Does the Defensive Shift Employed by an Opposing team affect an MLB team's Batted Ball Quality and Offensive Performance?

11/20/2019

Abstract

This project studies proportions of batted ball quality across the 2019 MLB season when facing two different types of defensive alignment. It also attempts to answer if run production is affected by shifts. Batted ball quality is split into six groups (barrel, solid contact, flare, poor (topped), poor (under), and poor(weak)) while defensive alignments are split into two (no shift and shift). Relative statistics come from all balls put in play excluding sacrifice bunts in the 2019 MLB season. The study shows there to be differences in the proportions of batted ball quality relative to defensive alignment. Specifically, the proportion of barrels (balls barreled) against the shift was greater than the proportion of barrels against no shift. Barrels also proved to result in the highest babip (batting average on balls in play) + slg (slugging percentage), where babip + slg then proved to be a good predictor of overall offensive performance measured in woba (weighted on-base average). There appeared to be a strong positive correlation between babip + slg and woba. MLB teams may consider this data when deciding which defensive alignment to play over the course of a game. However, they will most likely want to extend this research by evaluating each player on a case by case basis.

Background and Signifigance

Do MLB teams hit the ball better when facing a certain type of defensive alignment? As the shift becomes increasingly employed in Major League Baseball these types of questions become more and more important. In this study, we investigate the relationship between defensive alignment and the batted ball quality against that alignment. Baseball players, coaches, and front office staff would be interested in this research as it could help answer the broader question that is "How is offensive performance affected by defensive shifts?"

Studies that examine offensive performance relative to defensive alignment have yielded mixed results. Some suggest that despite concern about the shift limiting offense, hitters are becoming increasingly better against shifts. Sportswriter Travis Sawchik (2019) argues that despite the increase in shift frequency, batting average on balls in play (a widely used batting metric that attempts to measure offensive performance) has remained relatively constant. Contrarily, baseball writers Rob Arthur and Ben Lindbergh found that players who see the highest frequency of shifts also tend to experience the largest decline in offensive performance as they age.

In this paper, we test the null hypothesis that states there is no difference in batted ball quality when a team faces a certain defensive alignment against the alternative hypothesis that states there is a difference in batted ball quality when a team faces a certain defensive alignment. If we can reject this null hypothesis, we will then test a different set of hypotheses specifically related to the proportion of barreled balls between each defensive alignment. The null hypothesis will state the proportion of barreled balls is the same regardless of defensive alignment. The alternative hypothesis is the proportion of barreled balls is greater when a team faces a shift.

Methods

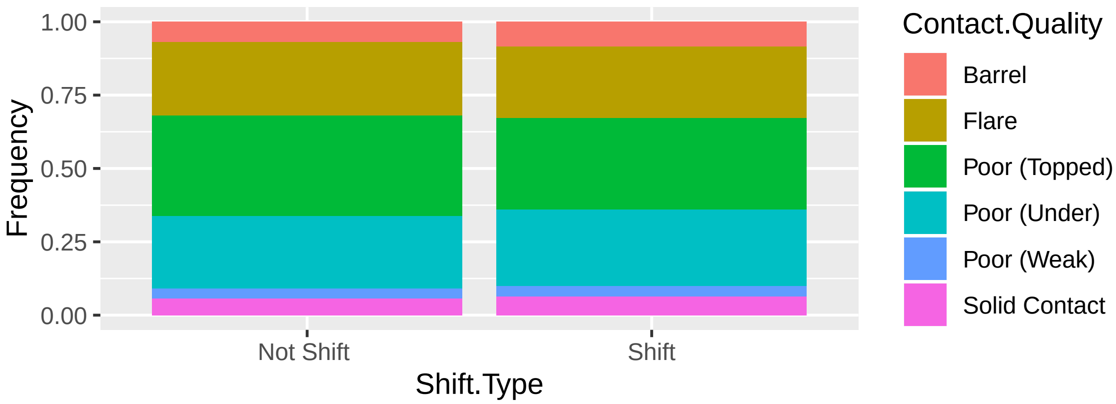
Data was gathered using statcast search from Baseball Savant (a baseball statistics database). Data collected include batted ball quality proportions, babip (batting average on balls in play), slg (slugging percentage), and woba (weighted on-base average) for all balls in play excluding sacrifice bunts in the 2019 MLB season when facing two different types of defensive alignments (no shift and shift). Defensive fielding shifts in major league baseball can be categorized as any player realignment from the standard baseball positions. Batted Ball Quality is measured in six different categories "barrel", "solid contact", "flare", "poor(under)","poor (topped)", and poor "(weak)"). The batted ball quality proportions and the two types of defensive alignments were both categorical variables. All balls put into play excluding sacrifice bunts in 2019 are considered in the study.

