Business 620

Allegheny College
Meadville, Pennsylvania
16335

THE EFFECTS OF THE DEFENSIVE SHIFT BAN ON MAJOR LEAGUE BASEBALL HITTERS' BATTED BALL PROFILES AND OFFENSIVE SUCCESS

Score Perkins
April 8, 2024

THE EFFECTS OF THE DEFENSIVE SHIFT BAN ON MAJOR LEAGUE BASEBALL HITTERS' BATTED BALL PROFILES AND OFFENSIVE SUCCESS

by

Score Perkins

Submitted to the Department of Business and Economics

Project Advisor: Russell Ormiston

Second Reader: Tomas Nonnenmacher

Date: April 8, 2024

I hereby recognize and pledge to fulfill my responsibilities as defined in the Honor Code and to
maintain the integrity of both myself and the College as a whole.

_____

Score Perkins

Acknowledgments

Many individuals have shaped my experience at Allegheny College. My coaches, professors, and friends have shaped me into who I am today. Meadville, Pennsylvania, has become my home for the past four years, and I will cherish the memories I have made here forever.

I extend my sincere gratitude to Professor Ormiston and Professor Nonnenmacher for their guidance and support throughout the process of completing my Senior Composition project. Their dedication to their students in the Allegheny College Business and Economics Department sets graduates up for success in the future. I felt prepared for this project as they made the process transparent.

I am grateful for having the opportunity and freedom to choose a topic that aligns with my interests and passion. I want to thank my parents for allowing me to attend a four-year college and continue playing football while completing my academic studies. Playing football at Allegheny College has helped me meet many lifelong friends and connections.

I am also thankful for the Allegheny Fishing Club and the opportunity it gave my best friends and me to live in our house together. I would not trade my time at 606 North Main Street for anything. I have created bonds that will last beyond my time at Allegheny College.

Score Perkins

Table of Contents

List of Tables

List of Figures

Abstract

The defensive shift was banned at the beginning of the 2023 Major League Baseball season to increase offense. The study aims to see how the ban on the defensive shift affects different types of hitters based on their profiles and offensive success. To test which kinds of hitters were affected most, the change in batting average on balls in play in 2023 was noted for all hitters who had at least 100 at-bats in 2022 and 2023. The statistics tested in the study are average launch angle in 2022, pull percentage in 2022, ground ball percentage in 2022, and fly ball percentage in 2022. There is room for future research as the results contradict each other. The P-value is significant and shows a slightly positive change in BABIP for hitters with higher average launch angles and batters who hit ground balls at a higher rate.

Chapter I: Introduction

1.1 Introduction

To regain baseball's status as America's Pastime, the Major League Baseball Competition Committee implemented three significant rule changes at the beginning of the 2023 season. The committee voted to increase the size of the bases in an attempt to reduce injuries and to keep the game's stars on the field as much as possible. Additionally, Major League Baseball added the pitch clock to improve the pace of play and keep more fans engaged. The third and most important rule change the Joint Competition Committee voted for was to ban the shift. The shift is a realignment of defensive players from their original or traditional fielding positions. Typically, the defensive shift is a strategy that overloads a side of the infield with three or more fielders. Often, defenses deployed three infielders on the right side of second base for left-handed hitters and three infielders on the left side of second base for right-handed hitters. Teams use this philosophy to reduce the probability of a batter reaching base by placing fielders in positions where the batter is most likely to hit the baseball based on previous tendencies. This rule change now requires teams to deploy two infielders on each side of second base with their feet in the dirt. Teams often deployed the shift against left-handed pull hitters, but there were many variations for all batters. The idea was to create a "blanket" of infielders on the pull-side of the infield. Major League Baseball voted on the shift ban to help increase the number of baserunners in a game. These changes to the rulebook were part of an attempt to create more action and less downtime. Major League Baseball is a business to maximize its profits and expand the game to new markets while attempting to put out the best product possible. Most

average fans desire more offense and less time between plays, and Major League Baseball gave fans what they asked for in an increase in offense. In 2023, the average runs scored per game increased, average game times decreased by about thirty minutes, and attendance at the ballpark increased.

In this study, different types of hitters will be tested across Major League Baseball to see which types were affected the most and which were affected the least by the shift ban. The variables in the model are pull percentage, ground ball percentage, and fly ball percentage. Based on what variables create changes in BABIP, predictions can be made about what types of hitters are most affected by the shift. After running regression models, any trends that cause a positive or negative change in the batting average on balls in play will be noted.

Although variations of the shift were born over one hundred years ago, the shift became much more common in Major League Baseball in the twenty-first century as more teams began to use more advanced analytics. Baseball analytics are not relatively new, as stats have been recorded for decades. Due to advancements in mathematics and technology, many Major League Baseball teams and organizations have begun using sabermetrics to gain advantages on the field. According to the Oxford Dictionary, sabermetrics is "the application of statistical analysis to baseball records, especially to evaluate and compare the performance of individual players" (sabermetrics, 2024). Today, the term has evolved into much more. Baseball organizations use metrics to help predict outcomes by using statistics to give teams a competitive advantage. Managers, players, fans, and front offices access this data through Statcast and other tracking devices and use it as a form of probability. By using

these metrics, teams would attempt to position fielders based on a hitter's spray chart or where he was most likely to hit the ball. With more organizations using this data for a competitive advantage, the use of the shift skyrocketed in the second half of the 2010s. In 2011, there were 209 plate appearances where the defense deployed the shift (Kelly, 2023). By 2018, variations of the defensive shift were used in over 40,000 plate appearances (Kelly, 2023). One prominent way the game of baseball drastically changed in recent years was due to the rise in the shift and eventual ban. Major League Baseball and Commissioner Rob Manfred implemented this rule as part of their effort to increase offense and scoring. To better understand how the shift, and the ban of it, affects different types of hitters, it is essential to know why teams began to utilize the shift so much in the first place.

