Bullpen Strategies for Major League Baseball

Paper Track: Baseball
Paper ID: 1684

1. Introduction

As never before, Major League Baseball (MLB) teams are turning to analytics in an attempt to gain a number of small advantages that, in the composite, may result in significantly altering the odds of winning in their favor. This changing mindset was on full display during the 2016 MLB postseason, where teams showcased several strategies attributed to the field of sabermetrics [3, 6, 8]. One of these was their interesting use of relief pitchers; post-season managers removed starting pitchers from games earlier than in any other postseason to date [2], perhaps to avoid high pitch counts [12] or to prevent opposing lineups from seeing the same pitcher too many times during a single game [10]. The result was, according to popular media, one of the most exciting World Series in recent memory [11].

Several issues arise when witnessing paradigm shifts in how baseball is played, and this paper attempts to address two interesting ones surrounding the potential use of relief and starting pitchers. Firstly, we look at the home-field advantage and propose a strategy for starters of visiting teams that can be used to remove roughly one half of the first-inning advantage to the home team. A byproduct of this analysis is a set of proper adjustments that must be made to calculate the true home-field advantage, which is roughly 0.429 runs/game, rather than the 0.133 runs/game suggested by the scoring data. Secondly, we wish to tackle the age-old question of when to remove a pitcher from a game by proposing a pitcher-by-pitcher analysis that utilizes data that can be easily measured during a game for real-time decision making.

Since each of these two topics can be addressed independently, they are discussed separately in Sections 2 and 3, respectively. We then conclude the paper in Section 4 with a summary of the pitching strategies of the paper, and a list of potential bullpen tactics for future research.

2. Mitigating the Home Field Advantage

In this section, we discuss sources of the home field advantage, and identify the first inning as the most advantageous inning for the home team. We propose a novel strategy for starters when pitching on the road that can remove a significant piece of the first-inning advantage from the home team, and discuss the implications of our findings. The data used to form the results of this section are from retrosheet.org [1], and specifically include game log data from 1980--2015. These data were selected because it was shown in [16] that the home team's advantage in scoring has been roughly the same since the 1980s, but experienced a (thus far unexplained) jump in the 1970s. Data selection was, therefore, made to keep our findings and strategies applicable to the game in its current state. For simplicity, only games that lasted at least 51 outs in which the home team batted last are included in the analysis of this section, and following the convention in [16], all extra innings were compiled together and are presented as the "10th inning."

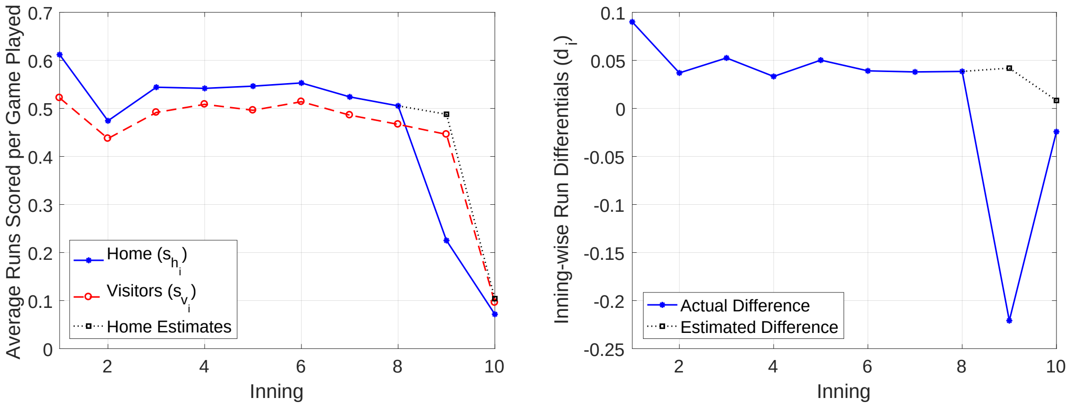

Figure 1: LEFT: Average number of runs scored by inning for home and visiting teams. RIGHT: Average run differential by inning calculated as home teach score minus visiting team score. Since the home team does not bat in the bottom of the last inning once they have the lead, the true advantage to the home team in the 9th and 10th innings must be estimated using other techniques, and these estimates are also included in the figures.

2.1. Background

Over 80,000 regular-season baseball games were played in the Major Leagues from 1980--2015, and home teams won 53.93% of those games. Home teams scored a total of 370,331 runs (4.597 runs per game) during that stretch compared to 359,589 runs (4.464 runs per game) produced by visiting teams. The overall scoring differential can then be calculated to be 0.133 runs per game, and could have been higher if the bottom of the last inning were always played to three outs. Much research has been done to attempt to identify the sources of the home-field advantage (see [4] and references). Everything from team travel schedules to umpire bias [13] to familiarity with the home ball park has been analyzed, and many of these factors have been shown to be possible sources of the bias toward the home team. When one investigates this advantage by inning, it is surprising to find that a large portion of the scoring differential between home and visiting teams occurs in the first inning [16, 17, 19]. Let $n =$ 80,554 be the number of total games in the data sample, and let s_{h_i} (s_{v_i}) be the average number of runs scored by the home (visiting) team in inning i for $i = 1, 2, \ldots, 10$; i.e.,

$$s_{h_i} = \frac{\text{total runs scored in inning } i \text{ by home teams in } n \text{ games}}{n}, \tag{1}$$

and similarly for s_{v_i} for $i = 1, 2, \ldots, 10$. Then inning-wise scoring differentials are calculated as

$$d_i = s_{h_i} - s_{v_i}, \qquad \text{for } i = 1, 2, \ldots, 10. \tag{2}$$

Figure 1 shows both s_{h_i} and s_{v_i} in the left-most figure, and d_i in the right-most figure for $i = 1, 2, \ldots, 10$. Perhaps the most glaring problem with these calculations is that the data appears to indicate that the visiting team has a scoring advantage in innings 9 and 10, but this is actually not true and will be addressed later on. For now, simply note that the estimates represented with black dotted lines and square markers are likely more accurate when measuring the true advantage to the home team for innings 9 and 10.

The first inning differential between home and visitor scoring in the regular-season data set from 1980--2015 is $d_1 = 0.09$ runs, which seems to indicate that the first inning is responsible for roughly

67.5% of the overall home team advantage of 0.133 runs per game. This approach is actually taken in [16] by Smith (although his figure comes out to be only 58% when he takes data from 1909--2013 into account), where he argues that the biggest reason that a home-field advantage exists is because there is a first inning (although we feel this number is actually much too high and will provide more accurate estimates later). Furthermore, in [17] Smith goes on to show that the advantage of a home team in the first inning is highly correlated ($R^2 = 0.86$) with the length of the top of the first inning in time (the linear fit indicates an extra run for the home team in the bottom of the first inning for every 90 seconds beyond 6 minutes for the top of the first). The connection is that it appears the starter on the visiting team is effectively "cooling down" in the dugout during the top of the first, and when he finally pitches in the bottom half of the inning, he is less prepared to face the opposing lineup. Similar results were found in [19], where it is shown that pitchers' velocity in the first inning tends to be much higher when the pitcher is at home verses when he is on the road. Also, strike-out-to-walk ratio (K/BB) in the first inning is affected, tending to be smaller when pitchers are on the road versus at home.

