# Does Handedness Still Matter In The MLB?

Max Brown, Evan Ginsburg, Luke Lorentzatos, Raymond Xiong

# Introduction

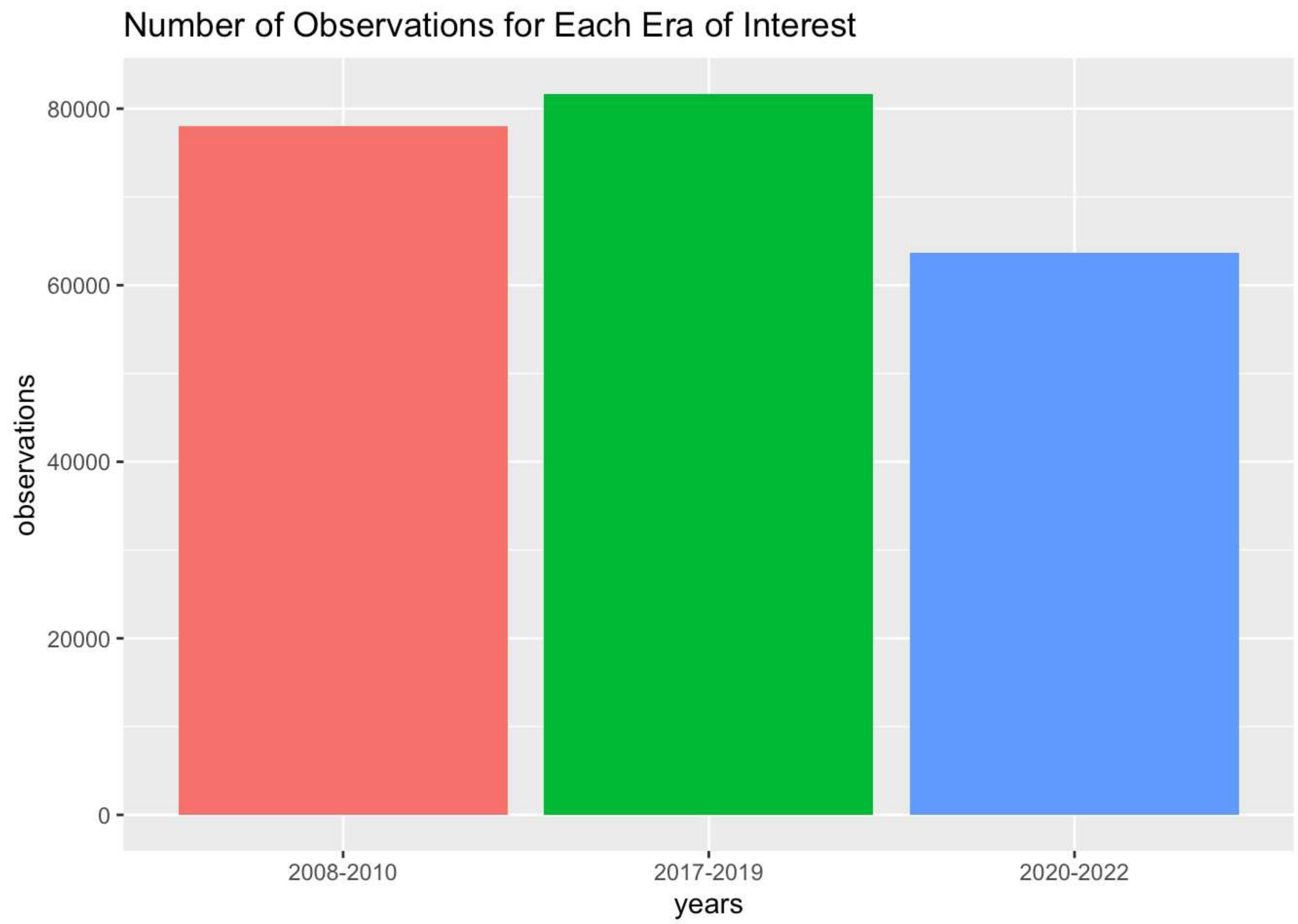
Baseball has historically been considered a game of skill, but over the last two decades, strategy and analytics have grown into integral aspects of an overall outcome. Managers of teams are responsible for setting an example and establishing a culture for their respective team, but they also have the final say in any lineup or bullpen related decision. Bullpens are vital to a team's success, as 9 of the 10 teams with the lowest regular season bullpen ERA reached the postseason in 2022. Today, organizations have analytics teams that conduct reports on a daily basis of how their relievers are best used against a given opponent, but there is a fundamental theory that hitters are less likely to be successful against pitchers who throw on the same side as the hitter's stance. Our goal is to test that theory, and evaluate how hitters' performances have varied against pitchers that throw from different sides across different time periods.