First, we consider a contingency table that shows the proportions of contact quality relative to shift. Then, to display the results graphically, we create a stacked bar graph that shows each shift and their batted ball quality proportions. Shift type is displayed on the horizontal axis while the batted ball quality proportions are shown using various colors. We then perform a chi-squared test to examine if there is a difference between batted ball quality proportions among shifts and non-shifts. We use a chi-squared test because we are testing for differences in proportions among two categorical variables. We then use a test for differences between two proportions to specifically test for a difference in the proportion of barrels (balls that the batter hit using the barrel of the bat) between non-shift and shift alignments. In this case, we collected the total number of barrels from all balls put into play relative to defensive alignment. We then address the context of the proportion test by asking the question: Does barreling a ball actually improve overall offensive performance? To do this we first create a stacked area chart that examines how batting average on balls in play (babip) as well as slugging percentage (slg) is affected by contact quality. We then create a multiple linear regression model that tests whether batting average on balls in play plus slugging percentage is a good predictor of weighted on base average (a widely used statistic that measures the overall performance of a hitter).

Results

The contingency table labeled Table 1 in the appendix shows the different proportions of batted ball quality between non-shifts (labeled Not Shift) and shifts (labeled Shift). From this table, we can see that the largest differences in proportions of contact quality arise when comparing balls barreled (Barrels) and poorly hit balls resulting from hitting the top of the ball (Poor(Topped)).

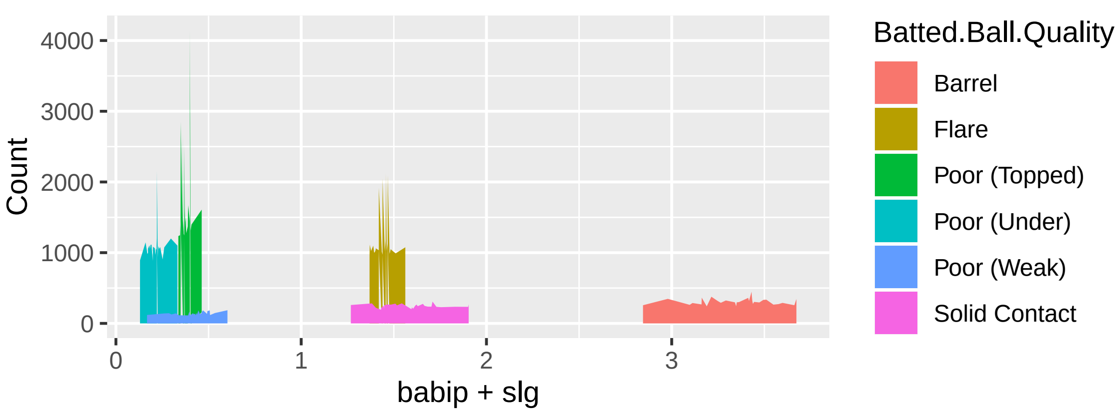
The different proportions of batted ball quality between defensive alignments are also displayed graphically using a stacked bar chart.



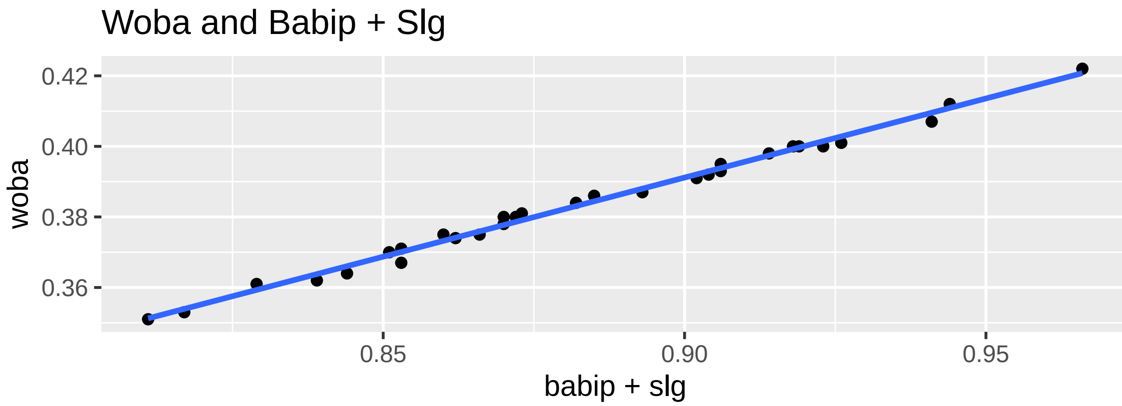
To evaluate statistical significance between contact quality proportions relative to defensive alignment, we use a Chi-Squared test. From the Chi-Squared test results, we obtain an extremely large Chi statistic of 262.74 and a correspondingly small p-value of almost 0 (2 × 10ˆ-16). From this, we can reject the null hypothesis, concluding that our data provide evidence of differences in the proportions of batted ball quality between non-shifts and shifts.

