Quantifying Hitter Plate Discipline in Major League Baseball

January 11, 2022

Abstract

We use Statcast data from the 2017-2021 Major League Baseball seasons to quantify the ability of players to make correct decisions about whether or not to swing at each pitch. Using player hitting ability and pitch characteristics, we use an xgboost model to predict the likelihood of each possible outcome from a swing, and from not swinging. From each outcome, we calculate the expected runs scored in the inning. Combining these, we get the expected runs if the player swings, and the expected runs if the player takes the pitch. We quantify a hitter's plate discipline by averaging the total increase in expected runs from each pitch thrown to the hitter. We show that this metric is stable over time, explains current year hitting performance, and can improve predictions of future hitting performance.

1   Introduction

Plate discipline, the ability of a hitter to select the correct pitches to swing, at is widely acknowledged as an important part of batter performance and evaluation. Players have earned reputations for being particularly bad or good at choosing which pitches to swing at. For example, Vlad Guerrero is widely recognized as being particularly free-swinging, and having a poor eye [1, 2]. However, the statistics used to measure plate discipline are relatively unsophisticated, the percent of pitches outside the zone (O-Swing %) and in the zone (Z-Swing %) that are swung at, being the most commonly used measurements. Other measurements include strikeout rate, walk rate, or strikeout to walk ratio. All of these measurements are flawed for a variety of reasons: they ignore game situation, quality of pitches seen, pitch type, and/or hitter ability. Further, they are season aggregates, and do not directly measure the decision of the hitter, whether or not to swing at each pitch.

Plate discipline, as measured by walk rate and strikeout rate, has been shown to be an important determinant of hitting performance, as measured by slugging percentage, on base percentage, walk, and strikeout rate [3]. Improving plate discipline, as measured by O% and Z% could add 68 points (0.068) to the average hitter's On-base Plus Slugging (OPS) [4]. "We find that 33/50 metrics demonstrate signal. However, these metrics are highly correlated with one another and related to traditional notions of performance (e.g., plate discipline, power, and ability to make contact)." - plate discipline is predictive, but "plate discipline" is O% and Z% and walks and strikeouts.

Further, measuring literal eye-quality, Laby et al. found that hitters with better visual function are more likely to be successful when batting, and were more productive hitters [5]. This is because having better visual function helped hitters determine which pitches to swing at, but was not correlated with contact rates [6]. In other words, having better vision abilities helped hitters choose which pitches to swing at, but didn't impact how well they actually hit them: the biological eye, and the baseball "eye" are related. Over the season, eye-quality declines, probably due to fatigue, an effect that was pronounced after stimulants were banned in 2005 [7].

As a whole, batters' eyesight has been shown to impact plate discipline, and plate discipline is an important predictor of hitter performance. However the metrics used to assess plate discipline are crude and unsophisticated; the goal of a hitter is not to simply swing at strikes and take balls, but rather to maximize the (expected) number of runs his team will score. In this paper, we develop a plate discipline statistic called EAGLE that accounts for pitch location further than simply in/out of zone, game situation, and batter ability. EAGLE stands for Expected Additional runs Gained by Looking/swinging Estimate and it estimates the expected runs the team will score if the batter were to swing, and if the batter were to take the pitch, and measures the average increase in expected runs due to batter decision-making. We show that this metric is a better measure of plate discipline than the current measurements in terms of stability, explanation of performance, and prediction of future performance.

2 Methods

The goal of this paper is to predict the change in run expectancy for each decision that a hitter makes. This means measuring the predicted runs added or lost by each swinging or taking decision.