# Data Description

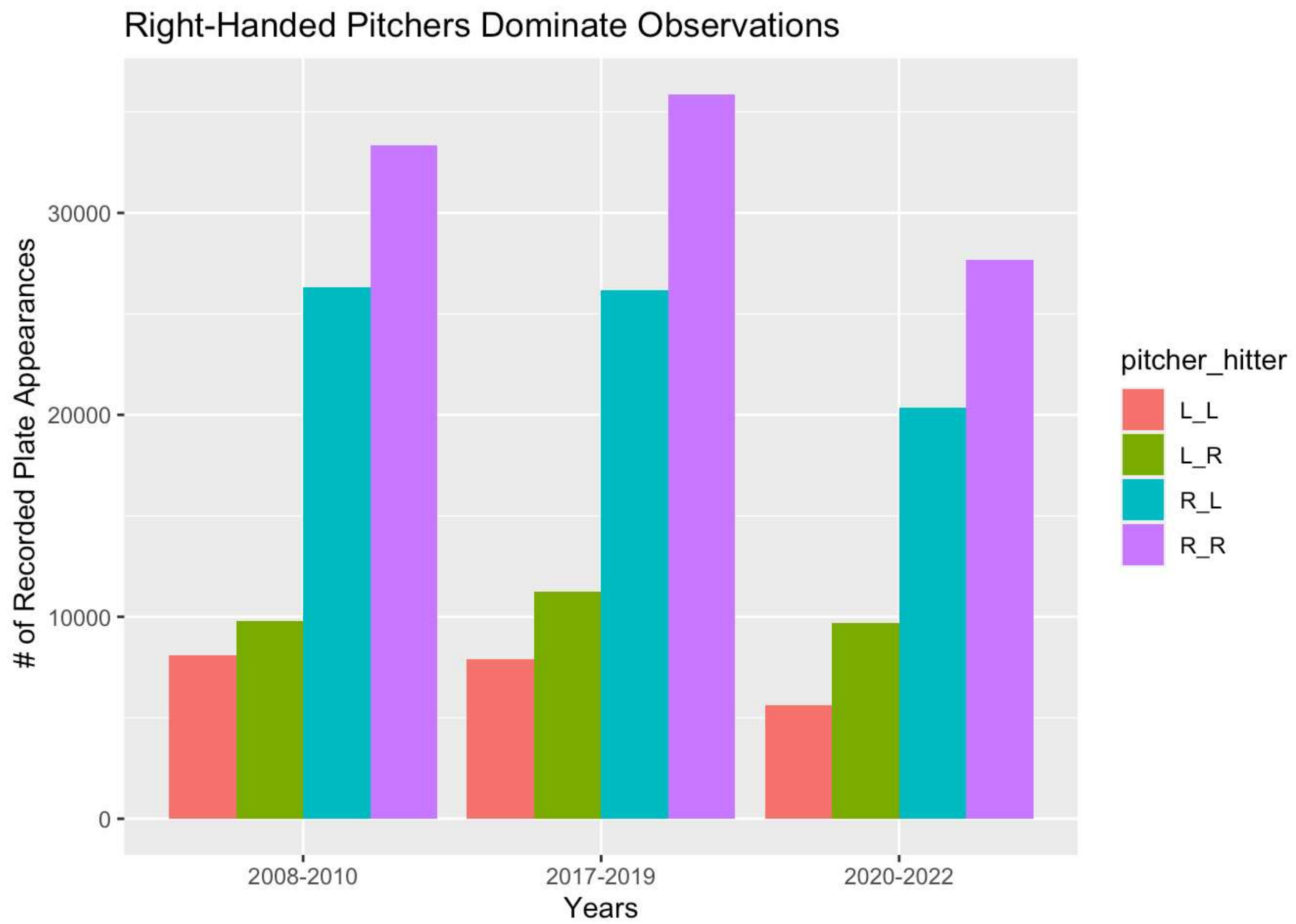
We sourced our data from baseballsavant.com, which provides the ability to filter for certain factors and allowed us to narrow our focus before pulling data into R Studio. We wanted to focus our data on intentional matchups, or cases in which a manager calls upon a reliever to face the next hitter. However, Baseball Savant does not contain a filter for pitchers following entry into the game, but instead allows for inclusion of a pitcher's first time through the batting order, which we included in our data selection. We only selected observations with relief pitchers, as we figured that a higher percentage would be a strategic decision, and we also filtered for the 7th inning onwards to remove instances of openers or relievers entering earlier in games that may already be lopsided in score. To further narrow our focus, we filtered for observations in which the hitting team was within 2 runs of the other team, assuming that these observations were significant in context and increasing the probability that strategically deployed relievers were called upon in these particular situations.