In an attempt to understand the baseball significance of the Chi-Squared test, we will now examine the specific proportional difference in barrels between non-shifts and shifts. The null hypothesis states there is no difference in the proportion of barrels when facing non-shifts and shifts. The alternative hypothesis states there is a difference in the proportion of barrels when facing non-shifts and shifts. We perform a test for difference in proportions, obtaining an X-squared statistic of 113.26 and a p-value of almost 0 (2 × 10ˆ-16). Consequently, we can reject the null hypothesis and conclude that our data provide evidence that the proportion of barrels against the shift is greater than the proportion of barrels against no shift.

Since we can now conclude that the proportion of barrels against the shift is greater than the proportion of barrels against no shift, we will attempt to answer if barrels affect offensive performance. We will use the statistic woba (weighted on-base average) as our measurement for overall offensive performance. Because of the nature of our data, it is difficult to directly compare batted ball quality to woba. Therefore, we instead compare batted ball quality to one combined statistic of babip + slg. These results are shown below using a stacked area chart. From the chart, we can see that barrels result in the highest babip + slg.



We will now use a linear regression model to examine the relationship between woba and babip + slg. As we look at the graph below, we can infer that there is a positive linear relationship between woba and babip + slg. We can conclude this by obtaining a correlation coefficient of .996. We then perform a standard linear regression test to see how well the model fits the data. Through the test, we obtain a Multiple R-squared value of .997, concluding that 99.7 percent of the variability in woba is explained by babip + slg.



Discussion and Conclusions

The analysis shows that there is a difference in batted ball quality proportions relative to the opposing defensive alignment. The data also suggests that on average in 2019 MLB teams barreled the ball more often when facing a shift than when facing a standard defensive alignment. It also appears that barrels lead to higher overall performance which was seen in the stacked area chart and mulitple linear regression model. Therefore, if the shift results in more barrels and that type of contact leads to the highest offensive performance, our data suggest that MLB teams are shifting more often than they should be.

However, we must remain careful about making this conclusion as optimal defensive alignment likely depends on the specific player batting. It is also important to note that the shift is a relatively new concept in the game of baseball and the data on it is limited to the last 5-7 years. Our study also only evaluated balls in play excluding sacrifice bunts from the 2019 season so we must be aware that these correlations might not carry over on a yearly basis. Additionally, we cannot prove causation as batted ball quality and defensive alignment were collected through an observational study.

To increase the practical use of this study, it will be important that further research on the topic is conducted. Repeating this study over different time periods will be one necessary step in the research. Breaking this data into subgroups such as left-handed batters vs right-handed batters and top hitters vs average hitters will also allow us to better understand the desired case by case results. Like the methodology of economic science that suggests economic phenomena cannot be truly understood until broken down all the way to the individual decision-maker, baseball phenomena can only be truly understood when broken down to an individual batter and pitcher. My initial research serves as an important first step in quantifying the results of the shift. However, to further understand the full effects of the shift it will be crucial to repeat the study over different time periods, break the data into more specific subgroups, and study the potential reasons why the shift results in a higher proportion of barrels.

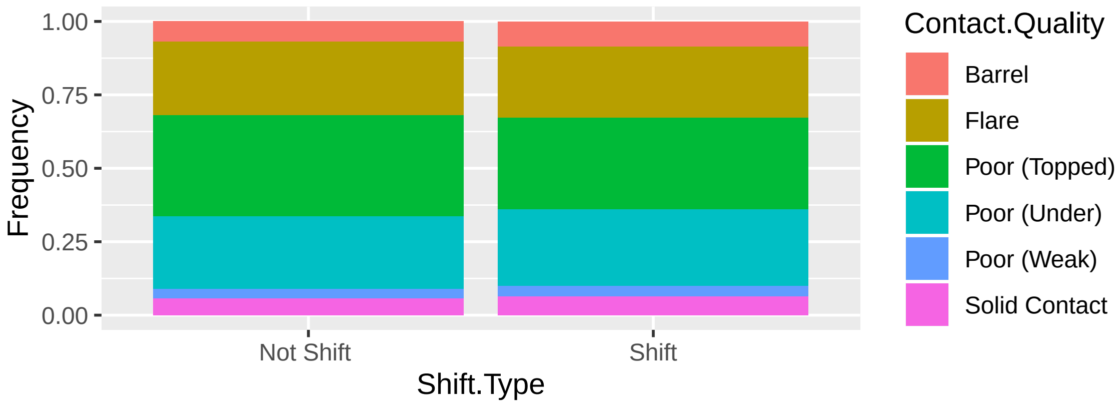
Bibliography

Travis_Sawchik. (2019, January 17). Don't Worry, MLB - Hitters Are Killing The Shift On Their Own. Retrieved from https://fivethirtyeight.com/features/dont-worry-mlb-hitters-are-killing-the-shift-on-their-own/.