Inspired by the expected points added metric in football [8], this metric, EAGLE, would account for situation and measure the predicted change in the scoring total, in this case, runs. In order to do this, we construct a decision tree for the decision faced by Major League hitters.
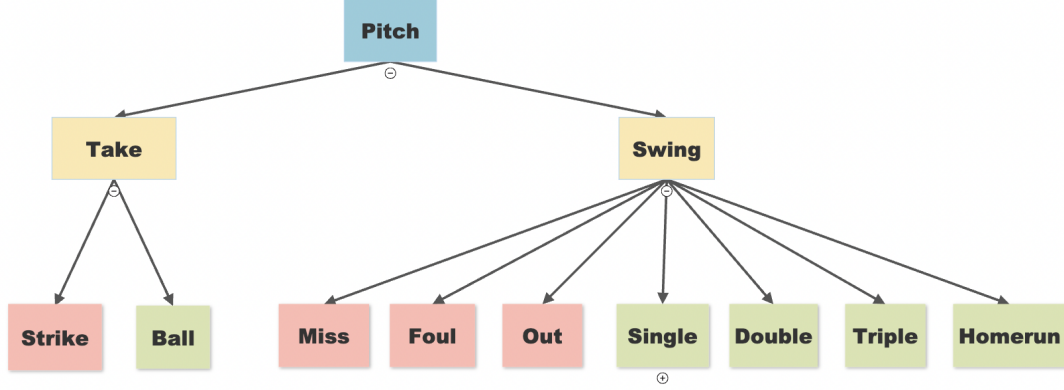

Figure 1: The decision analysis tree for the optimal swing/take decision.

The analysis is modeled by the tree shown in Figure 1. The decision of take or swing is made by the hitter. When taking a pitch, there are only two outcomes: ball or strike. When swinging, however, there are many different outcomes which we distilled to seven fundamental results: miss, foul, out, single, double, triple, and home run. To determine the value that a player contributes with each decision, we used change in run expectancy. For each pitch, we take the specific combination of count, number of outs, and runners on base and estimate the expected runs scored in the rest of the inning from that pitch on. These estimates are drawn from the baseballR package [9] and stored as the original run expectancy. We then adjust the run expectancy for each hitters based on their hitting ability (as measured by OPS). The current run expectancy with a good hitter at the plate is higher than with a bad hitter at the plate. Ignoring this will lead to EAGLE penalizing good hitters for taking borderline pitches, because the actual run expectancy of a take for a good hitter is higher than the league average run expectancy.

In order to measure the effect that the hitter had in that situation based on their decision, we compute the change in run expectancy from before the pitch to after the swing/take decision has been made.

In our approach to combining the decision tree with this change in run expectancy variable, we must understand the effect on run expectancy for of the nine outcomes we specified. We calculate

3

the change in run expectancy for each pitch if each of the nine outcomes occur. To calculate the run expectancy for each decision, we need the probability each outcome will occur, and the expected runs if it does. We predict the probability a pitch will be called a strike or a ball if the hitter does not swing, and calculate the expected runs for each of those outcomes. Combining those gives us the expected runs in the inning if the hitter takes the pitch. We also need to estimate the probability of each of the seven outcomes if a hitter were to swing at each pitch. We combine those estimates with the expected runs for each outcome to calculate the expected runs if the hitter were to swing. Below, we discuss how the probabilities are estimated.

We use the baseballr R package to scrape Statcast pitch location and outcome data from the 2019-2020 seasons. We scraped all of the hitter and pitcher data, and merged them together so that pitcher information and hitter information were in the same dataset. We also scraped basic player season stats for both the pitcher and the hitter and merged them to each pitch in order to have descriptive features representing hitter ability.

For the model to predict strike probability, the best model simply used the x and z coordinates of the baseball when it crossed the plate and the top and bottom of the strikezone for the particular batter. The classification model for swing outcomes was much more complex, due to the number of outcomes and the randomness of the outcome. To predict the outcome of a swing, we used variables containing information on the pitch, the game situation, pitcher ability, and hitter ability. Pitch information included the x and z coordinates when crossing the plate, horizontal and vertical movement, the pitch type, velocity, effective velocity, and spin rate. Situational variables included count, outs, and the pitcher and batter handedness. We used basic statistics from the 2019-2020 seasons including Plate Appearances, Batting Average, On-Base Percentage, and Slugging Percentage to measure batter ability, and Innings Pitched, Earned Run Average, Groundball to Flyball ratio, Walks plus Hits per Innings Pitched, Strikeout percentage, and Walk Percentage for Pitchers. To place an additional level of depth on the batter skill level, we grouped several statistics by the Statcast zone in which the ball was thrown including Barrel Percentage, Exit Velocity, Launch Angle, Expected Batting Average, Expected Weighted On-Base Average, Weighted On-Base Average, and Isolated Power for batters and Overall Velocity, and Overall Spin Rate for Pitchers.