As mentioned earlier, teams in recent years have hired analytics departments to manufacture reports that aid managers in their decision making process. As a result, we wanted to determine if hitter results have changed over different eras of analytics in baseball. The earliest era we chose to evaluate was between 2008-2010, mainly because that is the earliest recorded data on Baseball Savant. From there, we broke the more recent era into two sections. Before the 2020 season the MLB voted to implement a rule that a relief pitcher must face at least three hitters once entering a game, a change with the intention of shortening the length of the game. Therefore, we chose to include 2017-2019 as a "middle era" and 2020-2022 as the "new era". Downloading three years of regular season data can be overwhelming, so we chose to only include the last pitches of matchups in our dataset, as we are only concerned with the final result and not necessarily the sequence of pitches prior in the matchup. Each member was responsible for loading in data from Baseball Savant in different time periods, and we proceeded to combine as necessary in R Studio.

# EDA




It is apparent that the latest era, 2020-2022, has significantly less observations than the other two data sets. This is important to note, and is the case because of the shortened season of just 60 games in the 2020 season. Despite this difference, we will not adjust our analysis because of the size of each data set.



While the number of plate appearances with right-handed pitchers is significantly greater, the proportionality does not change over each era, which is interesting considering the three-batter minimum rule implemented before the 2020 season.

# Analysis

# Basic Summary Statistics

Before looking deeper into models, we wanted to first evaluate summary statistics for each matchup in each era. In doing so, we needed to dive deeper into the "events" variable in the data set which outputted a string for the result of the play. The definitions of each summary statistics are as follows:

- AVG constitutes the overall batting average, which can be calculated by taking the number of hits and dividing by the number of at-bats. There are certain cases in which a batter can face a pitcher and an at-bat is not recorded, so we account for these situations in our statistics.
- OBP signifies on base percentage, which is the amount of times a hitter reached base (hits + walks + hit-by-pitch) divided by the total number of plate appearances. Similar to batting average, there are special cases in which the final result did not constitute a plate appearance.
- SLUG, known as slugging percentage, is similar to batting average but provides extra significance to doubles, triples, and home runs. This statistic takes total bases in the numerator and at-bats in the denominator.
- OPS is an abbreviation for on base percentage + slugging percentage, and has become one of the most predominant statistics for evaluating the success or contribution of a batter.

Since different events are considered differently in each of the first three statistics, the approach had to slightly vary. Therefore, we created different binary variables that constituted whether an observation is a plate appearance, an at-bat, reached base, or a hit. From there, we could filter for only plate appearances, or only at-bats, and count the number of cases that represent the numerator for the statistic of interest.