2.2. Methodology

2.2.1. Estimating the 9th and 10th Inning Differentials

We wish to first properly estimate the percentage of the home-field advantage that exists in the first inning, and then propose a simple strategy for visiting teams to mitigate much of that first-inning bias. The problem with stating that the first inning is worth 67.5% of the overall home-field advantage is that the true baseball data shows advantages in run differentials for visitors in both the 9th and 10th innings, but if this were true we would expect visiting teams to win more extra inning games. For the 7,246 extra inning games in our data set, however, the home team still managed to win 52.43% of them. The negative differentials in Figure 1 for innings 9 and 10 arise because when the home team wins, it doesn't use all of its outs in the bottom of the last inning; and thus estimates favor the visiting team even though the visitors remain at a disadvantage throughout the game.

Consider, for example, an extra inning game in which the visitors score six runs in the top of the 10th, and the home team scores zero runs in the bottom of the 10th. In this case, all six runs scored contribute to the statistic $s_{v_{10}}$. Consider now the exact opposite, a game wherein the home team would have scored six runs in the bottom of the 10th, and the visiting team scores zero runs in the top of the 10th. Except in the case of a multi-run walk off home run, the home team would be credited with only one run to contribute to $s_{h_{10}}$, and the overall statistic is devalued even though the home team wins the game.

Some strategies for estimating the 9th and 10th inning data points are (a) to consider only a subset of the innings played, or (b) to give estimates based on the number of outs used in the inning; but both of these techniques are flawed and will result in biased estimates in favor of the visitors by discounting cases where home teams were likely to score more runs and counting cases where they were unlikely to score more runs. Thus, our approach to making these estimates is to trust the overall home-field advantage (since home teams still win a majority of extra-inning games), and apply an average differential to s_{v_9} and $s_{v_{10}}$ to estimate average home team scoring potential in these innings. We define d_{ave} as the average scoring differential using innings three through eight, discounting innings one and two because of irregular behavior in the differential (d_1 is an inflated differential and d_2 is a deflated differential, probably due to the home team being closer to the bottom of the order in the second inning). We must be careful in our calculation of the 10th inning because only a portion of games actually have extra innings, and thus there is an extra weighting parameter required to accurately estimate the 10th inning advantage. To be specific,

$$\hat{s}_{h_i} = s_{v_i} + \Pr(\text{inning } i \text{ is played}) \underbrace{\left(\frac{1}{6}\sum_{j=3}^{8} d_j\right)}_{d_{\text{ave}}}, \qquad \text{for } i = 9, 10, \tag{3}$$

and $\hat{s}_{h_i}$ signifies the estimate of s_{h_i} if the bottom of the last inning was always played to three outs. The probability that the 9th inning is played is one for our data set, and the probability that the 10th inning is played is 0.1973, which is the number of extra innings divided by the number of games in the data set. Estimates are given in Figure 1, and $d_{\text{ave}} = 0.0419$ for our dataset. Using these estimates, we are able to get a better idea of the total advantage to the home team, and likewise the overall percentage of the advantage experienced in the first inning. If we choose to ignore this subtlety, we ignore the fact that the home team has an advantage throughout the bottom of the last inning, whether or not they actually use it in a real game.[1]

2.2.2. First-Batter Starter Strategy

Now, we propose a novel strategy for removing a significant chunk of the home-field advantage, assuming that the chief reasoning behind the elevated advantage in the first inning is due to suboptimal timing of the warmup for a visiting starter. Our strategy allows the effective "starter" to warmup both before the game and during the top half of the first inning, because he will not actually start the game. Some [5] have advocated for the use of relief pitchers to pitch the first inning in the role of an "opener," but this strategy would not solve the problem of the opener cooling down in the top of the first. Instead, we propose a "first-batter starter." Although some people will hate this strategy because it means one more relief pitcher in the game, it stands to reason that if a starter is made less effective by sitting in the dugout during the top of the first inning rather than warming up, a visiting team should want to be subject to that disadvantage for as little time as possible. The rules of baseball actually allow for pitching substitutions mid at-bat, but only if the pitcher to be replaced has already pitched to at least one complete batter or is injured [14]. We assume that the first-batter starter will be at the heightened disadvantage of d_1 runs/inning for one batter (probably a high estimate since this pitcher can throw as hard as he likes knowing he only needs to face one batter), and can then be replaced with the effective starter fresh from the bullpen, who will only be subject to the regular differential of d_{ave} runs/inning. Using data from 2015, we calculate the average number of plate appearances per inning as

$$\bar{\text{PA}} = \frac{3}{\text{PO}} \left(\text{AB} + \text{BB} + \text{IBB} + \text{HBP} + \text{SF} + \text{SH} + \text{CI}\right), \tag{4}$$

where AB: at-bats, BB: walks, IBB: intentional walks, HBP: hit by pitch, SF: sacrifice flies, SH: sacrifice hits (bunts), CI: first base awarded for catcher interference, and PO: put outs by opposing team, and all these stats are at the per-game level. Notice, this is simply the fraction of all plate appearances that we'd expect to see over the span of three outs in a baseball game. This calculation gives $\bar{\text{PA}}_h = 4.30$

[1]Several works actually make this error by calculating the number of runs they feel a source of the home-field advantage is worth, and then simply taking a ratio against actual scoring, like in [13] where the authors claim that umpire bias is worth 2/3 of the home-field advantage, but their methods actually give a number close to 1/5 when the total advantage is properly calculated. One sure effect of this miscalculation is the overestimation of the importance of specific items to the home-field advantage in baseball.

($\bar{\text{PA}}_v = 4.21$) plate appearances per inning for home (visiting) teams. Now we have everything we need to estimate a new first inning differential given the first-batter starter strategy as

$$\hat{d}_1 = \frac{d_1}{\bar{\text{PA}}_h} + \frac{d_{\text{ave}}(\bar{\text{PA}}_h - 1)}{\bar{\text{PA}}_h}. \tag{5}$$

2.3. Results and Discussion

The results of our study indicate that the first run differential of $d_1 = 0.09$ is actually only 21% of the total home-field advantage, which of course is quite at odds with the 67.5% calculated earlier (or even the 58% reported in [16]), but has a much more intuitive feel. After all, if the first inning differential is around twice the differential for any other inning, it stands to reason that most innings account for about 10% of the advantage, and the first inning should be around twice that number. A full listing by inning of the average advantage to the home team in runs/game and percentage of the whole is given in Table 1.

Table 1: Run differentials after the adjustment and percentage of total home-field advantage in each inning.

