The Effect of Starting Pitcher Quality on Individual Game Attendance

Introduction

The game of baseball is structured in a unique way to other sports. In most sports, the same players play every game, or at least can play every game. Pitchers in baseball do not have this luxury. Throwing is an extremely unnatural motion for the human body to perform, which creates lots of stress on the arm. This prevents starting pitchers, who throw many pitches throughout a ballgame, to pitch every day. Relief pitchers who throw less in each game may be able to pitch on any given day, but starting pitchers are typically on a schedule that will allow recovery between starts. In Major League Baseball, a team will typically have a rotation of 5 starting pitchers. With games almost every day throughout the season, each starter will pitch every 5 days, sometimes 6 if there is an off day for the team. Additionally, since a pitcher impacts every single play of the game when their respective team is playing defense, a team will perform very differently each game, depending on who the starting pitcher happens to be that game. This creates a unique fan experience each game. The scheduled starting pitchers for each game will mostly be widely available to fans, allowing fans to choose which starting pitcher they would like to see play. The incentive to attend a game featuring a better pitcher is clear, and fans have the opportunity to act on that preference and produce larger crowds for higher quality starting pitchers.

The effect analyzed in this paper does not quite equate to the superstar effect. While very similar, quality is strictly performance, whereas the superstar effect relies more on a player's public perception. They are most likely related, and a high-quality pitcher should have a greater chance to be labeled a superstar. The effect of starting pitcher quality will be assessed

in the context of the 2010 Seattle Mariners. Interestingly, this is the year that Felix Hernandez of the Seattle Mariners won the Cy Young Award for best pitcher in the American League.

Literature Review

There has been some literature on the superstar effect on attendance across sports. The superstar effect occurs when more fans attend games that have more superstars, but what defines a superstar? Is it purely based on quality? How much is public perception a factor? These may seem similar, but some performance stats are more widely publicized than others. One measure of star power of which fans take notice is awards. Russell Ormiston (2012)[1] compiles all of a pitcher's career awards and accomplishments, and gives stars for each, adding up to a total star power measure for that pitcher. These awards and accomplishments include All-Star Games, Cy Young awards, Most Valuable Player awards, no-hitters, All-Star Game MVPs, World Series MVPs, whether they won the Rookie of the Year award, and whether they have won 300 or more games. All of these titles are noticed and valued by baseball fans. One issue with this measure is that a pitcher could be a young star, with not enough years in the league to have won a significant number of awards. Ormiston posits another measure, an age-adjusted star power, that gives higher ratings to younger players.

Another measure of star power comes from Lewis and Yoon (2016)[2] uses Cy Young votes (or MVP votes for hitters) as the basis, with the current year's star power building off of a diminished version of the previous year's number. They discuss how star power affects salary

---

[1] Ormiston, Russell. "Attendance effects of star pitchers in Major League Baseball." *Journal of Sports Economics* 15, no. 4 (2014): 338-364.

[2] Lewis, Michael, and Yeujun Yoon. "An empirical examination of the development and impact of star power in Major League Baseball." *Journal of Sports Economics* 19, no. 2 (2018): 155-187.

and attendance and posit that the entertainment value of a player is the difference between their actual salary and what their salary should be based on their performance. They measure performance using an increasingly popular stat called Wins Above Replacement, or WAR. If star power were to be measured purely on performance rather than awards or salary, then WAR would be a good choice, as it is an all-encompassing stat capturing offensive and defensive performance into one number. Chris Russo and Fangraphs (2019)[3] used their own variation of WAR to estimate a pitcher's performance and compared that to the attendance of the pitcher's games. They grouped each pitcher into one of 5 groups from best to worst in terms of WAR, and then analyzed the relationship between the groups and attendance.

