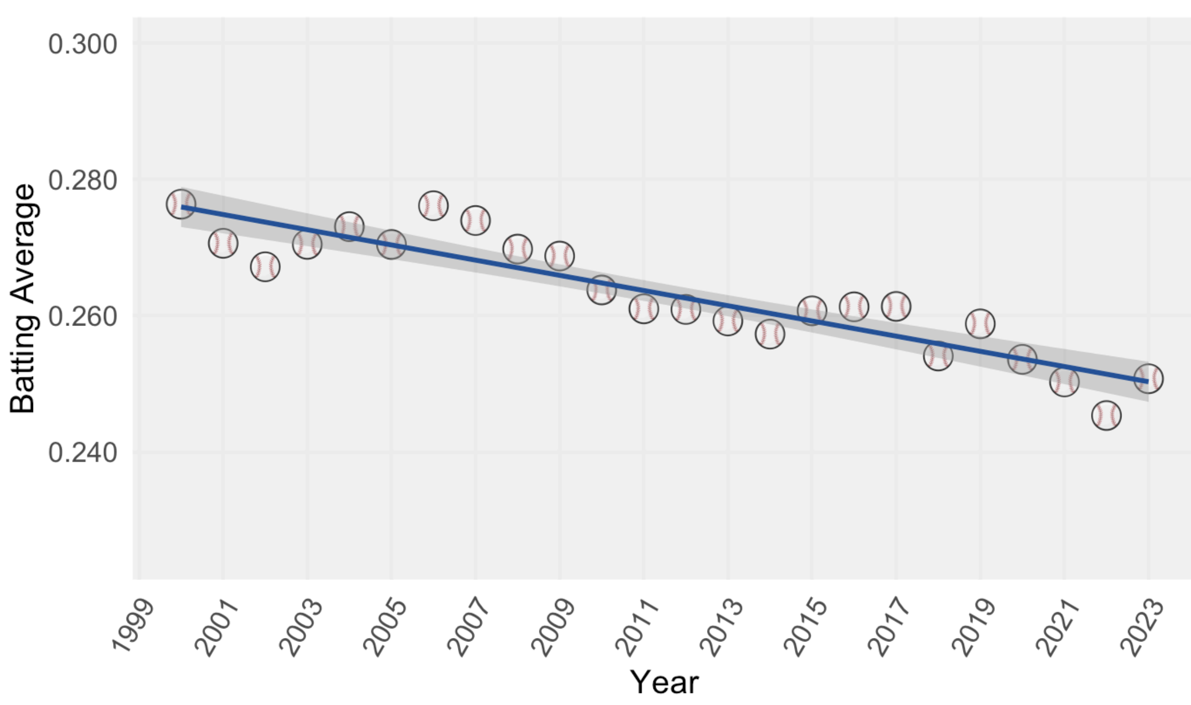
What’s Driving the Decline in Batting Averages in Major League Baseball?

Samantha Archie, Judy Cohen, Wayne Dudding, Raj Shah
April 28, 2025

Introduction

Over the past 25 years, batting averages in Major League Baseball (MLB) have been declining[1]. In 2000, the average batting average was .276 and there were 81 players whose batting average was over .300. In 2023, the average batting average was .251 and there were 16 players whose batting average was over .300.