Inning	1	2	3	4	5	6	7	8	9	10	Total
d_i (runs/game)	0.09	0.04	0.05	0.03	0.05	0.04	0.04	0.04	0.04	0.01	0.43
d_i (%-age of whole)	21.0	8.6	12.2	7.8	11.7	9.1	8.9	9.0	9.8	1.9	100

We also found that the new value of d_1 that could be possible by having a first-batter starter for away games is $\hat{d}_1 = 0.053$ runs/game, which could be a high estimate. This has the potential to remove approximately 0.037 runs/game (about 8.6%) of the total home-field advantage. While this may appear to be a small amount, we remind the reader that this shift requires no training, hard work, or discipline of any kind; only a (perhaps uncomfortable) trip to the mound for the manager during the first inning of each and every away game of a season. Over the course of 81 away games per year, this amounts to roughly 3 runs, which could easily result in an extra win (17.8% of the games in our database were lost by the visiting team by exactly one run).

3. Comparison of Starting to Relief Pitchers using Data Analytics

We now shift our attention to the second question addressed in this paper: when should a manager remove a pitcher from a baseball game? It is safe to say that this question has yet to be answered in a way that satisfies managers, players, owners, and fans alike. It was discussed at length in [18] how managers have been in possession of various data sets for years as well as indicators that should apply when deciding whether or not to remove a pitcher; and yet, managers still feel that their "gut instinct" is more valuable at times than all the data in the world. This is likely because good managers are, in fact, internalizing crucial pieces of data that help them to make good clutch decisions, and the analytics has yet to incorporate all of these data into a workable system. This section provides a simple system based on pitch counts and the strike-to-ball (STB) ratio that make up a nice start towards this end. While simple, it is surprising what can be learned through these two observable variables. This section provides evidence that pitchers' effectiveness can, in fact, be monitored in real time, giving

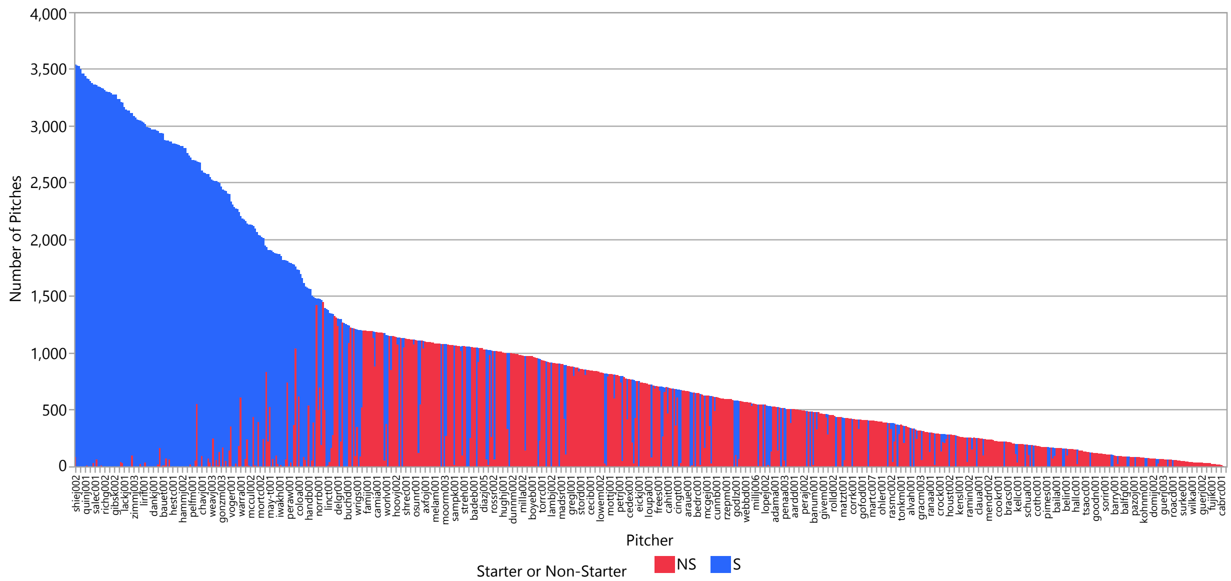

Figure 2: Stacked bar graph of the number of pitches by pitcher in 2015, sorted in descending order. Blue sections of bars indicate starting pitchers, while red sections of bars represent non-starting pitches. A large majority of pitches from a player will be of one of the two types (i.e. Starter or Non-Starter).

managers straightforward techniques for deciding when to replace a pitcher. We use play-by-play files from Retrosheet.org [1] for the 2015 season for the analysis of this section.

3.1. Methodology

Almost 700,000 pitches were observed and recorded during the 2015 MLB season. These pitches came from 712 different pitchers for an average number of just under 24 different pitchers per team. Since a team's active roster is comprised of only 25 players (usually with only five starters), it is evident that franchises are indeed using and seeking to leverage the bullpen in many ways to help relieve, close, and save games.

Figure 2 shows the sorted distribution in descending order of all pitches across the 712 pitchers. Here we see that 325 different pitchers started at least one of the 2,429 games in 2015 (counting both home and visiting teams there were 4,858 total starts for the season). Collectively, these 325 starters (45% of the 712 pitchers) hurled over 64% of the balls originating from the mound before leaving the game. These 461,317 pitches are colored in blue on Figure 2. The non-starter pitches are colored in red (36% of the total area) on the same figure.

Interestingly, of the 325 "starters" 174 of them (53.5%) also pitched in a non-starting role at least once during the season. Analysis reveals that an additional 48,923 pitches come from these "starters" when they are not starting (10% of their pitches are in relief of other starters).

Figure 2 suggests that there is indeed mixing of roles in the individual games. Naturally, teams are responding to the season's game schedules, unforeseen injuries, and experimental contracts with minor league players. We propose that, among other strategies, this "mixing" can be extended such that a franchise employs twice as many starters (10) and uses them in much more varied sequences, specialized match-ups, and different rotations.