2008-2010 Summary Statistics

| Hitter | Pitcher | AVG | OBP | SLUG | OPS |
|---|---|---:|---:|---:|---:|
| L | L | 0.234 | 0.312 | 0.347 | 0.659 |
| L | R | 0.254 | 0.350 | 0.373 | 0.723 |
| R | L | 0.252 | 0.348 | 0.372 | 0.720 |
| R | R | 0.244 | 0.313 | 0.365 | 0.678 |

2017-2019 Summary Statistics

| Hitter | Pitcher | AVG | OBP | SLUG | OPS |
|---|---|---:|---:|---:|---:|
| L | L | 0.232 | 0.306 | 0.349 | 0.655 |
| L | R | 0.236 | 0.316 | 0.382 | 0.698 |
| R | L | 0.245 | 0.319 | 0.382 | 0.701 |
| R | R | 0.237 | 0.305 | 0.379 | 0.684 |

2020-2022 Summary Statistics (After Rule Change)

| Hitter | Pitcher | AVG | OBP | SLUG | OPS |
|---|---|---|---|---|---|
| L | L | 0.216 | 0.302 | 0.317 | 0.619 |
| L | R | 0.231 | 0.313 | 0.369 | 0.681 |
| R | L | 0.243 | 0.322 | 0.391 | 0.713 |
| R | R | 0.231 | 0.305 | 0.360 | 0.665 |

In evaluating each of the tables, it is apparent that hitters cumulatively have better results-based statistics against pitchers of the opposite hand (L v. R) across each time period. As shown in the EDA, there are significantly fewer L v. L matchups in our data set, however, it is important to note that particular matchup results in the values for each statistic in each timeframe. Analytics departments have most likely been aware of this decrease in offensive production, and as a result, one of the most common strategic decisions has been bringing in a left-handed pitcher to face a left-handed hitter. As previously mentioned, the 2020 rule implementation restricted managers from bringing in a pitcher for a singular hitter, and while there was a downward trend in production between 2017-2019, one matchup that does not reflect this pattern is right-handed hitters against left-handed pitchers, as it is the only matchup that saw an increase in OBP and SLUG in 2020-2022 from 2017-2019. While there can be many explanations for this difference in overall trend, one may be that left-handed pitchers are now obligated to face potentially right-handed hitters they would not have before the rule change in order to reach their three batter obligation.

# Weighted On Base Average (wOBA)

wOBA, or Weighted On-Base Average, is a rate statistic similar to OPS. However, instead of arbitrarily determining a double to be twice as valuable as a single, and one point of OBP and SLG to be equivalent, it attributes weights to each offensive outcome (walk, hit by pitch, single, double, triple, home run) based on their relative values. These values change slightly each year based on the run environment (or the number of runs scored per game). In general, an increase of 20 points in wOBA is worth around 10 runs over 600 PA. An average wOBA is around .320. After counting the number of at bats, walks, sac flies, hit by pitches, singles, doubles, triples, and home runs for each year's batter-pitcher matchups, we added up the walks, hit by pitches, singles, double, triples, and home runs multiplied by their respective weights for the numerator. Then, for each era, we would add the numerators (of each year and specific pitcher-hitter combination) and divide by a denominator that included the sum of the at bats, walks, sac flies, and hit by pitches for all the years in that era.

wOBA By Matchup Across Time Periods

| year | LL | LR | RL | RR |
|---|---|---|---|---|
| old | 0.303 | 0.320 | 0.325 | 0.306 |
| middle | 0.327 | 0.346 | 0.349 | 0.334 |
| new | 0.322 | 0.356 | 0.345 | 0.331 |

In evaluating the wOBA table, it is evident that left-right and right-left matchups are advantageous for hitters, with wOBAs well above league average, while same handedness pitcher-hitter matchups lead to a wOBA hovering near league average. This justifies the common managerial trend to seek a DH of the opposite handedness as the pitcher if at bat or to change the pitcher to the same handedness of the hitter if in the field. While the 3 batter minimum has made this more difficult to implement on the pitcher side since 2020, it is clear that the ability to optimize the handedness is as important as ever in limiting the opponent's wOBA, or maximizing your own.