A further measure of star power comes from Humphreys and Johnson (2019)[4] in their analysis of the NBA. They use a player's salary relative to the rest of the league as a definition of potential star power. The league leaders in salary are identified as potential stars. From that group, actual stars are picked out as those that pass the Weekend Test, that is if they have the same effect on attendance that a weekend game has in comparison to a weekday game. Salary can be a good measure of star power, especially when considering the element of public perception. A player is worth to a team what revenue that player can bring to the organization. This revenue includes attendance, as well as merchandise. If a team believes that a player can attract more fans to games, they might be inclined to pay that player more money. This study is limited in its application to baseball as basketball allows for each player to make a bigger

---

[3] Russo, Chris. "Ballpark Attendance and Starting Pitchers." *Fangraphs Community Research*. July 2, 2019. https://community.fangraphs.com/ballpark-attendance-and-starting-pitchers/

[4] Humphreys, Brad R., and Candon Johnson. "The effect of superstars on game attendance: evidence from the NBA." *Journal of Sports Economics* 21, no. 2 (2020): 152-175.

impact on the game with only 5 players on the court at once, as opposed to 9 in baseball. This will likely cause one superstar in the NBA to have a greater effect on attendance than a superstar in MLB.

Superstars are not the only players that can draw extra fans. In many sports, and especially baseball, rookies and top prospects and bring excitement to fans. People look forward to seeing new players that have chances of becoming stars in the future. One article analyzed how top prospects in baseball affect Minor League Baseball attendance. The measure of a top prospect that Gitter and Rhoads (2011)[5] used is the standard measure: the prospects ranking relative to other prospects. Every year, *Baseball America* compiles a Top 100 Prospects list. Another article by Ormiston (2016)[6] studies the effect of rookies and number one draft picks as starting pitchers on MLB attendance. Again, the measure of rookies and number one draft picks are within their definitions.

The results of most of these aforementioned studies show expected results. Ormiston's (2012) analysis showed that for every star a pitcher had, there was a statistically significant increase in that game's attendance by .56% for the home team's pitcher and .63% for the visiting team's pitcher. The age-adjusted star power yielded similar results. In Lewis and Yoon (2016), both the home star power and the visiting star power had positive effects on attendance, but this effect had diminishing returns as the star power increased. This makes

---

[5] Gitter, Seth R., and Thomas A. Rhoads. "Top prospects and minor league baseball attendance." *Journal of Sports Economics* 12, no. 3 (2011): 341-351.

[6] Ormiston, Russell. "Hype and Hope: The Effect of Rookies and Top Prospects on MLB Attendance." *Society for American Baseball Research*. 2016. https://sabr.org/journal/article/hype-and-hope-the-effect-of-rookies-and-top-prospects-on-mlb-attendance/

intuitive economic sense, as performance is a product, and this can be seen as diminishing marginal utility from the fans' perspective. Russo (2019) showed that when measuring star power by WAR, only the top group of pitchers will have positive effects on attendance, and that even above-average pitchers will not have this effect. From 2010-2018, the best pitchers add 2.8% to the attendance. In their analysis of the NBA, Humphreys and Johnson (2019) found that attendance increased with more superstars on each team. Additionally, they found that attendance increased as the number of All-Star votes on each team increased, as this proves to be another measure of star power. For Minor League Baseball, Gitter and Rhoads (2011) concluded that only the Top 5 Prospects had an effect, and only at the AAA level. This effect turned out to be quite small, as a Top 5 Prospect only added 161 fans per game. After controlling for several outliers in the study, Ormiston (2016) showed that only rookie pitchers with a WAR of 5.0 or more in their rookie season had a positive effect on attendance. Number one draft picks had an effect on attendance with a 13% increase. As we see, the literature insinuates that a top pitcher should have a positive effect on attendance.

Empirical Model

In this paper, we are not looking at the superstar effect. Rather, we will dive into the effect of pitcher quality on attendance. This is not a variable based on awards or public recognition, but purely success on the mound. Baseball, and all sports, continually search for statistics that capture overall success and measure the value of a player. In baseball, many consider WAR to be the best statistic to date, but there are other statistics such as wOBA

(weighted on base average) that also do a good job of encompassing a player's success. WAR measures how many more wins a player produces compared to a replacement level player from the AAA minor league level. It includes defense, offense, and pitching into its equation. Statistics websites vary in their equation for WAR, but the goal of the statistic remains the same. For the purposes of this paper, we will use the Baseball-Reference version of WAR.