In order to predict the probability of a strike, we created a logistic model that uses the x and
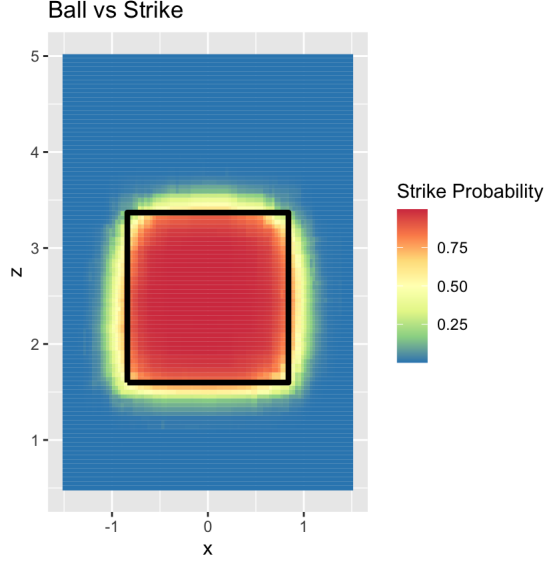
Figure 2: Estimated strike zone for a hitter with league average top and bottom of strike zone.

z coordinates of the baseball as it crosses the plate. However, since the strike zone as called by umpires is not purely linear, we used an xgboost with a logistic objective in order to capture this nonlinearity. In the end, the independent variables for the model were x, z, and variables for the top and bottom of the strikezone. The model has an AUC of 0.982, and the estimates of the strike zone for a league-average height hitter are shown in Figure 2.

To predict outcome probabilities we trained an xgboost classification model. It was tuned using cross validation to determine the optimal meta-parameters for the model. In this way we were able to reduce the mean squared error and end up with the resulting confusion matrix shown in Table 1 of probabilities, the true outcomes being the columns and the average predicted probabilities given in the rows. For example, for each pitch which was actually hit for a double, we predicted a 16.0% chance of a swing and miss, a 41.2% chance of a foul ball, and a 4.2% chance of a double (among others). For missed pitches, we estimated a 40.98% chance of a miss, a 32.8% chance of a foul, and only a 1.06% chance of a home run.

To combine the outputs into a single comprehensive metric, we multiply each outcome probability by the change in run expectancy for that outcome. We add the results together and do this

Table 1: Outcome Probabilities

Outcome	Miss	Foul	Out	Single	Double	Triple	Homerun
Miss	0.4097761	0.3283625	0.1814942	0.05319165	0.01491320	0.001579695	0.01068269
Foul	0.2170178	0.4213399	0.2429398	0.07507962	0.02340903	0.002355106	0.01785872
Out	0.2269750	0.3890573	0.2651662	0.07840765	0.02280722	0.002229124	0.01535743
Single	0.1726865	0.3956535	0.2718372	0.11046822	0.02891459	0.002696567	0.01774349
Double	0.1602147	0.4066227	0.2662421	0.09320279	0.04228762	0.003741214	0.02768888
Triple	0.1524018	0.4120729	0.2667650	0.09066951	0.03970592	0.007907397	0.03047746
Homerun	0.1638953	0.4198078	0.2471371	0.07638592	0.03659626	0.003848514	0.05232918

for both the take outcomes and the swing outcomes as in the formula below.

$$RE(Take) = \sum P(Call) * RE(Call)$$

$$RE(Swing) = \sum P(Outcome) * RE(Outcome)$$

We then end up with a run expectancy for each decision, for swing and for take. We judge the correct decision as the change in run expectancy that is the larger of the two and the player's EAGLE for that pitch is the difference between the decision taken and the other choice. This is the expected Runs Gained metric defined here. Then EAGLE is just the expected Runs Gained based on the decision made.