This study will test Major League Baseball hitters based on their average launch angle, ground ball and fly ball percentages, and pull rates. Any trends in the groups that show they benefited or suffered from the shift will be noted. With this information, the individual hitters affected by the shift the most can be found, as well as if the shift was effective in the first place.

Chapter II: Literature Review

2.1 Research Introduction

The research question in this paper is to find out how the shift ban in Major League Baseball affects batted ball profile and success for different types of hitters. It is essential to understand if the ban on the shift is increasing offense, which is the primary goal of Major League Baseball and the competition committee in the Rob Manfred era. By grouping hitters into different categories based on average launch angle and different pull rates, predictions can be made for which hitters may have suffered from the shift or possibly benefited from it. This section reviews the literature on similar topics to help better understand previous research and how data analytics are valuable to Major League Baseball organizations. Because the shift was just banned one season ago, in 2023, there needs to be more published research on the topic. In the future, more academic studies will be available and published on the effects of the defensive shift in baseball.

2.2 Data and Analytics in the Major Leagues

Baseball has become one of the most data-driven sports in the world. Data has been used as an advantage for teams since the 1920s (Mizels et al., 2022). Sabermetrics began to increase in popularity among fans following Billy Beane's stint as GM of the "Moneyball" Oakland Athletics. Today, sabermetrics have evolved into data-driven sports sites such as FanGraphs, Statcast, and Baseball Savant, which all fans can access.

Sabermetrics were made to quantify player performance but are limited to the balls put into play. New technologies like Statcast have allowed the development of newer and more accurate defensive metrics and the use of probability to improve both offensive and defensive metrics. Metrics in baseball are categorized into four categories: standard or traditional, advanced, value, and Statcast (Mizels et al., 2022).

Since 2015, all 30 Major League Baseball ballparks have been equipped with Statcast technology to track the baseball and all offensive and defensive players on the field. Statcast uses a Doppler-radar technology to track the baseball (Mizels et al., 2022). Sensor data tracks players' movements on the field (Mizels et al., 2022). The data is then reported back in real time.

In the future, more analysis and understanding of how to use this data will be vital to the game. Advancements in baseball technology already allow organizations to track and gather information on the kinetic movements of players. This information can then be used to prevent injury and better understand post-injury rehabilitation's effects. Additionally, AI and machine learning can help predict when an injury may occur or when a player may be at a much higher risk of becoming reinjured. Baseball metrics have transformed the game (Mizels et al., 2022). With new wearable technologies beginning to emerge, the opportunities for new baseball metrics are endless.

2.3 Analytical Success

Belief in sports analytics seems to lead to more success. Based on data from the 2014-2017 regular and postseason, the results of all 30 teams are studied. Based on The Great Analytics Rankings in 2015 by ESPN, the teams are listed under one of

five categories: All-in, Believers, One-foot in, Skeptics, and Non-believers (Chu & Wang, 2019). The authors then count how many analytical staff each team employs. Some of their work "may include analytics, statistical data analysis, mathematical modeling, data science, data architecture, decision sciences, informatics, performance science, research, and development" (Chu & Wang, 2019, p. 206). This factor must be controlled as the correlation coefficient to address the most significant variable of teams spending more money leading to success. Then, analytics and team success were calculated as the partial coefficient. A partial F test was then used to calculate whether or not it is helpful to add sports analytics to a team with a high payroll (Chu & Wang, 2019). According to ESPN's previous study in which Chu and Wang based their research, ESPN concluded the team has the highest chance of making the playoffs if they spend at least 29.7 million dollars on pitching and are classified as All-in or Believers.

2.4 How the Game Changes

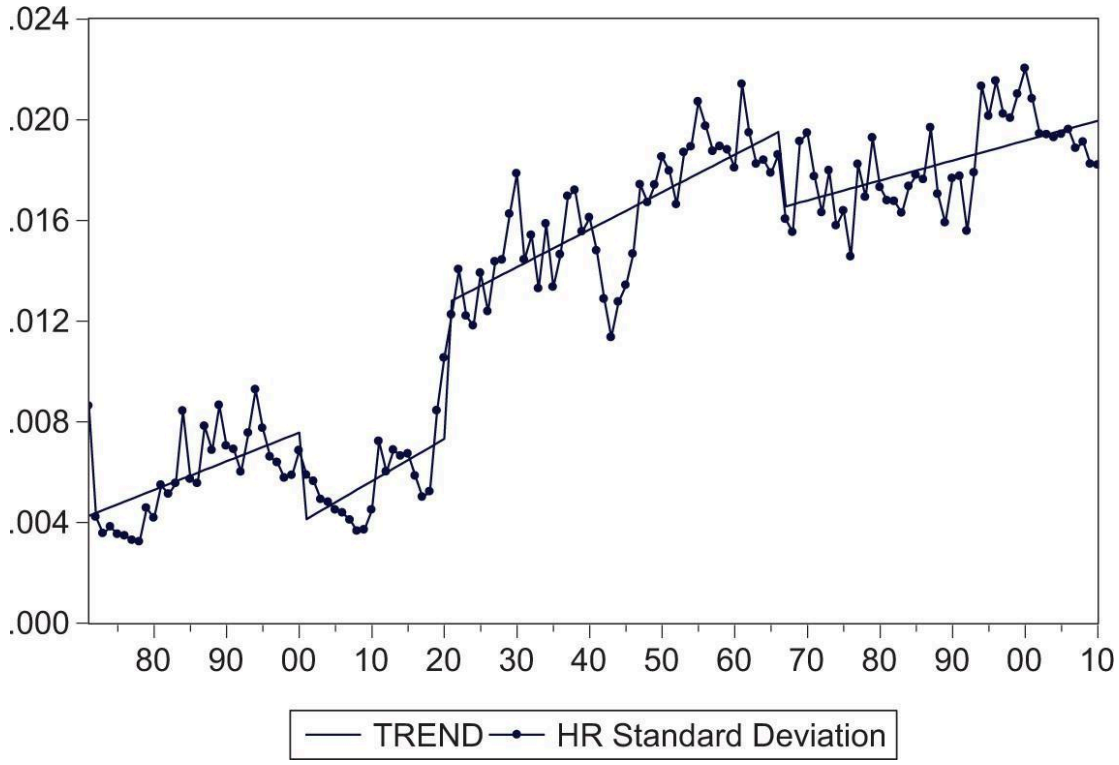
Sports leagues go through changes and different eras. Time-series tests show structural breaks in Major League Baseball performance. The mean and standard deviation can be found for each athlete's season using a database with every MLB player from 1871-2010 with at least one hundred at-bats. These statistics are used to create a time series. Time-series tests are used to measure if a player's performance is stationary (deterministic trend) or nonstationary (stochastic trend) to identify a structural break. The breaks can also be identified when plotted as a significant change in stats occurs at a certain period. These structural breaks are the points between different