We must account for the many factors that affect attendance at MLB ballparks. We expect weekend games to draw larger crowds as more people have the time to attend. Another factor is time of day. An argument can be made that either day or night games might draw larger crowds. Night games can be too late for some people, but they also do not conflict with the average work shift of most people. Night games might also have less preferable weather, while day games have the chance to be in the sun. This brings us to the next variable: weather. The weather forecast for a game should have an effect on attendance, as nicer days should create higher demand for tickets. The next two variables depend on the opponent of the home team. Division rivals might draw higher crowds, but on the other hand the home team will play divisional teams more often, which could decrease demand for each individual game. Interleague games between an American League team and National League team could increase demand as these games are rarer. Home team quality and visiting team quality should both have positive effects on attendance, as fans likely prefer to see good teams play. Fans might also enjoy watching familiar players, so how long the home team's pitcher has been with the team would likely be positively correlated with attendance. The final two control variables relate to how competitive a team is within their league. How many games back from first place in their division (winning the division automatically sends a team to the playoffs) and whether

or not they are eliminated from the playoffs could affect ballpark attendance. Lastly, many teams see attendance surges for their first home games. Our variables of interest, home team starting pitcher quality and visiting team starting pitcher quality, are expected to both have positive effects on attendance. This leaves us with the following economic model:

$$Attendance = \propto_1 Weekend + \propto_2 Time + \propto_3 Weather + \propto_4 Divisional + \propto_5 Interleague$$
$$+ \propto_6 Home\ TQ + \propto_6 Visiting\ TQ + \propto_7 Years\ SP\ with\ Team$$
$$+ \propto_8 Games\ Back + \propto_9 Eliminated + \propto_{10} Home\ SP\ Quality$$
$$+ \propto_{11} Visiting\ SP\ Quality + \propto_{12} Home\ Opener$$

Data

Some of the variables used have multiple ways to represent, so we will explain how we measure each variable. A weekend game, which includes Friday, Saturday, and Sunday, will be represented with a "1", while a "0" signifies a weekday game. Time will also be a binary variable, day or night, with a daytime game holding the "1" value. Weather will be reduced to temperature, since rain is not as much of a factor with the Seattle Mariners ballpark as they have a retractable roof. This is not an indoor stadium when the roof is closed, however, so the temperature outside still affects the fans inside. Divisional games, interleague games, and the home opener also take on binary variables, with "1" representing divisional and interleague games. Home and visiting team quality will be measured using a lag win percentage, which is their respective winning percentage from the prior year. This equation is as follows:

$$TQ = \frac{[Lag\ Win\ \% * (162 - Games\ Played) + Win\ \% * Games\ Played]}{162}$$

This puts more weight on the prior year's success, with the weight slowly shifting to the current year as the season progresses and creates a variable scaled to winning percentage. The

years a starting pitcher has played for the Mariners will not include the 2010 season, but rather just the years before then. Games back will simply be the number of games back from the first place team in their division. The eliminated variable will have a "1" if the team is mathematically eliminated from the playoffs and a "0" if they mathematically still have a chance.

Home and visiting pitcher quality can be measured with career WAR. WAR, however, is a cumulative statistic, as opposed to an average. This means that a mediocre pitcher who has pitched for 10 years could have a higher WAR than an extraordinary second year player. In order to account for this, we measure the pitcher quality as WAR divided by the number of innings thrown. Unfortunately, data could not be collected on WAR for individual games, only on a yearly basis. This left us with the choice of whether to include the 2010 season or not, as a pitcher's current season would likely affect attendance in the later parts of the season, but for games at the beginning, this could distort results. With no perfect decision, the 2010 season was left out of the calculations. This decision also reflects the fact that season tickets are purchased prior to the season, so if prospective season ticket holders are interested in pitcher quality, they have only the seasons before 2010 to help them make their decision whether or not to buy season tickets.