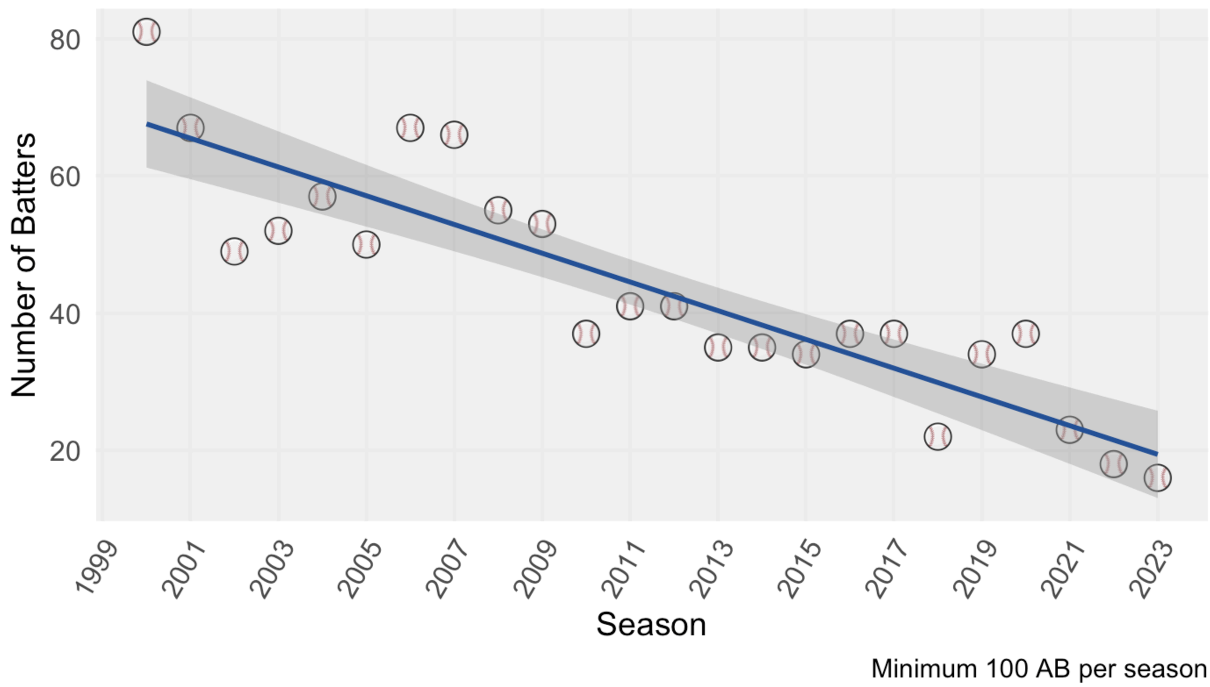
Figure 1. MLB Batting Average by Season (2000-2023)



---

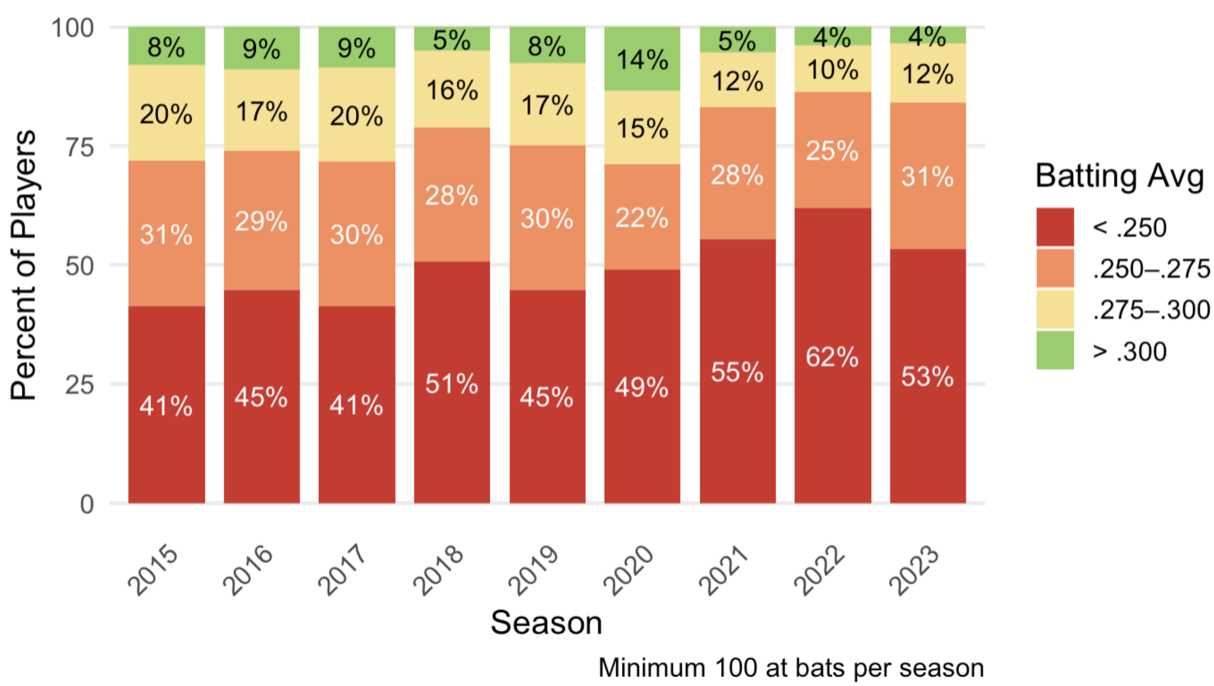
1 https://lastwordonsports.com/baseball/2024/06/04/mlb-hitting-stat-batting-average/

Figure 2. Number of Batters Hitting Over .300 by Season (2000-2023)



Over the past 10 years, the percentage of players who hit over .275 has declined by about 50%, while the percentage of players hitting below .250 has increased by about 12 percentage points.