eras. Major League Baseball statistics' most significant structural breaks coincide with the early years of Babe Ruth's dominance and revolutionary swing in the 1920s and the steroid era in the 1990s (Groothuis et al., 2017). The game continues to change with new philosophies, rules, and trends.

Over time, innovation changes the way games are played and approached. Historians assumed significant changes in the past based on particular events, such as Ray Chapman dying from the spitball (Groothuis et al., 2017). In contrast, Groothuis et al. used time-series tests to identify breaks and changes based on slugging, home runs, batting average, and runs batted in through the years.

The approach found an upward trend in batting, beginning in 1921 and 1992. In 1921, there was an increase in Slugging percentage, Home Run percentage, and RBI percentage. **Figure 1** shows trend lines where the breaks in HR Standard Deviation occur. Causes for the 1921 shift can be attributed to the 1919 Chicago Black Sox scandal, Babe Ruth's revolutionary swing, and the banning of the spitball (shiny white baseballs began to be used and replaced more frequently) (Groothuis et al., 2017).

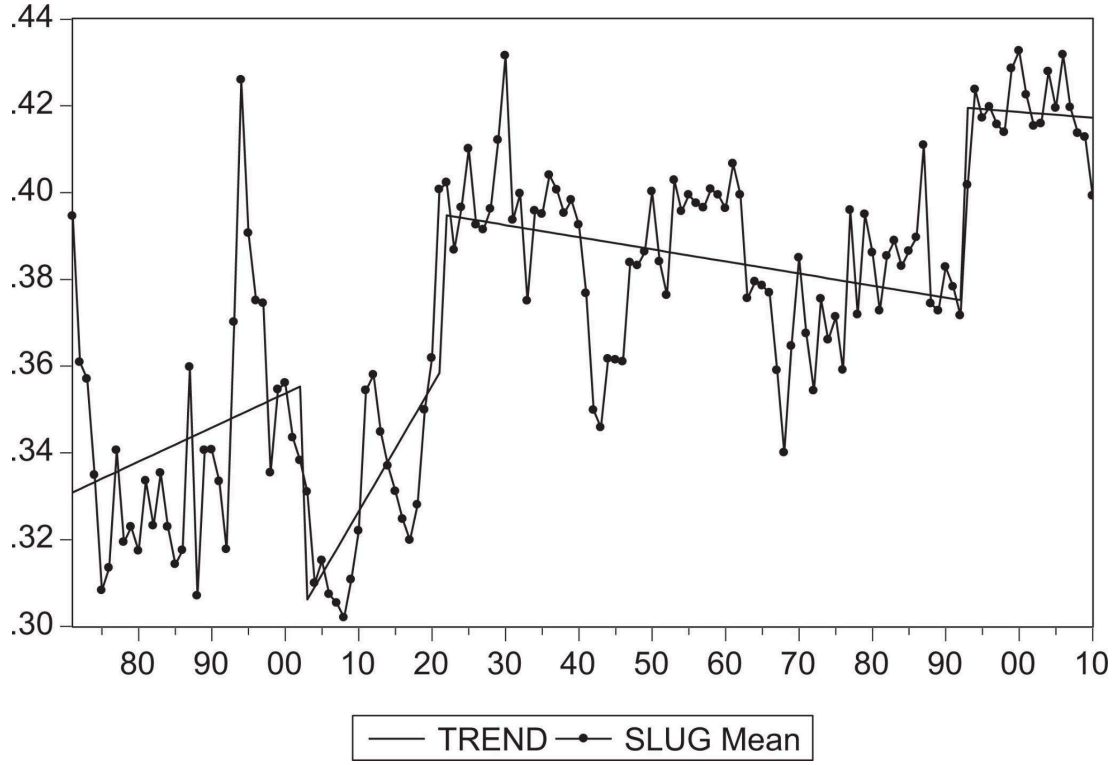
Figure 1: Breaks in Home Run Standard Deviation



(Groothuis et al., 2017)

In 1992, there was another break in mean slugging percentage across baseball. **Figure 2** below shows a significant change in the trend line. The major changes occurred in the years 1921 and 1992. Because the exact date of increased steroid use cannot be pinpointed, it is hard to define when the steroid era began. Interestingly enough, the upward shift in offensive statistics started in 1992, a year after they were banned (Groothuis, et al., 2017). Although banned, Major League Baseball did not begin testing players until 2003 (Groothuis et al., 2017). The time series test coincides with significant changes in the game. Due to rule changes in 2023, structural changes may again be present.