Baseball-Reference[7] provided all of the data for the variables weekend, time, weather, divisional, interleague, years starting pitcher has played for the team, games back, and pitcher

---

[7] Baseball-Reference. "2010 Seattle Mariners Schedule." *Baseball-Reference*. https://www.baseball-reference.com/teams/SEA/2010-schedule-scores.shtml

quality. Baseball-Reference and SHRP Sports[8], used for league standings on a certain date, combined to supply data for team quality and elimination.

Results

Table 1 in the Appendix provides the results of the regression analysis on the stated empirical model. For reference, the average attendance for home Seattle Mariners games in the 2010 season was 25,491 fans. Most of the control variables had expected effects. Weekend games increased attendance by 7,457 fans per contest, a significant amount considering this makes up close to one third of an average crowd. Day games produced an extra 2,578 fans, a somewhat large difference. An increase in one degree Fahrenheit increased attendance by 200 fans, which would mean that 10 degrees warmer should cause an extra 2,000 people to come to the game. This could help explain the positive effect held by day games, and there could be some multicollinearity involved between these two variables that enhances each effect. Divisional games interestingly did not have any effect on attendance, possibly because of the aforementioned pros and cons to attending a divisional game, as they are important but also high in quantity. The novelty of interleague games, however, did draw more spectators to the ballpark, with an extra 5001 fans choosing to attend each of these games. As expected, home team quality and visiting team quality both had economically significant positive effects, but the home team's quality showed much more importance the visiting team's quality. A weighted winning percentage of .100 greater for the home team would increase attendance by a staggering 31,996 fans. The same change in the visiting team's weighted winning percentage

---

[8] SHRP Sports. "MLB Standings." *SHRP Sports*.
http://shrpsports.com/mlb/stand.php?link=Y&season=2010&divcnf=div&month=Oct&date=2

would boost ticket sales by 5,478 fans, a large amount in its own right. This means that a good visiting team with a .600 winning percentage would expect to draw almost 11,000 more fans than a poor team sitting around .400. The number of years the starting pitcher has played for the Mariners had a small, but statistically insignificant positive correlation with attendance. One variable that opposed the hypothesized effect was the number of games back from first. This showed a positive effect of an extra 1401 fans per contest for every additional game back from first. This is surprising because more games back means the Mariners are less likely to make the playoffs and each game matters less and less when it comes to the team's overall season goals of success. The variable capturing whether the Mariners had been eliminated from the playoffs or not showed the expected negative effect, with a decrease in attendance of 3,913 fans per game. There should be some multicollinearity between the games back and elimination since elimination is directly related to the number of games back.

The results regarding starting pitcher quality did not come back as expected. The home team actually has a large negative effect, while the visiting team has a lesser positive effect. We would expect both home and visiting starting pitcher quality to have positive effects, with the home pitcher having a larger effect as the home crowd likely wants to see their own team's leading players more than those of the opposing team.

Even though neither effect we found had any statistical significance, the results still pose the question: why might we not have gotten statistically significant results for the expected positive effect of the home pitcher quality? The first possible reason is that our hypothesis could simply be incorrect. While the literature shows that fans care about star power, they might not be as responsive to strict performance. The average fan does not likely

follow the WAR of players. Using anecdotal evidence as a fan, the biggest star amongst the starting pitchers was Felix Hernandez, who had the third highest starting pitcher quality. However, he drew the fourth largest average attendance out of the eight different starting pitchers in 2010. This limits this theory as a reason for the results, since Hernandez drawing large crowds with less pitcher quality would distort the results. A second reason might be that the Seattle Mariners fanbase, or at least the 2010 Seattle Mariners fanbase does not care about pitcher quality as much as other fanbases. They might just go to a game to enjoy watching their team without a desire to see a specific player. A further reason comes from the lack of success from the Mariners in 2010. Hernandez, the team's best pitcher and 2010 Cy Young Award winner, had a record of 13 wins and 12 losses in 2010[9]. Even when they had their best pitcher playing, they still only won about half of their games. While this is much better than their 2010 team winning percentage of .377, it still does not create as much incentive to go watch the best pitcher if the team loses half of the time. This posits another question however: would fans enjoy a game with a good starting pitcher more because it gives their team a greater chance to win or because they enjoy watching a player put on a good performance? In the case of the former, the team's lack of success even when good pitchers pitched would disincentivize buying tickets to a game simply because that's when a good pitcher would play. The latter scenario would seem to interest bigger baseball fans more compared to people who just enjoy watching their team. People who dive deeper into baseball should appreciate a great performance more than the casual fan since they know more about what it takes to perform at that level.