$$xRG(Swing) = RE(Swing) - RE(Take)$$

$$xRG(Take) = RE(Take) - RE(Swing)$$

$$EAGLE = xRG(Decision)$$

As an example, for a given pitch if the estimated expected runs gained from swinging is 0.2 and the estimate of expected runs gained from taking is -0.05, the EAGLE for swing would be 0.25, and the EAGLE for a take would be -0.25. We then average these values over each hitter over all pitches seen for the season. The resulting average is the player's EAGLE. This gives the average expected runs gained (or lost) by the swing/take decisions for each player. Importantly,
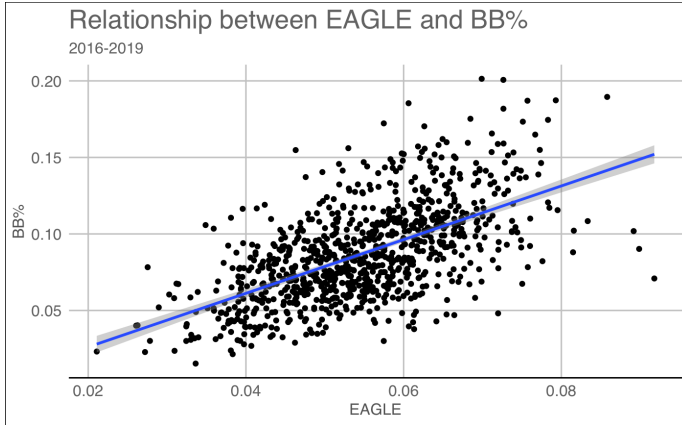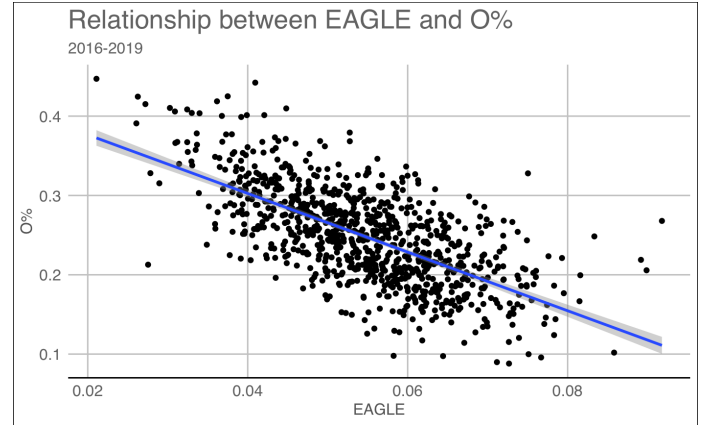
Figure 3: EAGLE vs. Walk rate.

Figure 4: EAGLE vs. O%.

EAGLE does not measure the *actual* outcomes of the swing - EAGLE does not measure hitting ability, simply the ability to swing at the best pitches.

3   Results

As a measure of plate discipline, EAGLE tells us how good a player is at deciding which pitches to swing at and which pitches to take. We expect it should correlate with commonly used statistics of plate discipline and hitting performance. Figure 3 shows the relationship of hitters' EAGLE statistic with their walk rate, with a correlation coefficient of 0.563. EAGLE also incorporates how good a players chase rate is (measured as percent of pitches out of the strike zone swung at, called O%), shown in Figure 4. EAGLE and O% have a correlation of 0.683. The hitters who have the best EAGLE and the worst EAGLE for the years 2016-2019 are also listed in Tables 4 and 3.

An important feature of a player-level evaluation metric is its stability - if a new statistic is measuring an ability inherent to a player, it should be relatively stable from year to year, whereas if the metric is driven by chance or other factors outside of the player's control, the statistic will vary from year to year [10]. We show in Figure 5 that EAGLE is stable from year to year, as a player's plate decision making is a skill developed over time and rarely changes drastically.