Figure 2: Breaks in Mean Slugging



(Groothuis, et al., 2017).

2.5 History and Approach to the Shift

The shift was first created in 1946 (Doan, 2019). Initially, it was designed to increase the chances of getting baseball great Ted Williams out over seventy years ago. Many decades later, the shift was revived due to many revolutions in statistical analysis. High usage of the shift immediately struck controversy between baseball purists on whether the shift was a good or bad thing and if it even gave defenses an advantage. Until the 2023 MLB season, when the shift was banned, hitters and pitchers had to change their patterns and approach. The shift affects left-handed hitters (LHH) much more than right-handed hitters (RHH). On average, LHH hit against the shift three times as often as RHH did (Doan, 2019). Because they faced the shift more often, this could

explain the reasoning behind having a lower batting average. RHH also had a higher batting average and slugging percentage than LHH (Doan, 2019). Two main reasons can be attributed to these numbers: RHH may be better batters than LHH, and there are more physical benefits to shifting LHH. As the infielders move closer to first base, their throws are shorter and take less time (Doan, 2019).

In the past, pitchers often pitched into the shift. This means pitchers would pitch inside to the batter, hoping they pull the baseball to the side of the infield where more fielders are positioned. This would intuitively make sense, but the problem with this approach is that even with the shift in effect, batters tend to hit the inside and middle pitches much better than outside pitches. LHH has a batting average of .246 on inside and middle pitches and an average of .197 against outside pitches (Doan, 2019). Likewise, RHH hit .264 on inside and middle pitch locations while only batting .242 on the outside pitch (Doan, 2019). Even though most of the opposite side of the infield is empty, it was still better for pitchers to use the outside part of the plate rather than to rely on the shift. Another surprising number is lefties pull the outside pitch more often than the middle and inside pitch. For middle and inside pitches, LHH hit 35.6% of balls to right field, 27.35% to left field, and 37.05% to center field (Doan, 2019). Lefties pulled the ball at over a 50% clip for outside pitches, taking only 17.67% of outside pitches the other way (Doan, 2019). Consistent with where most batted balls are put into play, lefties record extra-base hits over three times more often on pitches in the middle or in comparison to outside pitches. Batters hit for more power by pulling all pitches, including outside pitches, but they have the most success on middle and in pitches because they hit the ball where it is pitched. These numbers go against the age-old

philosophy of taking outside pitches the other way from gap to gap and emphasizing pulling the baseball and hitting for more power. Instead, hitters should pull all pitches for the most success, and pitchers should attack the outside part of the plate for the most success.

Similarly to lefties, RHH pull the ball twice as often as they go to the opposite field (Doan, 2019). The batting average based on pitch location was almost identical for right-handed hitters compared to left-handed hitters. Hitters perform better on inside and middle pitches, where pitchers tend to pitch the baseball. Inversely to what most would initially assume, hitters pull the ball more on outside pitches and take the ball to all fields on middle and inside pitches. RHH and LHH generate the most power by pulling pitches in any location (Doan, 2019).

Many different circumstances can explain the counterintuitive approach to beating the shift. As average pitch velocity increases, more hitters get jammed and hit the ball to the opposite field (Doan, 2019). Hitters may also be more comfortable controlling the inside pitch, as their overall stats are better than when they hit the outside pitch (Doan, 2019). The more significant variance in stats on the outside pitch is likely due to hitters having more trouble handling outside pitches.

Most hitters, however, did not change their approach to hit around the shift. Instead, the transition was to hit the baseball over the shift. Hitters have tried to increase their average launch angle, which has led to the flyball revolution. A pitcher's approach to the shift is more straightforward because pitchers typically pitch toward their defense (Doan, 2019). This was done to get a hitter to hit the ball to the three-person side of the infield. The data shows that this approach is backward, and

pitchers should attack the outside part of the plate, away from the shift, until hitters adjust to make the pitcher pay (Doan, 2019).

2.6 Mathematical Approaches to Defensive Alignment

In the 2018 Major League Baseball season, hitters faced a form of defensive shift in 34% of their plate appearances (Bouzarth et al., 2021). The idea is that putting more fielders where the batter tends to hit the ball should reduce the batting average on balls put in play (BABIP). A study from the 2018 MLB season shows that batters hit .276 on balls put in play against the shift while they hit .306 on balls put in play when facing traditional defensive positioning (Bouzarth et al., 2021).

In the 2018 MLB season, batters were shifted against 34% of the time. Typically, the shift is used to reposition infielders to the pull-side, where the batter is most likely to hit the baseball. To best know where to position defensive players, 5,021 possible disks, or starting areas, were created to cover the surface area of the baseball field. These disks are circles with a five-foot radius created by the researchers as areas where fielders can be placed. This allows the seven defenders (pitcher and catcher excluded) to be placed anywhere on the field instead of their traditional alignments. Once positioned, each player is given a coverage zone based on the average range of an infielder or outfielder. This approach allows defensive players to be aligned based on each batter's high-percentage hit locations. Although it varies from hitter to hitter, the average hit position for both right and left-handed hitters was charted for 10,000 balls put into play from the 2019 MLB season. A standard or traditional fielding coverage was created by averaging ten random LHH and RHH weight vectors. These results

determined that a standard defensive positioning is optimal for the average hitter using four infielders. The author created simulations to compare the change in BABIP of a hitter when he faces a three-person vs. a four-person infield. Their simulations showed that “optimal positioning with three infielders lowered predicted BABIP by 5.9% for right-handers and by 10.3% for left-handers on average when compared to a standard four-infielder placement of players” (Bouzarth et al., 2021, p. 7).