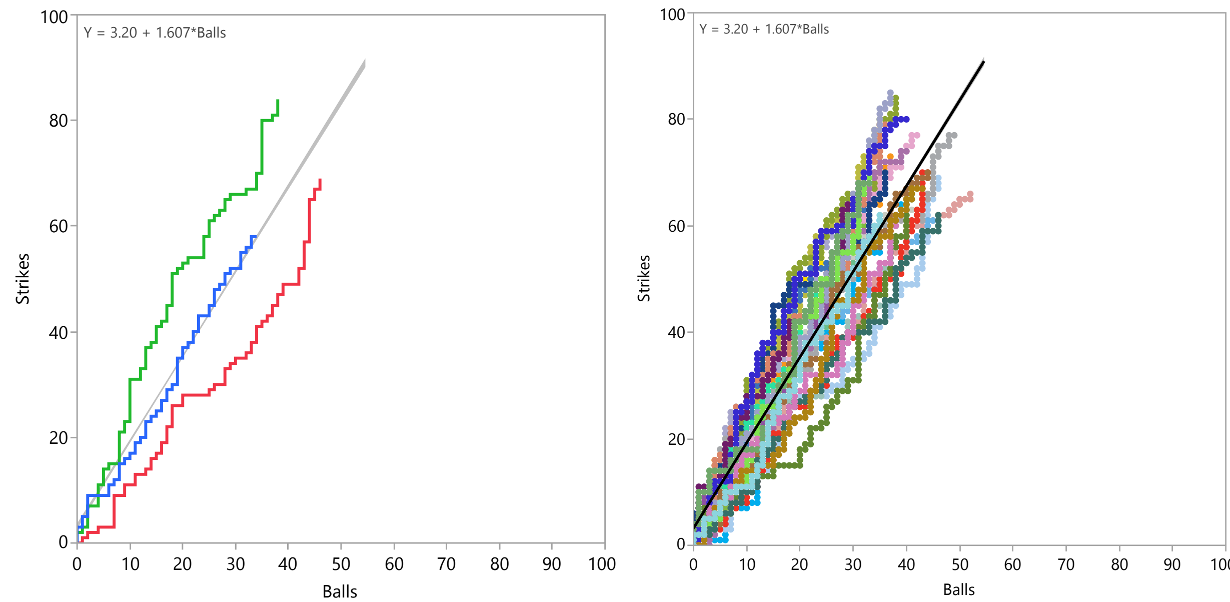

Figure 3: Plots of strikes versus balls for three games (LEFT) and all games (RIGHT) in which Jake Arrieta started during the 2015 MLB Season (colored by game).

In order to quantify the validity of this proposition, we look at the pitch-by-pitch performance of the starters versus the non-starters. We track the strikes, balls, and STB ratio with respect to the pitch count to ascertain if indeed closers could start, i.e., perform equal to or better than starters, but for a smaller number of innings. Data is aggregated across different groups and categories with comparisons between teams and individual pitchers.

3.2. Results and Discussion

The STB ratio can be a useful metric to evaluate the current state or health of a pitcher. The authors recognize that this ratio is not the only pitching statistic that matters and have even developed derivatives of the popular ERA statistics (see [15]). However, for purposes of initial comparisons between pitchers within the scope of this paper, this statistic provides a wealth of information that should be analyzed prior to complicating the matter. Undoubtedly, similar comparisons and analyses using other measures, including the component ERA statistic, will be explored in the future.

The left-most plot of Figure 3 displays balls to strikes throughout three games in which Jake Arrieta of the Chicago Cubs started in 2015. The straight grey line, with accompanying equation at the top left (Y stands for strikes), indicates the linear regression of all of his games in 2015 based on the relationship between balls and strikes. In other words, this line represents an expected strike-to-ball ratio at which Arrieta should be at for different pitch counts. For one of the three games represented, Arrieta followed this line almost exactly the entire game. In the other two games, Arrieta was pitching either above or below his average strike-to-balls ratio for much of the game. When a pitcher pitches significantly below his average trend line, this may be an indicator that he is fatigued or experiencing other issues that may lead to a bad start. On the other hand, when a pitcher pitches above this line, this may be an indicator that he will pitch at levels above his average throughout the night. Of course, at anytime a pitcher may regression back towards his mean, and therefore early signs of superior or

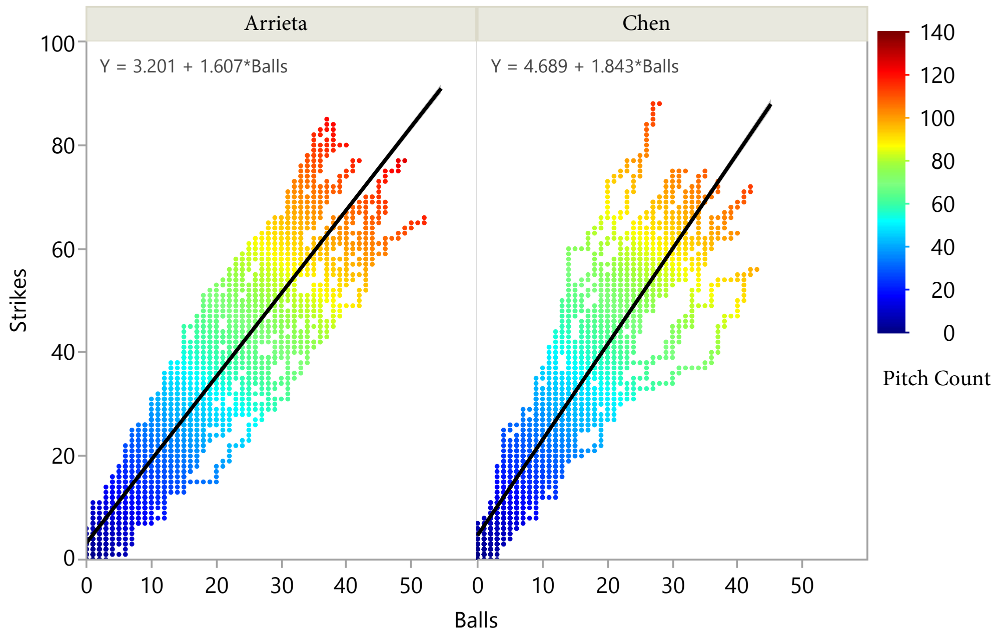

Figure 4: Plot of strikes versus balls for all pitching sequences for Jake Arrieta (LEFT) and Wei-Yin Chen (RIGHT) during the 2015 MLB Season colored by pitch count.

inferior pitching should be taken with a grain of salt.

The right graph in Figure 3 extends this concept to all 33 games in which Arrieta started in the same year with the best fit line reproduced from the left hand figure. Although, each game is represented in a different color, certain strike and ball combinations are reached in multiple games and therefore many "strike-to-ball profiles" are observable in his overall dataset.

The most significant number in the equation for fit lines is the slope coefficient, which appears as a multiplier to the "Balls" variable. This slope is a measure of the strike-to-ball ratio (STB), which is generally desired to be as large as possible (although we also note that hits are counted as strikes, so is it theoretically possible to have an excellent STB while still being ineffective as a pitcher).

By overlaying the pitch count on top of the strike-to-ball region for Arrieta as shown in Figure 4, one can readily see the final ratio with respect to the pitch count in multiple games. In the same figure, the analysis is performed for Wei-Yin Chen, who actually has a higher slope for the fit line but a much wider strikes-to-balls region, with more games significantly deviating from the trend line. In terms of the number of pitches per game, Arrieta averaged 106 pitches per game while Chen averaged 97. Also, Chen's ERA, 3.34 for 2015, was almost twice that of Arietta's, 1.77 suggesting that many of Chen's "strikes" resulted in more hits and runs in comparison.

Comparisons between pitchers using this method can now be explored. In Figure 5, the strikes-to-balls profiles for all Chicago Cubs pitchers are presented for the 2015 season. The familiar region and best fit line for Arrieta is contrasted with 24 other pitchers during 2015. The slopes and best fit lines are generally the steepest for starting pitchers, with the small regions very rarely composed of more than 40 pitches for relief or closing pitchers. Figure 5 shows that occasionally some non-starting pitchers, such as Richard, Wada, and Wood offer STB ratio fit lines that are in fact better than the starters. Although the reduced data at higher pitch counts suggests caution for predicting the performance of these pitchers over more innings, the potential to leverage these pitchers in other ways could be an advantage without loss of average team performance.