One of the most important parts of the EAGLE metric is its ability to predict OPS in the current year, and in the future. We feel this is a strong contribution of our metric - we don't just

Table 2: Best Players				
Name	Year	EAGLE	BB%	O%
Juan Soto	2019	0.088	0.164	0.177
J.D. Martinez	2017	0.084	0.108	0.276
Khris Davis	2018	0.083	0.090	0.235
Joey Votto	2017	0.082	0.190	0.126
George Springer	2019	0.081	0.121	0.217

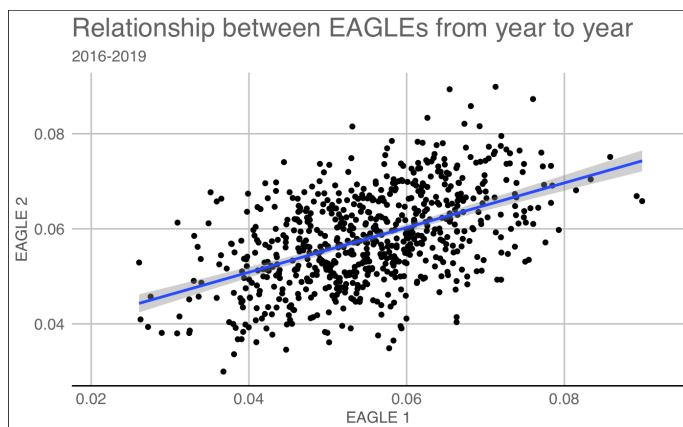| Table 3: Worst Players				
Name	Year	EAGLE	BB%	O%
Jose Iglesias	2016	0.023	0.054	0.372
Harold Castro	2019	0.023	0.024	0.427
Erick Aybar	2016	0.024	0.067	0.406
Ryan Zimmerman	2016	0.027	0.062	0.316
Jose Iglesias	2019	0.028	0.038	0.448

Figure 5: EAGLE from 2019 vs. EAGLE in 2020 for the each player

measure strikeouts and walks, but quantify the ability to pick pitches which can be hit for power. To our knowledge, there is no other work that account for plate discipline in relation to OPS. Figure 6 shows the relationship between EAGLE and a hitter's OPS, showing they are strongly correlated. We would expect a hitter in the 10th percentile of EAGLE to have an OPS of .705 while we would expect a hitter in the 90th percentile of EAGLE to have an OPS of .808. These numbers for predicting next year's OPS are for the 10th percentile .739 and for the 90th percentile .795. Other current statistics don't perform nearly as well. The relationship between EAGLE and OPS has a correlation coefficient of .461 while that of O% and OPS is only -.104. No other stats can match the power of EAGLE for predicting both discipline statistics and slugging statistics.

In addition to EAGLE's use with hitters, EAGLE can be used to tell how deceptive a pitcher is. We aggregated the EAGLE on each pitch thrown by pitchers from 2016-2019 and came out with the following results for the pitchers that are the best at causing hitters to make bad decisions.
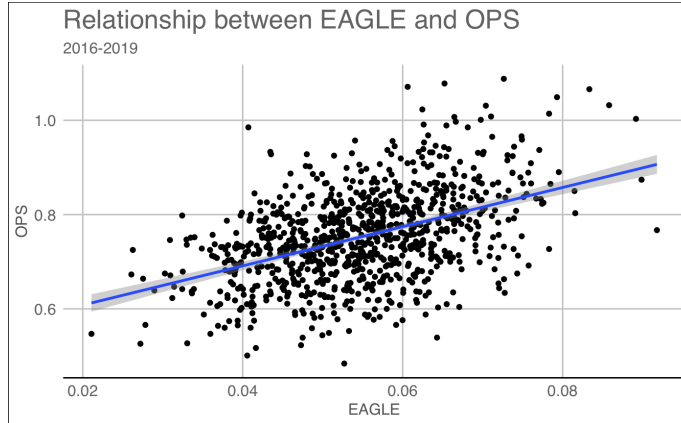
Figure 6: EAGLE vs. On-Base Percentage plus Slugging (OPS)

This "inverse EAGLE" correlates well with ERA, O%, BB% and K% as well as opponent slugging, making it an effective way to evaluate pitchers.