Using a mathematical programming approach, the author concluded that defensive players should be positioned with a balance where hitters most often hit the ball while still covering up most of the field. Repositioning against some of the game's biggest stars will decrease their batting average on balls in play, increasing the chances of getting out (Bouzarth et al., 2021). The table below shows the percent change in BABIP for various hitters if the defense was repositioned to areas where the hitter most often hits the ball instead of a more traditional alignment.

Table 1: Percentage Change in BABIP After Repositioning.

Player (RH)	BABIP Change	Player (LH)	BABIP Change
Altuve	-7.14%	Bruce	-13.68%
Arenado	-6.36%	Carpenter	-9.91%
Cain	+1.92%	Freeman	-11.82%
Encarnacion	-8.93%	Gallo	-5.13%
Frazier	-6.25%	Gordon	-9.73%
Longoria	-8.93%	Harper	-8.77%
Martinez	-3.85%	Heyward	-11.72%
Posey	-5.56%	Kiermaier	-9.91%
Pujols	-9.91%	Rizzo	-13.51%
Trout	-4.46%	Votto	-9.01%

(Bouzarth et al., 2021)

As **Table 1** shows, repositioning fielders according to the mathematical approach created by Bouzarth et al. would decrease the BABIP for nineteen out of twenty tested batters. This is essential information for clubs to learn more about the best way to keep runners off the basepaths. Major League Baseball banned the shift in its effort to increase offensive numbers to prevent the defense from having an advantage.

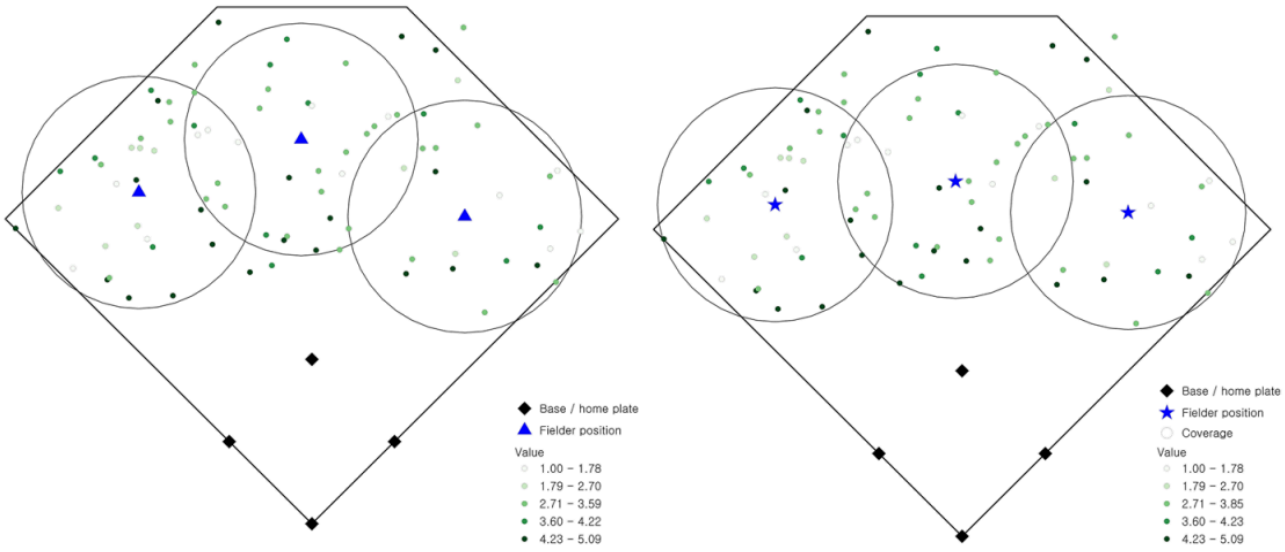
2.7 Shifting Outfielders

Location analytics are used to study where the ideal positioning of outfielders is. Doing so enhances strategic positioning and player performance. Enhancing strategic positioning by shortening the average distance to a batted ball and maximizing the importance of batted ball coverage may lead to outfield performance improvements of 1-3% (Murray et al., 2022). The idea is based on the widespread implementation of infield shifts to record extra outs throughout a season.

Baseball has entered the era of analytics through technology advancements and inventions such as the TrackMan, Statcast, and PITCHf/x (Murray et al., 2022). With more data, organizations demanded more studies on defensive positioning. Major League Baseball has released data on Baseball Savant, which gives the public ball-tracking ability. By using spatial analytics, managers can strategically position fielders to maximize the overall coverage of the outfield instead of playing their typical straight-up and symmetrical positions. The optimal positioning for fielders can be characterized by decision variables in a mathematical function. For this function, observed batted balls during NCAA Baseball games at the University of California at Santa Barbara during the 2018-2019 season were obtained from the TrackMan radar

system. Optimal positioning typically requires playing the center fielder at the same depth as the corner outfielders (Murray et al., 2022). Based on the batted ball profile used in the study, **Figure 3** shows how an outfield should be positioned to maximize coverage. By using this outfield maximization strategy, 89.04% of the total outfield can be covered by three outfielders.