Combining these strike-to-ball regions for all 30 teams and for all 712 pitchers from the 2015 season results in the left hand side of Figure 6. This aggregate region, with associated fit line, illustrates the typical range of performance expected by any professional major league pitcher. The coloring is based on frequency quantiles demonstrating that many more pitching states have, as expected, combinations of a low number of balls and strikes. On the other hand, certain strike and ball combinations may occur only once in a season such as the one pitcher's STB profile extending out to approximately 45 balls and 21 strikes towards the bottom. In fact, this particular profile by pitcher Tyler Matzek is a true outlier. After the first batter in the game was grounded out, Matzek threw 9 straight balls. Amazingly, the Colorado Rockies were able to finish the inning without giving up any runs. In the second inning, Matzek threw 8 balls in a row. Again, the Rockies' defense let no runs. Finally, in the third inning after 6 straight balls Matzek was eventually pulled, producing the lowest strikes-to-balls ratio for the whole season with more than 50 pitches. Arizona never scored against him but eventually went on to win the game.

The right hand side of Figure 6 shows the distribution of the STB for all 712 pitchers, with an average value of 1.32. This average is lower than the STB coefficient of 1.603 for the entire league on the left hand side. The main reason this is observed is an artifact of the larger number of events close to the bottom left corner of the strikes versus balls graph driving the slope higher. This effect is also seen in Figure 7 where the league pitchers are divided into starting pitchers (on the right) and pitchers which never start (on the left). The corresponding STB coefficients are again presented in the equations for the best fit lines at the top left. As expected, the starters are much more proficient overall at delivering more strikes to the plate than balls. In addition, the non-starting pitchers are often replaced more quickly if performance appears to be below average. Therefore, in the left hand side of Figure 7 above average performing relief pitchers are permitted to continue to higher pitch counts (i.e. profiles above the best fit line). However, in both cases the trend line is dominated by a majority of activity at low strike and ball counts.

Extracting the STB coefficients for all 712 pitchers and plotting these values with respect to the number of pitches in a season is shown in Figure 8. As before, the data points are colored based on starting (blue •) or non-starting (red +) pitchers. A distinct difference between the starters and non-starters is apparent in Figure 8. The categorization of these two groups was seen previously in Figures 2 and 7 where a "bend" in the order of pitchers sorted by number of pitches was observed and the STB coefficients were different for the two groups, respectively. However, Figure 8 also shows that a significant number of pitchers that never start are able to match or even exceed some starting pitchers with respect to STB over a shorter time frame. This suggests that relief or closing pitchers potentially have the capability to replace starting pitchers much earlier in the game if in fact the starter is showing signs of either fatigue or slumping performance. In addition, there is a distinctive absence of pitchers in the 1,200 to 2,000 pitch count and in a range of STB from 1 to 1.25, or where the sharp cut-off in the distribution is observed. This suggests that the potential to have pitchers train and perform in a type of hybrid role between the two traditional extremes of starters and closers may be untapped. Although this would necessitate alterations to pitching schedules and training regimens, a "three inning" pitcher that plays in every third or fourth game could increase team performance, and potentially be more cost effective as well.

Extending this STB analysis further can provide an additional metric in judging when to pull a pitcher from the mound. Figure 9 presents 10 pitchers from the Chicago Cubs, five starters on the top row, and five relief pitchers on the bottom row. Each subgraph has two parts. In the bar chart section of each player's subgraph, the bars represent the frequency with which the player reaches different pitch

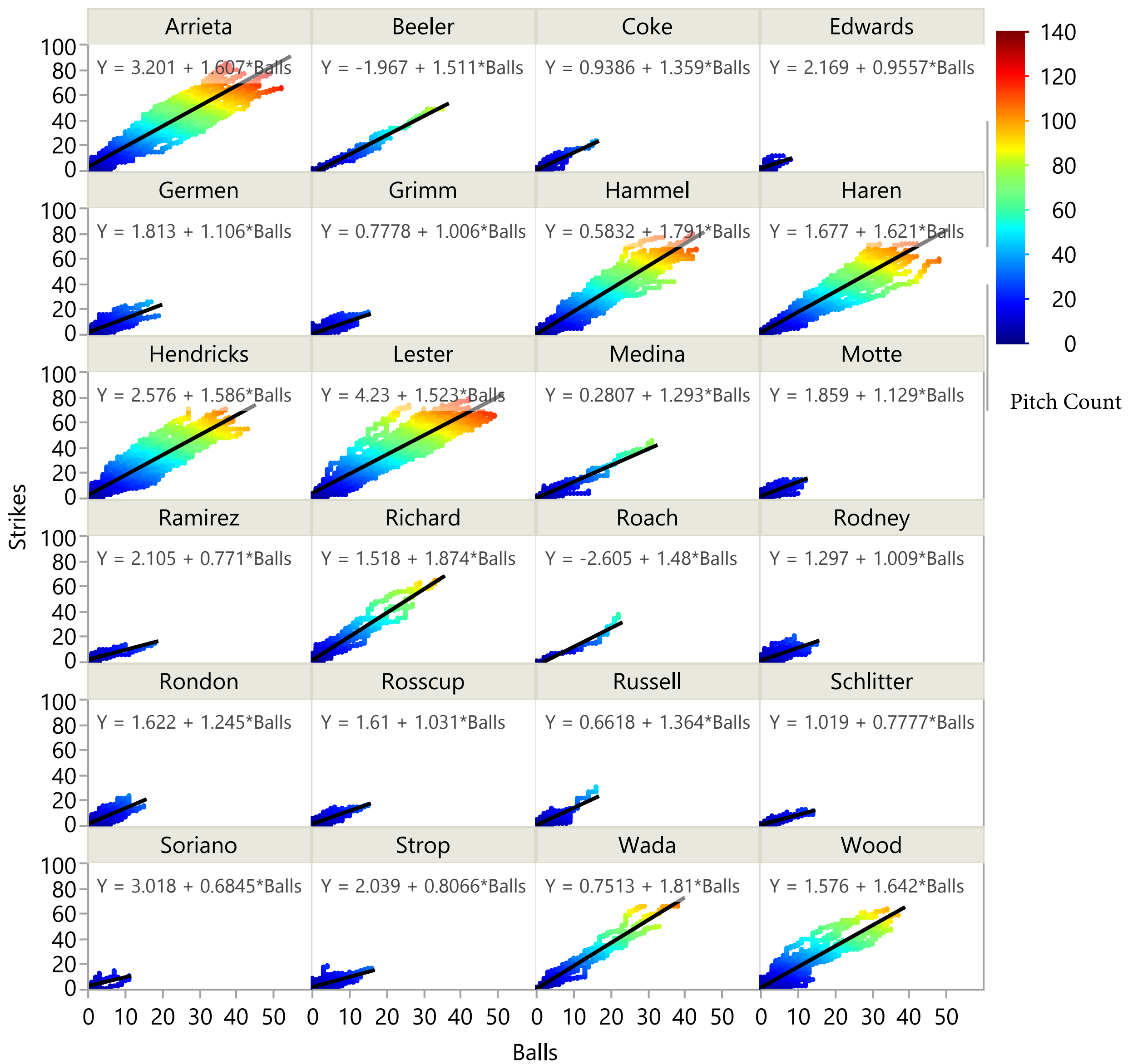

Figure 5: Plot of strikes versus balls for all pitching sequences during the 2015 MLB Season for the Chicago Cubs (colored by pitch count).

