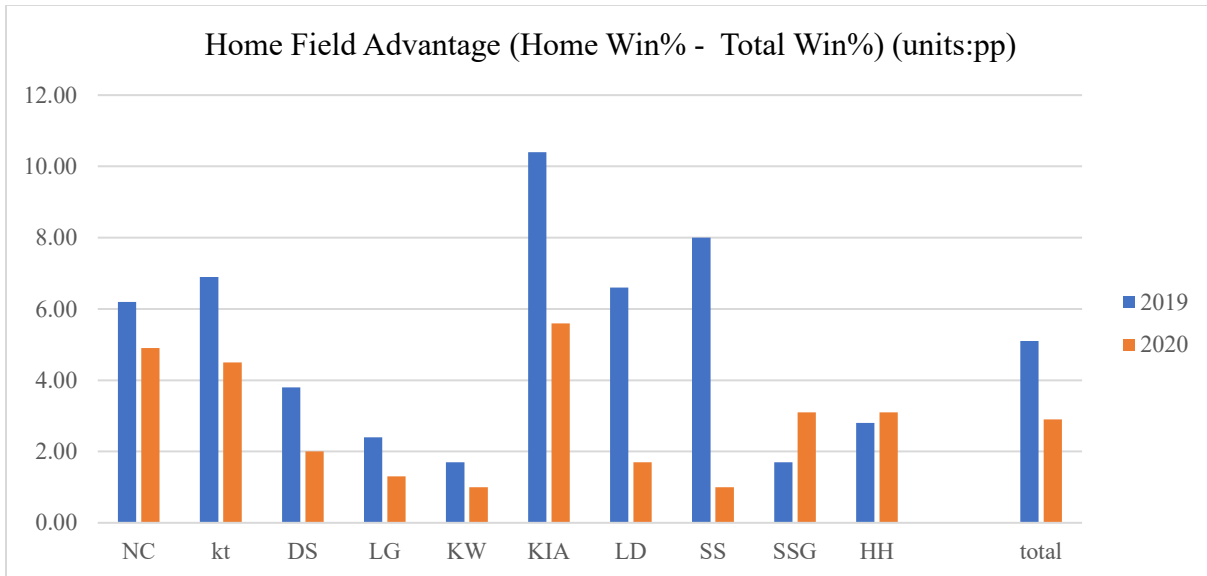**Figure 10.** 2020 Winning Percentage by Team