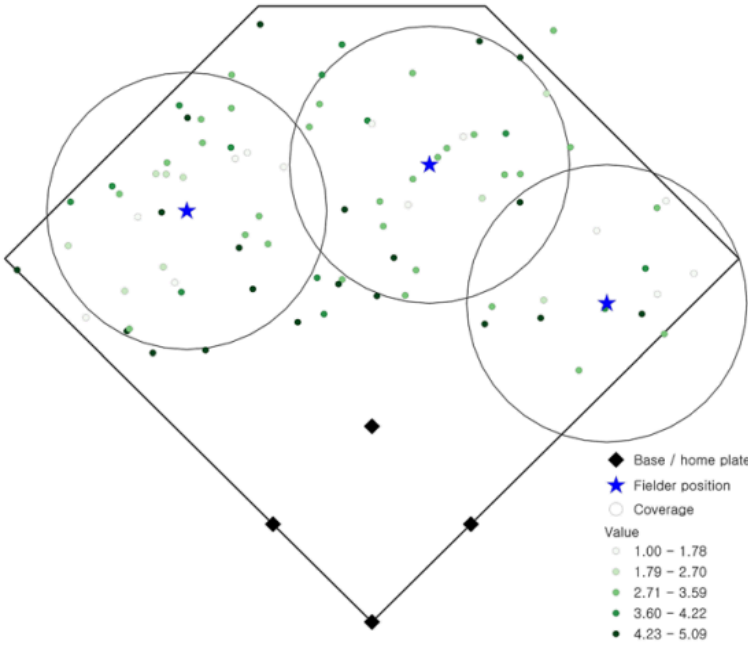
Figure 3: Traditional Outfield Alignment vs. Coverage Maximization Alignment



(Murray et al., 2022)

Another way fielders can be positioned is by distance maximization. Again, this maximization is based on the batted balls from the 2018-2019 baseball season at the University of California at Santa Barbara. By repositioning the outfielders based on these batted balls, teams can lower the average distance to a batted ball by 0.36-2.23 feet. As seen in **Figure 4**, this strategic positioning moves the center fielder into the gap more while the left fielder plays into left-center field and the right fielder starts closer to the foul line.

Figure 4: Outfield Positioning Based on Distance Maximization

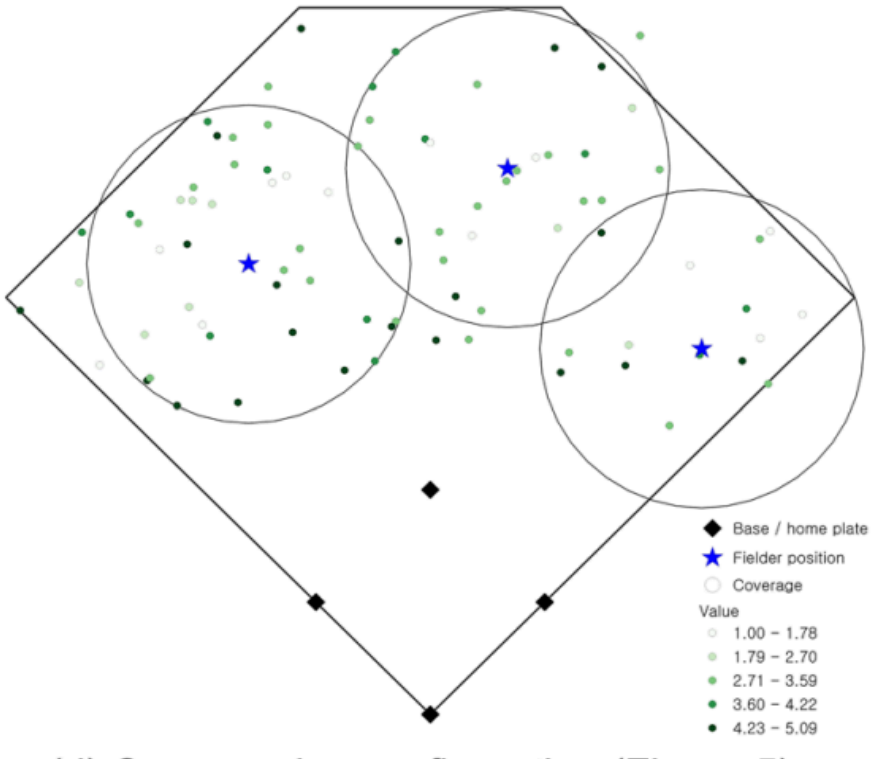


(Murray et al., 2022)

Moving the positioning of outfielders based on distance maximization may take away from the overall coverage of the defense (Murray et al., 2022). To compromise between the two, a tradeoff strategic repositioning was created. The center fielder can again be shaded toward the gap to maximize overall outfield coverage. To minimize the distance to the baseball, the center fielder can deepen up in the gap. The left and right fielders take the same position on the field as in the distance maximization scenario. The compromised configuration of defensive alignment can be seen in **Figure 5**. The ideal tradeoff decreases the average distance to a batted ball by 1.79 feet and increases average outfield coverage by 1.81%, which can be used as a defensive advantage (Murray et al., 2022). This study opens up more ideas as to whether using a fourth outfielder is useful. In the best cases, using four outfielders can increase coverage by ten percent and improve the average distance by over twelve feet.

Although further research needs to be done, a case could be made that there are some situations where using four outfielders would give the defense an advantage.

Figure 5: Compromised Configuration of Distance and Coverage Maximization



(Murray et al., 2022).

2.8 Conclusion

Analytics have taken over baseball. Teams use mathematical approaches to understand the most possible outcomes, which significantly influences their decision-making. Organizations are devoting more money, time, and resources to their analytical departments, believing that this can help save and generate money and lead to more success.

Previous research has been done on defensive positioning. Results show that nontraditional defensive positioning may be beneficial for ball clubs. As there are no

rules about outfield alignment, more data on distance minimization and maximum coverage is essential for further research that could still be relevant to Major League Baseball teams. This study aims to understand the effects of hitting against the shift vs. hitting against a traditional defense.

Other studies have also used the BABIP statistic to eliminate as many variables as possible. This research can help us understand which types of hitters may have had different career outcomes based on how much they faced the shift.