Table 4: Best Pitchers

Name	Year	EAGLE	IP	ERA
Andrew Miller	2016	0.022	74.1	1.45
Luke Gregerson	2016	0.024	57.2	3.28
Giovanny Gallegos	2019	0.028	74.0	2.31
Sean Doolittle	2017	0.032	51.1	2.81
Roberto Osuna	2017	0.032	64.0	3.38

4   Conclusions and Applications

In this work, we introduce EAGLE, an estimate of MLB hitter decision-making at the pitch-level. Using publicly available data, we measure the impact of each decision to swing or take a pitch on the expected number of runs scored in the remainder of the inning. This metric is stable from year to year, and predicts hitting outcomes, particularly OPS, significantly better than other measures of plate discipline. By taking into account not just a binary "in zone" vs. "out of zone" measurement of pitch location, but rather the full set of pitch location, speed, and type characteristics, we are able to identify which hitters best select pitches that can be hit for power

9

and help create runs. The point of hitting is not simply not to strike out, but to hit the ball hard. Measures of hitters' eye quality should account for this, rather than just their ability to walk or not strike out.

The potential uses for EAGLE are vast - first as a player-evaluation tool for identifying which players have good eyes at the plate. Second, the tool can be used for player development, instead of season-aggregates, EAGLE gives pitch-level estimates, so that each decision can be analyzed, and trends can be found. Identifying which pitches a hitter swings at that he shouldn't, or vice versa, can help inform the approach taken at the plate. Similarly, EAGLE estimates can be used for scouting planning how to pitch to each hitter, what types of pitches they swing at, and which they don't. Lastly, EAGLE could be calculated for pitchers to identify which pitches they have success getting hitters to chase, and which they do not.

In terms of limitations our prediction models use full-season hitting and pitching statistics, which is fine for an explanatory model, but if we would like to use EAGLE as a predictive tool, we would need to convert those to averages over the trailing 100 games of plate appearances, or some similar measure of recent past performance.

References

[1] Mlb's five least disciplined hitters — bleacher report — latest news, videos and highlights. https://bleacherreport.com/articles/439507-the-5-least-disciplined-hitters. (Accessed on 05/05/2021).

[2] The most patient hitter in baseball — fangraphs baseball. https://blogs.fangraphs.com/the-most-patient-hitter-in-baseball/. (Accessed on 05/05/2021).

[3] Blakeley B McShane, Alexander Braunstein, James Piette, and Shane T Jensen. A hierarchical bayesian variable selection approach to major league baseball hitting metrics. *Journal of Quantitative Analysis in Sports*, 7(4), 2011.

[4] David Michael Vock and Laura Frances Boehm Vock. Estimating the effect of plate discipline using a causal inference framework: an application of the g-computation algorithm. *Journal of Quantitative Analysis in Sports*, 14(2):37–56, 2018.

[5] Daniel M Laby, David G Kirschen, Usha Govindarajulu, and Paul DeLand. The effect of visual function on the batting performance of professional baseball players. *Scientific reports*, 9(1):1–11, 2019.

[6] Sicong Liu, Frederick R Edmunds, Kyle Burris, and Lawrence Gregory Appelbaum. Visual and oculomotor abilities predict professional baseball batting performance. *International Journal of Performance Analysis in Sport*, 20(4):683–700, 2020.

[7] Scott Kutscher, Yanna Song, Lily Wang, Raghu Upender, and Beth Malow. Declining plate discipline during the major league baseball season may be the result of fatigue (p01. 260), 2013.

[8] Ronald Yurko, Samuel Ventura, and Maksim Horowitz. nflwar: a reproducible method for offensive player evaluation in football. *Journal of Quantitative Analysis in Sports*, 15(3):163–183, 2019.

[9] *baseballR: Functions for acquiring and analyzing baseball data.*

[10] Alexander M Franks, Alexander D'Amour, Daniel Cervone, and Luke Bornn. Meta-analytics: tools for understanding the statistical properties of sports metrics. *Journal of Quantitative Analysis in Sports*, 12(4):151–165, 2016.