Statcast Search. (n.d.). Retrieved from https://baseballsavant.mlb.com/statcast_search.

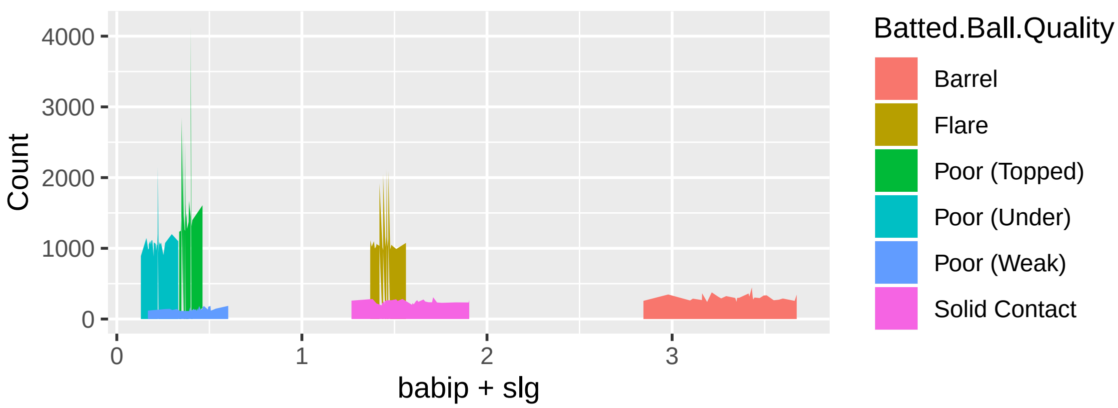
No_Little_Plans. (2016, June 30). Yes, The Infield Shift Works. Probably. Retrieved from https://fivethirtyeight.com/features/yes-the-infield-shift-works-probably/.

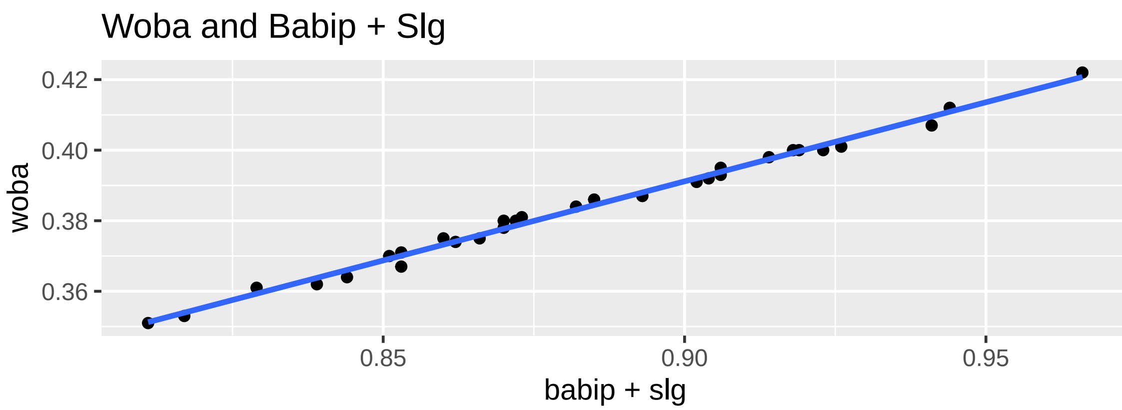
Appendix




```
##
##  Pearson's Chi-squared test
##
## data:  Shift.Table
## X-squared = 262.74, df = 5, p-value < 2.2e-16
```

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(4521, 5176) out of c(53555, 75488)
## X-squared = 113.26, df = 1, p-value < 2.2e-16
## alternative hypothesis: greater
## 95 percent confidence interval:
##  0.01336199 1.00000000
## sample estimates:
##     prop 1     prop 2
## 0.08441789 0.06856719
```






```
## [1] 0.9964978
```

```
##
## Call:
## lm(formula = woba ~ babip + slg, data = All.Regression)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -0.0021388 -0.0007497 -0.0001339  0.0006441  0.0020356
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.007441   0.005562   1.338    0.192
## babip       0.355565   0.017088  20.808   <2e-16 ***
## slg         0.461813   0.005438  84.926   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.00102 on 27 degrees of freedom
## Multiple R-squared:  0.9968, Adjusted R-squared:  0.9966
## F-statistic:  4215 on 2 and 27 DF,  p-value: < 2.2e-16
```

Table 1: Contact Quality Proportions

	Barrel	Flare	Poor (Topped)	Poor (Under)	Poor (Weak)	Solid Contact
Not Shift	6.856719	25.01325	34.39355	24.69134	3.348877	5.696270
Shift	8.441789	24.35627	31.16983	25.99197	3.674727	6.365419