Chapter III: Theory

3.1 Theory Introduction

This chapter relates a business or economic theory to the research project. This is the reasoning behind how the project relates to economics. Probability theory can predict outcomes based on hitters' pull percentage at a superficial level. Creating a mathematical formula can predict the batting average on balls in play. It also will determine whether hitters should attempt to pull the baseball more.

3.2 Equation

The following equation will be used to predict a hitter's batting average on balls in play:

BABIP = Ppull x (1-Dpull) + (1-Ppull) x (1-Doppo)

The batting average on balls in play for a hitter can be predicted using this formula. By dividing the field in half and classifying all balls in play as either a pulled baseball or an opposite-field baseball, the variables *Ppull* and *Doppo* should add up to one. *Doppo* can be found by using (1-*Dpull*) as they are inverses. By plugging in hitters' pull and opposite field percentages and comparing them between the 2022 and 2023 seasons, it can be noted which types of spray charts for hitters benefitted the most from the ban of the shift. In theory, hitters with a higher *Ppull* pull the baseball more and should benefit the most from the ban of the shift. This raises the question of whether pull percentage is significant in predicting a player’s batting average on balls in play. This is a simplistic equation compared to the regression used later in the study.

3.3 Economic Theory

Probability is assigning a number between zero and one to an event to analyze the chance it will occur. This can also be referred to as the relative frequency at which something may happen. The Bayes Theorem can be used in this scenario.

When using the equation in Bayes Theorem, whether or not a batter faces the shift can be used as the binary variable. The independent events will be whether the batter reaches base or gets out. The results are the probability of a batter getting a hit when facing the shift vs. hitting against a traditional defensive alignment.

3.4 Explanation

The variable *Ppull* is the probability that the hitter pulls the ball. *Dpull* is the probability that the ball is fielded if the batter pulls the baseball. *Doppo* is the probability that the ball is fielded if the batter hits the baseball to the opposite field. Each side of the equation is inversely related to each other. When a batter is shifted, the value for *Dpull* should increase, and *Doppo* should decrease. If there is no shift, *Dpull* should decrease, and the value for *Doppo* should increase.

3.5 Results

Based on the theory, hitters should succeed more when they hit the baseball to the opposite field when they are shifted against because there is less chance that the ball gets fielded. When there is no shift, hitters should ideally pull the baseball because they typically have better success doing so. The BABIP in the equation is based on the chance that the hitter is early and pulls the ball or is late and takes it the other way to

the opposite field.

3.6 Application

The results do not tell too much information as the mathematics is broken down into a straightforward form. Although the equation does not give the most realistic results, the numbers can be used as talking points. These numbers are based on success rates when pulling the ball vs. going to the opposite field when facing or not facing the shift. Results can influence the general approach of many hitters or even challenge some common hitting philosophies.

3.7 Conclusion

Based on Baye's Theorem, I expect batters to pull the ball more when shifted. When shifted, *Dpull* is high, and *Doppo* is low. In contrast, when there is no shift, *Dpull* is low, and *Doppo* is high. The equation is created to find the probability that the runner reaches base on a ball in play. Results should show why teams began to shift so much in the first place by showing a lower expected BABIP for hitters when they are facing the shift. The shift is expected to significantly affect hitters who pull the ball at a higher rate than those who take pitches to the opposite field. The regression model tests the pull rate and opposite field rate variables for statistical significance.

Chapter IV: Empirical Analysis

4.1 Introduction

Intuition would suggest that most hitters would benefit from rules that hinder where the defense may position itself. By banning the defensive shift, Major League Baseball has attempted to increase the number of baserunners. Previously, teams would position their defenders to take away hits through the 5-6 hole, the 3-4 hole, and back up the middle. By being forced to have a balanced infield, teams cannot overload aside to take more base hits away.

In order for a study to be created, data had to be built. The data was collected from Statcast. A player's eligibility threshold in the study was set at one hundred bats for the 2022 and 2023 seasons. The following statistical categories were recorded for each hitter in 2022 and 2023: batting average, expected batting average, average launch angle, pull percentage, straightaway percentage, opposite field percentage, ground ball percentage, line drive percentage, and fly ball percentage. The year 2022 was used because it was the last year the defensive shift was legal at the Major League level. The year 2023 was used to compare numbers, as 2023 was the first and only season with the defensive shift ban. Trends between seasons were noted, and stats were determined to be significant or insignificant based on the p-value and the r-squared value.

4.2 Data Summary

Various statistics were measured for all Major League Baseball hitters with at least one hundred at-bats in 2022 and 2023 to have a large enough sample size. This

created a list of three hundred fifty-three eligible hitters. The Statcast data was accessed through Baseball Savant. According to Baseball Savant, the “search page allows you to query MLB’s Statcast database on a per-pitch, per-game, per-player, per-team, and per-season basis” (Statcast, 2024).

Once the data was collected, it was inputted into an Excel spreadsheet. The spreadsheet included the player’s name, ID, year, change in batting average, batting average for 2022 and 2023, line drive percentage for 2022 and 2023, ground ball percentage for 2022 and 2023, flyball percentage for 2022 and 2023, average launch angle for 2022 and 2023, pulled percentage, opposite field percentage, and straightaway percentage for balls in play for 2022 and 2023.

Multiple regressions were run on Excel with this model. The change in BABIP between 2022 and 2023 as the y variable shows a change in hitters' success before and after the shift. The X variables used in the study were average launch angle, pull percentage, flyball percentage, and ground ball percentage. The data found that hitters with higher average launch angles and hitters who hit ground balls at a higher rate benefitted the most from the shift ban in Major League Baseball.