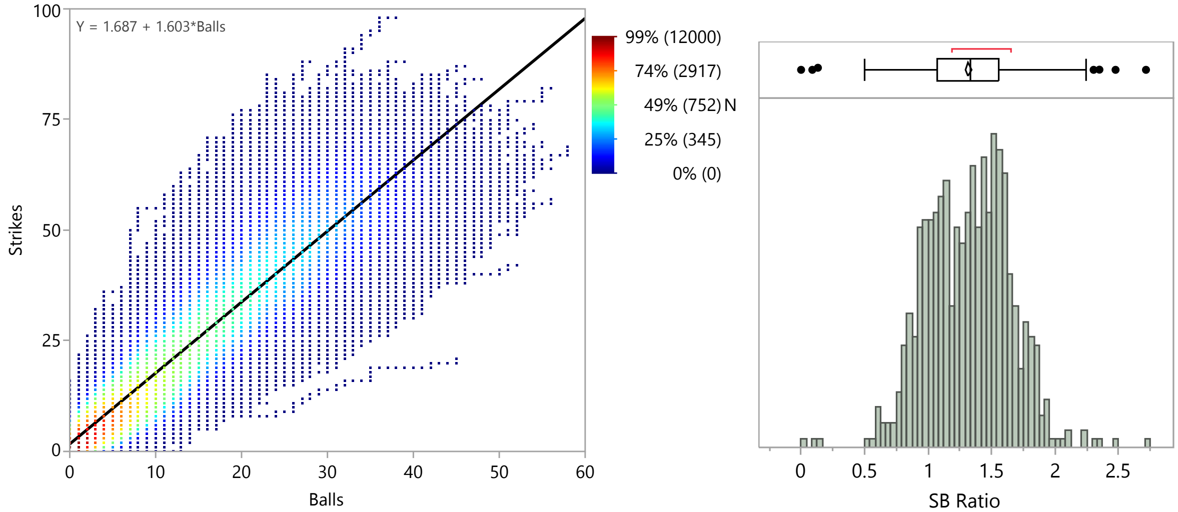

Figure 6: LEFT: Plot of strikes versus balls for all pitching sequences during the 2015 MLB Season. RIGHT: The distribution of strike-to-ball ratios for all 712 pitchers.

counts. For example, in all 33 starts by Arrieta, he reached at least 70 pitches. However, in fewer and fewer games he reaches 100 or 110 pitches. Below the bar chart is a mean strike-to-ball ratio for the 33 games at a particular pitch count. In other words, this line represents the typical strike-to-ball ratio expected with respect to Arrieta's pitch count. Interestingly, there seems to be a small strike-to-ball

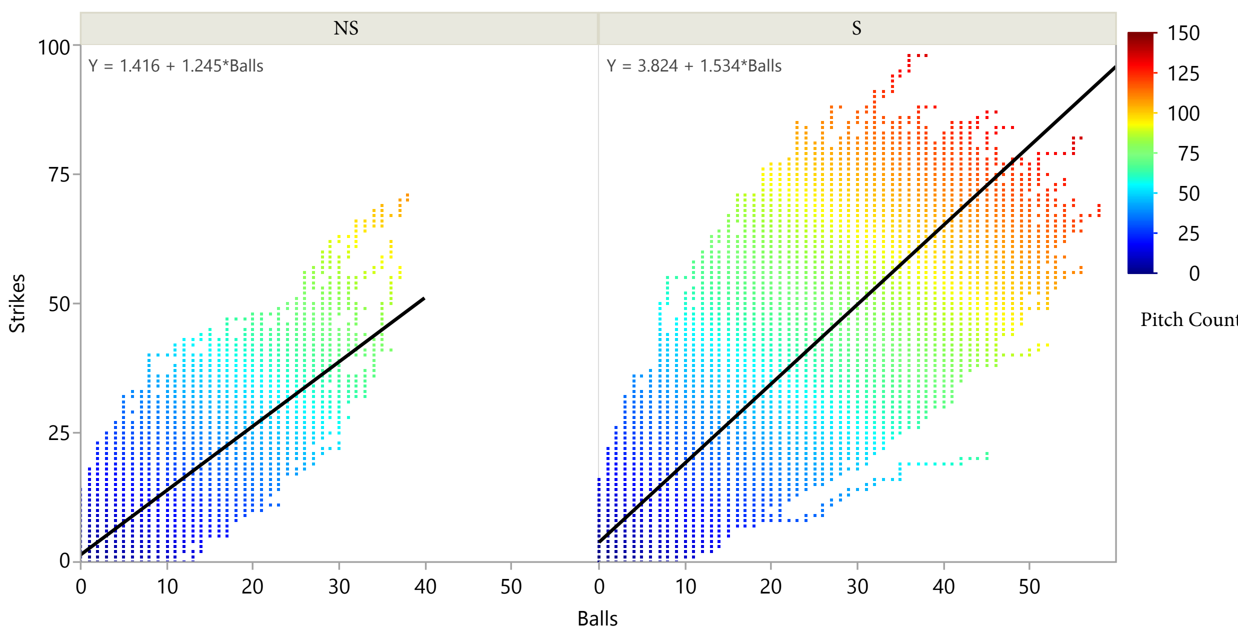

Figure 7: Plot of strikes versus balls for all pitching sequences during the 2015 MLB Season. Pitchers who started (S) at least once are shown collectively on the right, and pitchers who never started (NS) are shown collectively on the left.

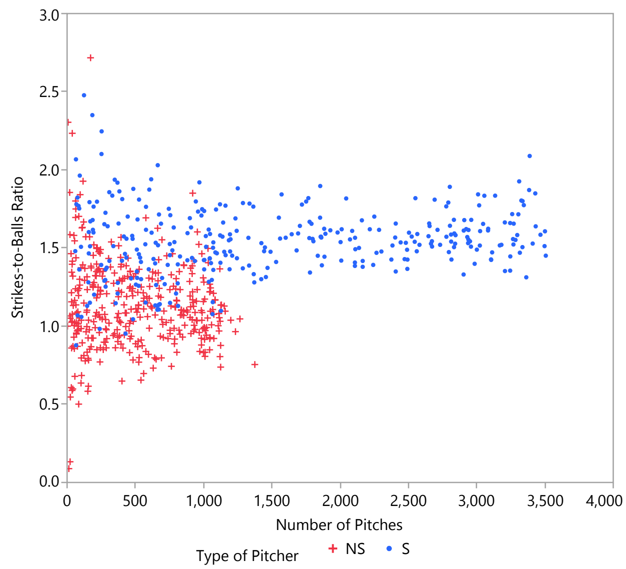

Figure 8: Plot of 712 pitchers' strikes-to-ball ratio versus the total number of pitchers for the entire 2015 MLB season. Pitchers who started (S) at least once are shown in blue, while pitchers who never started (NS) are shown in red.

ratio peak early in his games before he settles in to his average performance around 1.8. Note that this value is different than the STB coefficient using the best fit line described previously.