# Creating Linear Models In Search For Matchup-Based Significance

Our analysis attempted to determine the significance of batter- and pitcher-handedness by creating a series of multiple linear regression models and fitting them to the data from each era. There were a total of 12 models: in each of the three eras, there was one for left-on-right (meaning a lefty pitcher throwing to a righty hitter, referred to as LR), left-on-left, right-on-left, and right-on-right. The handedness of each party was captured in a binary variable called “matchup,” which was a 1 if the matchup was the one corresponding to the specific model and a 0 otherwise. The other variables considered were the release speed of the pitch, the number of balls in the count, the number of strikes in the count, the difference in scores at the time of the at-bat, and finally, the number of outs at the time of the at-bat (implemented as a factor). These other variables were included as a sort of context for the model in an effort to reduce some of the noise and find the most accurate result. The response variable was the change in expected runs due to the at-bat, mostly because this value is already calculated by Baseball Savant and was a quick way to determine if an at-bat had a positive or negative outcome for the hitter.

After fitting all of the models, we found that there were only two which had significant results for the matchup variable. In the old era, left-on-left matchups were significantly worse for the hitter than the average matchup, with a p-value of 0.086. Being a left-handed hitter against a left-handed pitcher reduced the predicted change in expected runs by 0.011 compared to any other handedness matchup. In the latest era, right-on-left matchups were significantly better for the hitter than a general matchup, with a p-value of 0.008. Being a lefty hitter against a righty pitcher increased the predicted change in expected runs by 0.016 compared to any other matchup.

Left (H) vs Right (P) Matchups From 2020-2022

| term | estimate | std.error | statistic | p.value |
|---|---:|---:|---:|---:|
| (Intercept) | 0.051 | 0.034 | 1.528 | 0.126 |
| release_speed | 0.000 | 0.000 | 0.320 | 0.749 |
| matchup | 0.016 | 0.006 | 2.632 | 0.008 |
| balls | 0.007 | 0.002 | 3.419 | 0.001 |
| strikes | -0.054 | 0.003 | -18.701 | 0.000 |
| bat_score | 0.004 | 0.002 | 2.381 | 0.017 |
| outs_when_up_X1 | 0.005 | 0.005 | 1.044 | 0.296 |
| outs_when_up_X2 | 0.006 | 0.005 | 1.170 | 0.242 |

Left vs Left Matchups From 2008-2010

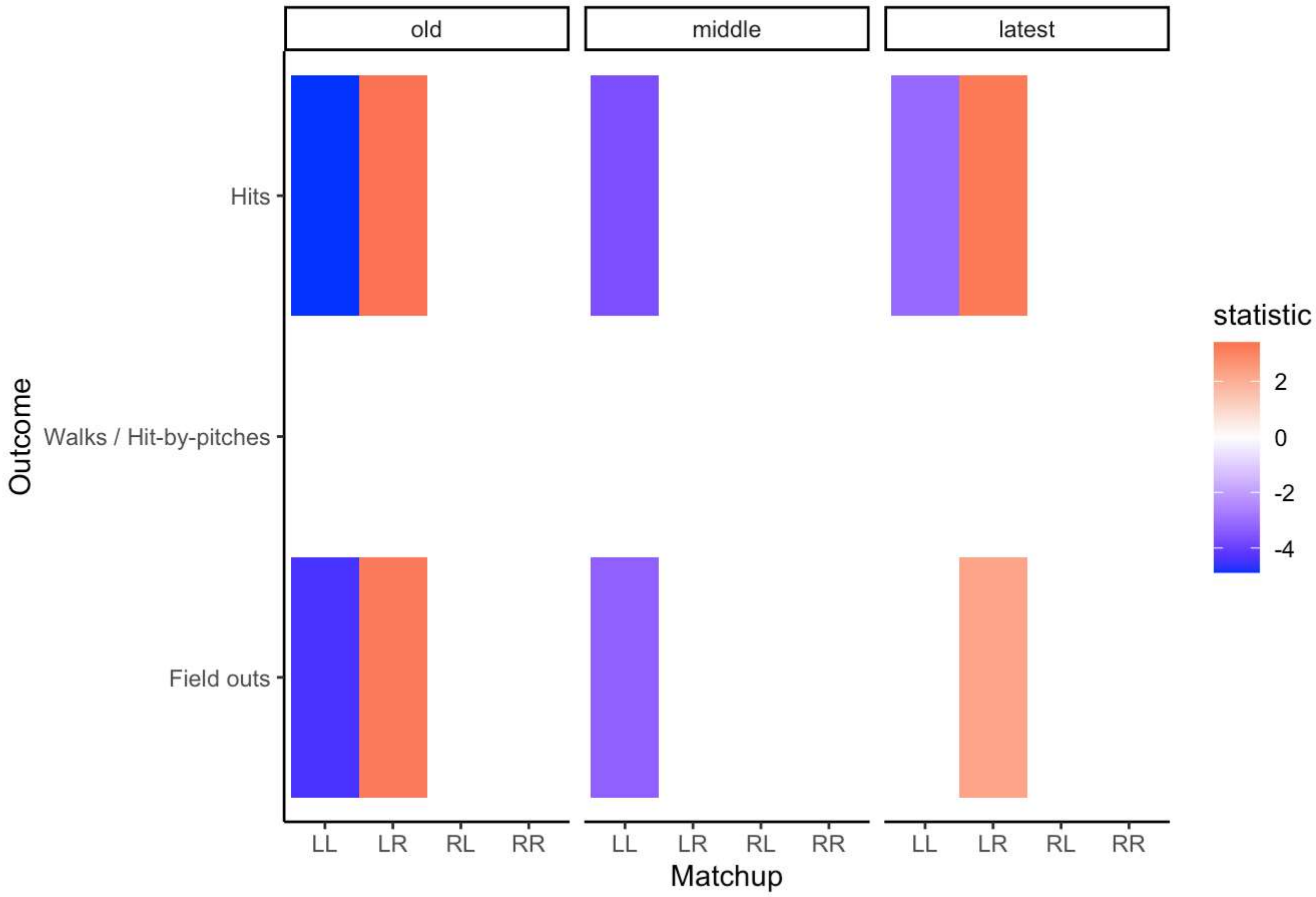
| term | estimate | std.error | statistic | p.value |
|---|---:|---:|---:|---:|
| (Intercept) | 0.014 | 0.031 | 0.457 | 0.648 |
| release_speed | 0.000 | 0.000 | 0.928 | 0.354 |
| matchup | -0.011 | 0.006 | -1.718 | 0.086 |
| balls | 0.004 | 0.002 | 2.080 | 0.038 |
| strikes | -0.042 | 0.003 | -16.242 | 0.000 |