4.3 Regression Results

Table 2: Results

Variable	Coefficient	P-Value	Minimum	Maximum
Average Launch Angle in 2022	**0.005**	**0.00006**	**0.0025**	**0.0072**
Pull Percentage in 2022	**-0.0002**	**0.59**	**-0.00091**	**0.00052**
Flyball Percentage in 2022	**0.0009**	**0.14**	**-0.00028**	**0.002**
Ground Ball Percentage 2022	**0.0035**	**0.00004**	**0.0018**	**0.0052**

4.4 Conclusion

The model shows statistical significance for average launch angle and ground ball percentage. In essence, hitters with a higher average launch angle and high ground ball percentages benefitted the most from Major League Baseball's shift ban. As the average launch angle and ground ball percentage increase, BABIP also slightly increases. These seem to contradict each other but may also apply most to a niche group of hitters who hit a lot of ground balls but still have a high average launch angle. Additionally, the R-square value is 0.06 which is low. This explains that the variables do not have a major impact on expected BABIP as hitting a baseball is mostly luck along with many other factors.

Chapter V: Conclusion

5.1 Results

Using a comprehensive data set of relevant statistics for Major League Baseball hitters with a sufficient sample size, regression analysis was made in Excel to help better understand the types of hitters most affected by the defensive shift ban. The model concludes that hitters with higher average launch angles and higher ground ball percentages benefitted the most from the shift ban. This goes against intuition, as hitting more ground balls would likely lead to a lower average launch angle. Because of this, future research is essential to understand the actual effects of the defensive shift ban. More research can be conducted with larger sample sizes as more seasons are completed under the new rules.

5.2 Limitations

Using actual baseball statistics, the test group had to remain small because of the limited number of hitters with at least one hundred at-bats at the Major League level in 2022 and 2023. The ban on the shift is still new, as the rules have only been in place for one season. Teams and players are still learning to navigate the changes in the new rule. As more data and information emerge, the philosophy and approach of players will continue to change.

Another considerable limitation of the study was the lack of qualified hitters. Only three hundred fifty-three Major League Hitters qualified for this study's one-hundred plate appearance threshold. In an ideal research study, more test subjects would be available for use.

The final limitation of the study is a data issue. With data from 2022, there is no way to differentiate between plate appearances where a hitter may or may not have faced the shift. Because of the limitations, at-bats were compared between seasons based on rules instead of whether a hitter faced the shift or a more conventional defensive alignment in their plate appearance.

One potential issue is that the shift ban also affects fielders. Without the shift, infielders are required to have more range. The shift allowed teams to substitute defense for more offense, as some players seemed to be playing out of position. Because teams may want to prioritize defense in the infield more now that the shift is banned, they may trade better defense over a better bat. Players with a stronger glove may see more at-bats throughout the season. This could slightly affect offense inversely, leading to less of an increase based on hitting statistics across baseball.

5.3 Final Thoughts

Major League Baseball implemented significant rule changes beginning in 2023. This was implemented by the league's Competition Committee to expand the game to more viewers and to create more offensive production. Although the shift ban did not make significant changes to baseball, other rule changes that were implemented simultaneously have changed the game of baseball. The goal of creating more runs and shortening game times has benefited Major League Baseball due to other rule changes such as bigger bases, pickoff attempt limits, and the pitch clock. In the coming years, baseball fans may still see more of a change in batting average and runs scored

as a result of the shift ban. For now, the new defensive rules do not significantly affect the success of different hitters.

References

Bouzarth, E., Grannan, B., Harris, J., Hartley, A., Hutson, K. & Morton, E. (2021). Swing shift: a mathematical approach to defensive positioning in baseball. Journal of Quantitative Analysis in Sports, 17(1), 47-55. https://doi.org/10.1515/jqas-2020-0027.

Chu, D., and Wand, C. (2019). Empirical Study On Relationship Between Sports Analytics and Success in Regular Season and Postseason in Major League Baseball. Journal of Sports Analytics, 5, 205-222. https://content.iospress.com/download/journal-of-sports-analytics/jsa190269?id=journal-of-sports-analytics%2Fjsa190269

Doan, C. (2019). Shifting Expectations: An In-Depth Overview of Players' Approaches To the Shift Based on Batted-Ball Events. Baseball Research Journal https://sabr.org/journal/article/shifting-expectations-an-in-depth-overview-of-players-approaches-to-the-shift-based-on-batted-ball-events/

Groothuis, P. A., Rotthoff, K. W., & Strazicich, M. C. (2017). Structural Breaks in the Game: The Case of Major League Baseball. Journal of Sports Economics, 18(6), 622-637. https://doi.org/10.1177/1527002515593113.

Kelly, M. (2023, February 16). *What is a defensive shift?*. MLB.com. https://www.mlb.com/news/infield-shift-a-definitive-history.

Mizels, J., Erickson, B., & Chalmers, P. (2022). Current State of Data and Analytics Research in Baseball. Current reviews in musculoskeletal medicine, 15(4), 283–290. https://doi.org/10.1007/s12178-022-09763-6.

Murray, A. T. & Ortiz, A & Cho, S (2022). Enhancing Strategic Defensive Positioning and Performance in the Outfield. Journal of Geographical Systems, Springer, vol. 24(2), pages 223-240, April. https://doi.org/10.1007/s10109-021-00367-1.

Sabermetrics. (2024). OED.com. Retrieved April 8, 2024, from https://www.oed.com/dictionary/sabermetrics_n?tab=meaning_and_use&tl=true#113034852.

Statcast Search. baseballsavant.com. (2024). https://baseballsavant.mlb.com/statcast_search.

About the Author

I am currently a senior at Allegheny College. I majored in Business and minored in Geology. I played football all four years and was a part of the Allegheny Fishing Club. After graduating from Allegheny, I plan to attend graduate school to pursue my MBA.