---

[9] Baseball-Reference. "Félix Hernández." *Baseball-Reference*. https://www.baseball-reference.com/players/h/hernafe02.shtml

Conclusion

The regression provided surprising results when it came to the variables of interest: home team and visiting team starting pitcher quality. Both had statistically insignificant effects, with the home pitcher having a negative effect. It is entirely possible that both have positive effects as we would expect, but these results cannot tell us that. We could make a few experimental design improvements. We are probably not including a few important variables, given that the R-Square is 0.77. A new variable with the record of the starting pitcher could be used to control for good starting pitchers with bad records and bad pitchers with good records. Additionally, finding a way to include their current season's WAR would help get more accurate results, especially in this particular season where fans may have flocked to watch their prospective Cy Young Award-winning pitcher. This could be a weighted WAR statistic similar to the equation for team quality using a lag WAR with more weight on their previous career WAR early in the season and higher weight on the current season later in the season. Lastly, including interaction variables between variables that could have collinearity might help the model explain attendance and control for multicollinearity. Unfortunately, no significant results came from this analysis, but we still see the effects of control variables and can build off of our model to find better ways to analyze the effect pitcher quality.

Appendix

Table 1:

Variable	Coefficient
Weekend?	7457.141***
Daytime?	2578.461**
Temperature	200.017**
Divisional Game?	-0.197
Interleague Game?	5002.342***
Home Team Quality	319957.671***
Visiting Team Quality	54778.017***
Years Home SP with Team	434.426
Games Back from First	1400.510**
Eliminated?	-3913.491*
Home SP Quality	-35722.665
Visiting SP Quality	9487.000
Home Opener?	23526.418***

Note: ***, **, and * label significance at the 99%, 95%, and 90% level, respectively.

References

Ormiston, Russell. "Attendance effects of star pitchers in Major League Baseball." *Journal of Sports Economics* 15, no. 4 (2014): 338-364.

Lewis, Michael, and Yeujun Yoon. "An empirical examination of the development and impact of star power in Major League Baseball." *Journal of Sports Economics* 19, no. 2 (2018): 155-187.

Russo, Chris. "Ballpark Attendance and Starting Pitchers." *Fangraphs Community Research*. July 2, 2019. https://community.fangraphs.com/ballpark-attendance-and-starting-pitchers/

Humphreys, Brad R., and Candon Johnson. "The effect of superstars on game attendance: evidence from the NBA." *Journal of Sports Economics* 21, no. 2 (2020): 152-175.

Gitter, Seth R., and Thomas A. Rhoads. "Top prospects and minor league baseball attendance." *Journal of Sports Economics* 12, no. 3 (2011): 341-351.

Ormiston, Russell. "Hype and Hope: The Effect of Rookies and Top Prospects on MLB Attendance." *Society for American Baseball Research*. 2016. https://sabr.org/journal/article/hype-and-hope-the-effect-of-rookies-and-top-prospects-on-mlb-attendance/

Baseball-Reference. "2010 Seattle Mariners Schedule." *Baseball-Reference*. https://www.baseball-reference.com/teams/SEA/2010-schedule-scores.shtml

SHRP Sports. "MLB Standings." *SHRP Sports*. http://shrpsports.com/mlb/stand.php?link=Y&season=2010&divcnf=div&month=Oct&date=2

Baseball-Reference. "Félix Hernández." *Baseball-Reference*. https://www.baseball-reference.com/players/h/hernafe02.shtml