Figure 11. Home Field Advantage by Team

Based on the regular season ranking, it was found that every team's home winning percentage was higher than its overall winning percentage. Therefore, a team's home-field advantage was calculated using the formula "home winning percentage – overall winning percentage." In 2019, the home-field advantages of NC, kt, KIA, LD, and SS were relatively significant, while in 2020, NC, KT, KIA, SSG, and HH were the teams that showed significant home-field advantages. However, considering both years, except for SSG and HH, all teams showed a lower home-field advantage in 2020. Assuming there was no effect of the crowd on home-field advantage, the same, if not similar, the home-field advantage would be expected from both seasons. Nevertheless, as seen in Figure 11, home-field advantage was seen to be less in 2020. This difference can be attributed to the presence or absence of spectators, given the conditions set prior.
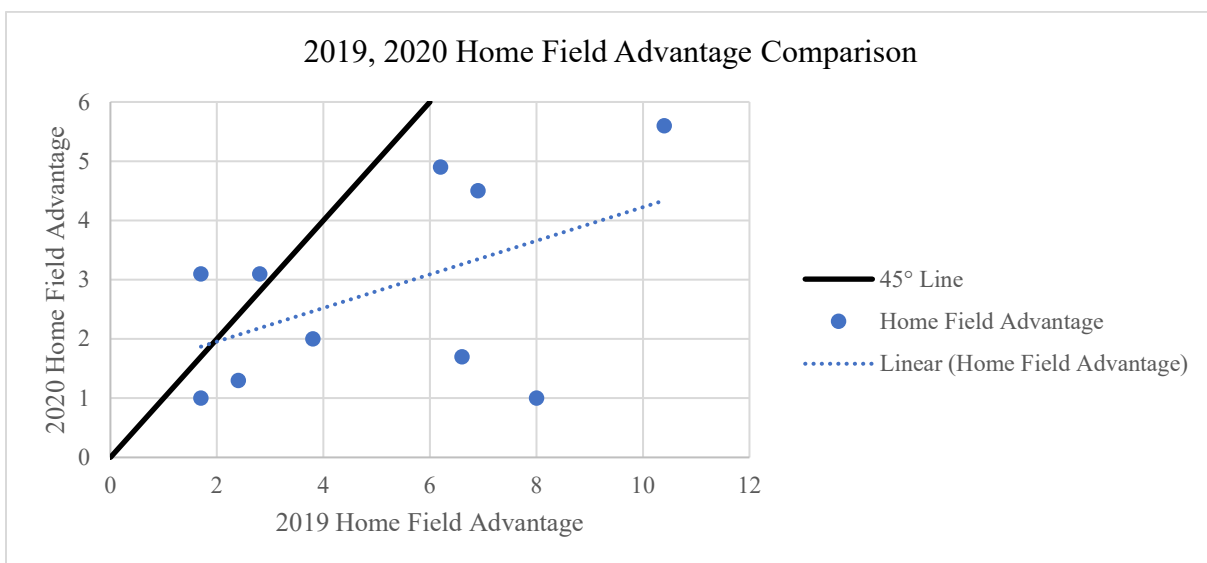

Figure 12. 2019, 2020 Home Field Advantage Comparison

Conclusion

Due to the odd year of 2020, baseball fans were presented with the unprecedented format of closed-door baseball games. Without any type of crowd affecting games, this was prime time to analyze the direct effects of the crowd given an ideally untainted, isolated dataset. This research sought to examine the immediate impact of the crowd on home winning percentage.

The basic analysis comparing the home winning percentage by year confirmed that the home winning percentage of the 2020 season was significantly lower than that of the 2019 season. There may be various factors causing the change in home winning percentage, including the sudden switch to closed-door games in 2020. Further analysis was conducted through additional research in multiple directions to determine the actual effect of spectators on the outcome of games.

Analysis 1 examined the effects of spectators considering the difference in skill amongst teams. While a significant skill difference ignored the impact of the crowd on the home winning percentage, games with little skill difference showed a lower home winning percentage in the 2020 season than in 2019. Hence, if there is no significant difference in skill amongst the teams, the crowd's cheering can positively affect the home team's winning percentage.

Analysis 2 explored and calculated the home-field advantage by season and team, notably to see if it existed for the 2020 season, and found that most teams saw their home-field advantage decrease in 2020. Even though home teams used the same stadium and playing field in both 2019 and 2020, the fact that the home-field advantage significantly reduced in 2020 can be credited to the crowd that had disappeared. Hence, the negative effect of closed-door games on the home team's games was identified.

Ultimately, by utilizing the results of KBO's 2020 closed-door season of baseball games, it was established that the home crowd cheering does help the home team win. Moreover, adjusting for the differences in opponents or home fields, the crowd still seems to have an imperative impact on the results. Go fans!

Questions for Future Studies

Now knowing the true effect fans have, what changes will the teams and leagues bring? Will players who have thanked their fans before the pandemic change their attitudes and expectations for them in the future? I hope to investigate further whether similar effects of spectators affect other sports.

Acknowledgments

References

Korea Baseball Organization, http://www.koreabaseball.com/.

Lee, Jung In. "Some Laughing, Others Quivering? Spectators' Impact on Game Performance." *Sporbiz,* 26 July 2020, http://www.sporbiz.co.kr/news/articleView.html?idxno=457650.

Naver Sports, https://sports.news.naver.com/index.nhn.

Professional Sports Data Portal, http://data.prosports.or.kr/.