In this case, we take the average of the strike-to-ball ratio at each pitch count value (i.e. 33 data points up until 72 total pitches, afterwhich the number of data points decreases). This means that near the end or tail of the distribution, lower confidence is expected in the results due to less data. However, before that point is reached some interesting observations can be made. At relatively high pitch counts, a local drop in the mean strike-to-ball ratio line is observed. This is most evident in Lester's STB ratio line around pitch number 100, but can also be seen in Hendricks's around pitch 80 and twice for Hammel's near pitch 65 and pitch 90. Arrieta's drops less so at pitch 100 and Haren's does too, although with more uncertainty in the latter case, due to less data.

These changes in performance can be used as trigger points to evaluate if a pitcher is tired or reaching his limit in other ways and perhaps needs to be pulled. Ideally, the signs should be recognized before an undesirable performance makes it obvious and therefore analyzing when and how each pitcher reveals his fatigue in such profiles presented here can be fruitful. Since relief pitchers will play in more games than a starter in a give season (e.g., 68 for Rodney), uncertainty should be lower, but over a shorter range of pitch counts. Still, the strike-to-ball ratio line can be used to ascertain when the pitcher is deviating from his expected performance, and modifications can be made if needed.

Lastly, analyzing the non-starting pitchers can also provide insight into when and how they should be used. For example. Rondon's STB line peak is higher than most starters but he quickly loses steam and falls back to an average starter after 10 or 15 pitches. However, falling "back to average" is still an attractive proposition if he can maintain that rate for more pitches. Although the data is sparse, there seems to be some evidence that he does not continue to fall since the line turns back up around 25 pitches.

Rosscup and Strop, on the other hand, have a similar peak but continue their burnout phase after only a few pitches. The others offer some indication that they could be used or trained to become more

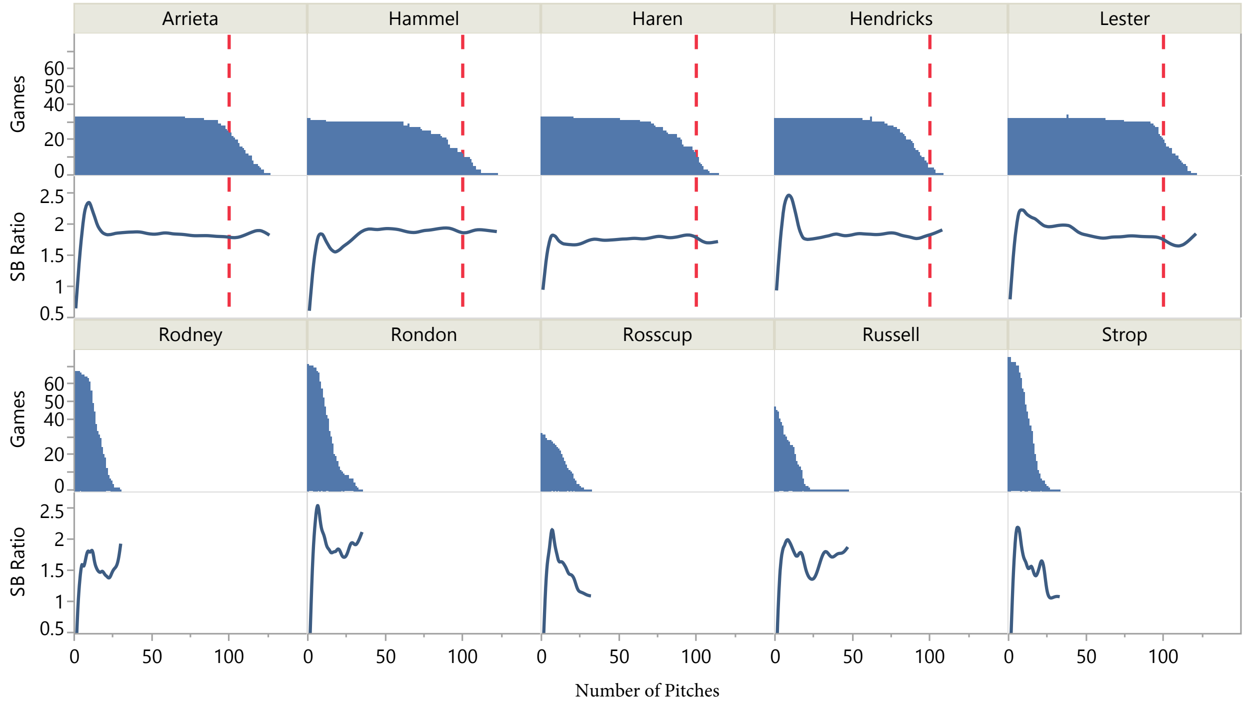

Figure 9: Chicago Cubs' pitchers with the five key starters on the top row and five relievers on the bottom row. The top half of each subgraph shows the number of games during which the pitcher reaches a certain pitch count. The bottom half indicates the mean strike-to-ball ratio for that pitcher at different pitch counts.

than just 10-or-20-pitch pitchers and fill that gap in the bullpen's collective capability.

4. Conclusion

By way of conclusion, we remark that our finding regarding the magnitude of the home-field advantage (0.429 runs/game rather than 0.133 runs/game) is significant, and is misrepresented in a number of works [13, 16]. It is now known that the first inning makes up approximately 21% of the overall difference in scoring potential between the home and visiting team, rather than 58% as was reported in [16]. A clear understanding of this principle prevents us from overestimating the importance of the first inning (or any other source), although it remains roughly twice as important as any other inning in terms of the home-field advantage.

The remainder of this section is used to succinctly list and summarize starting pitching and bullpen strategies that arise from the findings of this paper, and propose additional ones for future study.

4.1. Summary of Strategies in this Work

First-Batter Starters: Assuming the first inning bias really is largely a function of how long a starting pitcher sits in the dugout between warming up and pitching in a ball game, the strategy of using a relief pitcher to face only the first batter of the game for visiting teams allows the usual starter to enter the game directly from the bullpen, just as fresh as he would likely be when pitching at home. The analysis of this paper shows a reduction in the first inning home field advantage of 0.037 runs/game, or approximately 3 less runs scored by home opponents over 81 road games in a season when deploying this strategy. Since roughly 18% of games are won by home teams by only one run, this could easily provide one or more additional road wins to a team in a season. Furthermore, since the strategy does not require any extra skills or work by a team, it is an advantage that comes for free to anyone bold enough to use it.