| term | estimate | std.error | statistic | p.value |
|---|---|---|---|---|
| bat_score | 0.006 | 0.002 | 3.773 | 0.000 |
| outs_when_up_X1 | 0.004 | 0.005 | 0.955 | 0.340 |
| outs_when_up_X2 | 0.003 | 0.005 | 0.594 | 0.553 |

Because only two of the models had significant results, there is not too much that can be said about a lot of these matchups (especially the middle era, which had no significant results). However, one thing we can conclude is that the lefty-on-lefty matchup remains a very important one. Putting in a lefty reliever to face a lefty hitter was a statistically significant decision in favor of the pitching team in the old era, and that remains true in the latest era. While there was no statistical significance in the lefty-on-lefty matchup in the latest era model, because there was a significant increase in predicted change in expected runs when a righty pitcher faced a lefty hitter, it is still beneficial for the pitching team to have a lefty in to face a lefty hitter. This is only a general conclusion, as each individual pitcher has different splits, but it is still valid to conclude that on the whole, lefty hitters do better against righty pitchers, and thus worse against lefty pitchers. This result suggests that teams should make sure to have at least one or two good lefty pitchers in their bullpen, or at least one or two righty pitchers that fare as well against lefty hitters as lefty pitchers would, to avoid giving up a large advantage to opposing teams.