Pitch Until the Trigger Point: As demonstrated in the figures of the previous section, the different pitchers have characteristic triggers and "tells" regarding their state or condition. A pitching management style that reacts to the plethora of data available may be most optimal but also challenging to implement due to the various parameters and factors intrinsic in such a strategy. However, with advances in technology, including the areas of processing capability and computer vision, the potential to evaluate accurately and confidently process the multi-variable responses to all the signs, trigger points, and tells is closer to reality. Knowing exactly the conditions, both mental and physiological, of one's player on the mound can only improve decision making in this regard.

4.2. Pitching Strategies for Future Work

Closers as Starters: An extension of the first-batter starter tactic is to use a relief pitcher for the entire first inning as proposed in [5], and allow the usual starter to begin the game in the second inning. Although this technique does not focus directly on reducing the first-inning bias toward the home team, it may provide other benefits to a ball club. Some preliminary analysis of this strategy has indicated that relief pitchers generally pitching the 8th inning would likely be at least as effective as starters in pitching only the first inning. This may allow teams to match up their ace reliever against the top of the opposing lineup, and reduce the overall scoring in the first inning of their opponents, whether at home or on the road. Furthermore, it has been shown [10] that hitters tend to do better the third time facing a pitcher in a game. Since the top of the order is likely to be comprised of an opposing team's best hitters, this strategy prevents the best hitters from seeing the second-inning starter three times until much later in the game (and maybe not at all). While managers could adopt lineup changes to combat this idea, forcing the best opposing hitters down the lineup is always a good result as they will tend to bat less than those at the top of the lineup.

Three-Inning Pitchers: The data in Section 3 seems to indicate that starting pitchers are likely to be able to deliver a greater number of quality pitches in one outing than relief pitchers. This may be, at least partially, a function of relief pitchers intentionally throwing harder for a shorter number of pitches. In other words, the task of relief pitching is different from that of starting pitching, just as running a marathon is different from running a 100-meter dash. We propose a middle ground of grooming pitchers to pitch exactly three innings in an outing. Overall pitch counts for the starters would be reduced and this may allow for them to start more often. Regrettably, no MLB team has ever adopted such a strategy to our knowledge, but it has been shown recently that relief pitchers are now pitching more innings on average, and have a lower OPS+ for opposing batters compared to starting pitchers [7]. Maybe it's time to level the playing field between these two groups in the form of three-inning outings.

Pitch Once Through the Lineup: Taking the idea of the three-inning pitcher one step further, and motivated by the results in [9, 10] where we see that opposing batters perform better after seeing a pitcher multiple times in one outing, we propose the simple strategy of never allowing a batter on the opposing team to see any pitcher more than once per game. Thus, a pitcher would start with the leadoff batter in the first inning, and come out after facing the 9th batter in the lineup. In 2015, MLB teams had an average of 38 plate appearances in any one game; and therefore, this would necessitate 4--5 pitchers per side for every game. Surprisingly, this number is not all that different from the current trend of roughly 4 pitchers per game [7].

References

[1] Retrosheet. URL http://www.retrosheet.org/.

[2] Rob Arthur. In baseball, October is reliever season. *FiveThirtyEight and ESPN*, Oct. 2016. URL http://fivethirtyeight.com/features/in-baseball-october-is-reliever-season/.

[3] Phil Birnbaum. A guide to sabermetric research. *Society for American Baseball Research (SABR)*. URL http://sabr.org/sabermetrics.

[4] Phil Birnbaum. Examining home-field advantage. In *SABR Analytics Conference (SABR 41)*, Long Beach, CA, Jul. 2011. URL http://www.philbirnbaum.com/.

[5] Bryan Grosnick. Replacing setup men with "openers". *Beyond the Box Score*, Nov. 2013. URL http://www.beyondtheboxscore.com/2013/11/26/5144934/openers-bullpen-usage-closers-setup-men-weird-baseball-all-the-luke-hochevar.

[6] Bill James. *The Bill James Baseball Abstract*. Self-published by Bill James, Mar. 1982.

[7] Jonah Keri and Neil Paine. How bullpens took over modern baseball. *FiveThirtyEight and ESPN*, Aug. 2014. URL http://fivethirtyeight.com/features/how-bullpens-took-over-modern-baseball/.

[8] Michael Lewis. *Moneyball: The Art of Winning an Unfair Game*. WW Norton & Company, 2004.

[9] Mitchel Lichtman. Baseball ProGUESTus: Why having a quick hook helps. *Baseball Prospectus*, Nov. 2011. URL http://www.baseballprospectus.com/article.php?articleid=15549.

[10] Mitchel Lichtman. Baseball ProGUESTus: Everything you always wanted to know about the Times Through the Order Penalty. *Baseball Prospectus*, Nov. 2013. URL http://www.baseballprospectus.com/article.php?articleid=22156.

[11] Sam Miller. Are statheads responsible for the most exciting postseason in years? *ESPN*, Oct. 2016. URL http://www.espn.com/mlb/story/_/id/17870355/are-statheads-responsible-most-exciting-postseason-years.

[12] MLB.com. Guidelines for youth and adolescent pitchers. URL http://m.mlb.com/pitchsmart/pitching-guidelines.

[13] Tobias Moskowitz and L. Jon Wertheim. *Scorecasting: The hidden influences behind how sports are played and games are won*. Three Rivers Press, 2012.

[14] Official Playing Rules Committee. *Official Baseball Rules: 2016 Edition*. Office of the Commissioner of Baseball, 2016. URL http://mlb.mlb.com/mlb/downloads/y2016/official_baseball_rules.pdf.

[15] John L. Salmon and Willie K. Harrison. Tracking pitcher performance with instantaneous component ERA and moving averages. In *MIT Sloan Sports Analytics Conference*, Boston, MA, Mar. 2016. URL http://www.sloansportsconference.com/wp-content/uploads/2016/02/1564-Tracking-Pitcher-Performance-with-Instantaneous-Component-ERA-and-Moving-Averages.pdf.

[16] David W. Smith. Why do home teams score so much in the first inning? *Retrosheet Research Papers (SABR44)*, Aug. 2014. URL http://retrosheet.org/Research/SmithD/WhyDoHomeTeamsScoreSoMuchInTheFirstInning.pdf.

[17] David W. Smith. Home team scoring advantage in the first inning largely due to time. *Retrosheet Research Papers (SABR45)*, June 2015. URL http://retrosheet.org/Research/SmithD/HomeTeamScoringAdvantageRelatedToTime.pdf.

[18] Jayson Stark. The book on hooks. *ESPN*, May 2004. URL http://www.espn.com/mlb/columns/story?id=1798057columnist=stark_jayson.

[19] Jeff Zimmerman. First inning home field advantage. *FanGraphs*, Sep. 2013. URL http://www.fangraphs.com/blogs/first-inning-home-field-advantage/.