# Multinomial Logistic Models To Predict Pitch Outcome

In order to better understand how pitchers perform in different matchups between pitchers and hitters, we fitted multinomial logistic models to predict the outcome of a pitch. Outcomes that occur as a result of batting tactics like sac bunts and flies are eliminated because the focus of our analysis is pitchers. To simplify the prediction models, especially rare outcomes are eliminated, and the remaining outcomes are further classified into broad categories of “strikeouts,” “hits,” “walks or hit-by-pitches,” and “field outs.” The category “strikeouts” is set as the base level, mostly because strikeouts are ubiquitous and it would be more valuable to investigate how the probability of other more uncommon outcomes could change based on various predictors. Besides the matchup between pitchers and hitters, the predictors we chose are measurements of the pitch that are most intuitive and have clear physical meaning (i.e., release position of pitch, release speed, displacement on the home plate), as well as variables describing the game (i.e., whether there are runners on each base, number of outs, number of strikes, number of balls). There are 12 models in total, one for each era and each matchup (which lists the handedness of the pitcher and then that of the batter). For each model, we replaced the matchup variable with a binary variable encoding whether the matchup for a pitch corresponds to the one associated with the model. After fitting each model, we visualized the coefficients of the matchup variable. The following three plots visualized the t-statistic of coefficients for only statistically significant terms (i.e. terms with p-values smaller than the significance level $\alpha = 0.05$) for each of the three outcomes: hits, walks or hit-by-pitches, and field outs, in order. The x-axis specifies the model and is faceted by era, while the y-axis specifies the outcome. Warm colors indicate a positive coefficient (which indicates a positive association) and cold colors indicate negative associations. More saturated colors indicate greater coefficients in terms of magnitude.




Only matchups with lefty pitchers have significant coefficients, which is reasonable because righty pitchers outnumber lefties and thus would exhibit more variation in their skills and personnel. Walks and hit-by-pitches do not show significant associations with matchups. In the old era, the lefty-to-lefty matchup is associated with lower probabilities of hits and field outs, which also indicates a higher probability of strikeouts. This tendency lasted till the middle era, and in the latest era the association between the lefty-to-lefty matchup and probability of field outs becomes insignificant. The matchup between lefty pitchers and righty batters is positively associated with higher probabilities of hits and field outs in the old and latest era, indicating a negative association with this matchup and probability of strikeouts. Overall speaking, lefty pitchers perform better when playing against lefty batters in the sense that they tend to have more strikeouts and fewer hits in lefty-to-lefty matchups, and they tend to be less advantageous when playing against righty batters. This tendency is roughly consistent throughout all three time periods. Therefore, it would be beneficial for a team to have competent left-handed relievers in their bullpen to have an edge when playing against lefty batters.

# Discussion & Limitations:

In looking at a broad view of our analysis, it is apparent that left vs left matchups are advantageous across all time periods, and the strategy of bringing in left-handed relievers for these matchups should continue regardless of the rule change implemented in 2020. Although our work provides a picture for strategic decisions of the pitching team, context and individual matchups prohibit consistent usage of our matchup strategy. For example, there are hitters such as lefty Nathaniel Lowe of the Texas Rangers, who had a 2022 batting average of .330 against left-handed pitching compared to .288 against right-handed pitching (Fangraphs). The handedness of the hitter is not the only factor in determining the best matchup at the end of close games, which is precisely why teams have begun to invest millions in analytics departments to gain an advantage on a matchup basis. Hitters' success against certain types of pitches has emerged as another integral part of teams' scouting of their opponents, placing pitchers' strengths against hitters' weaknesses.

This confounding effect of “reverse platoon splits” players, those that actually perform better in matchups which are supposed to be unfavorable, could be one of the primary reasons why so few of the fitted models returned significant results. Future work could attempt to account for this by categorizing pitchers and hitters as “traditional lefty” or “traditional righty” based on their splits rather than their actual handedness. Regardless, it was interesting to see that in the middle era, which was the overlap of relievers being allowed to face a single batter and relievers being better than ever, none of the models were significant. Despite the advantages of shorter outings and being pre-crackdown on foreign materials used to increase spin rate, hitters still managed to fare approximately equally well on the whole against both same- and opposite-handed pitchers.

Our models predicting the outcome of a certain pitch are only additive linear models, in which interactions between predictors and nonlinearity of results are neglected. Future work could account for these aspects using more complicated models. The direct influence of this is that walks and hit-by-pitches are better predicted: introducing interaction between predictors and nonlinear functions like quadratic functions could help models address how far a pitch is from the center of the strike zone and thus better predict walks and hit-by-pitches. It is also interesting that matchups with righty pitchers do not have any significant coefficients for the matchup variable, which might be due to the greater variation in personal split among righty pitchers caused by the matchless outnumbering of righty pitchers. In future models, righty pitchers could be further classified according to their performances in previous pitches.