Figure 3: Distribution of Batting Averages by Season (2015-2023)



So what's driving this decline? Some reasons mentioned in the literature[2] are:

- Defensive positioning: infield and outfield shift
- Pitching: pitchers are throwing faster pitches and more strikeouts
- Batting strategy: batters are focusing more on power hitting versus singles

Better understanding what is leading to this decline is important for several reasons:

- Lower batting averages likely means fewer hits and more strikeouts, which could impact fan attendance and viewership.
- Understanding this trend and the reasons behind it could help teams develop better game and player evaluation strategies.
- MLB may want to institute new rule changes to address this decline but what rule changes to consider depends on what factors are driving the decline.

Our analysis aimed to test these hypotheses and quantify their impact on batting average. We used data from the last 10 years to create a Bayesian multilevel model to estimate a player's likelihood of getting a hit during a given at-bat, as a stand-in for season batting average. Our results find that team and league factors (like defensive shifts) are not as important as individual hitting and pitching approaches.

Data

The data we used is from Statcast, leveraging their search batters function. To appropriately scrape it, we used R:future to use threading to gather data from multiple months, and bind all of our rows together into a larger dataset. We then filtered our data to relevant at-bat events. The way we determine this is using the official MLB Definition - "an official at-bat comes when a batter reaches base via a fielder's choice, hit or an error (not including catcher's interference) or when a batter is put out on a non-sacrifice". Using this definition, we filtered events, and then downscoped our analysis to 2015-2024, ensuring we had the relevant data we needed for all years. Some mutation was done to determine hits, and the team the batter/pitcher belonged to, and then the data was ready for our model. We only included batters with at least 100 at-bats in a particular season.

Methods

The modeling technique we ended up utilizing was a Bayesian multilevel logistic regression model to predict whether a plate appearance results in a hit. We used trends in hit probability, as that can be measured at the event level and can then be aggregated to the season level to get a batting average. We assumed the following:

---

2 https://sabr.org/journal/article/plummeting-batting-averages-are-due-to-far-more-than-infield-shifting-part-one-fielding-and-batting-strategy/

- All hit events are the same: A hit single or a home run is equivalent. This made our analysis less focused on win probabilities or hitter performance, and more focused on “is this a simple hit”.
- Hit outcomes are Bernoulli distributed: Because of the previous assumption, we are able to use a Bernoulli distribution, as all values of our random variable are 0 or 1.
- We assume independence between our observations, after accounting for the batter, pitcher, team, year, and pitch characteristics.
- We assume our random effects for batter, pitcher, team, and year are drawn from normal distributions.
- We assume our random effects are not nested within each other, and are instead random intercepts, because many top performing players moved teams across years.

The final model we used was:

$$y_i \sim \text{Bernoulli}(p_i)$$

Where:

- Y is the probability a given pitch i results in a hit

$$\text{logit}(p_i) = \eta_i$$

$$\eta_i = \beta_0 + \beta_1 \cdot \text{release\_speed}_i + \beta_2 \cdot \text{plate\_x}_i + \beta_3 \cdot \text{plate\_z}_i + \beta_4 \cdot \text{pfx\_x}_i + \beta_5 \cdot \text{pfx\_z}_i + \beta_6 \cdot \text{age\_bat}_i + \sum_{k=1}^{K} \beta_{6+k} \cdot \text{pitch\_type}_{ik} + u_{\text{batter}[i]} + u_{\text{pitcher}[i]} + u_{\text{batter\_team}[i]} + u_{\text{game\_year}[i]}$$

Where[3]:

- plate_x is the horizontal position of the ball when it crosses home plate from the catcher's perspective of a given pitch i
- plate_z is the vertical position of the ball when it crosses home plate from the catcher's perspective of a given pitch i
- pfx_x is the horizontal movement in feet from the catcher's perspective of a given pitch i
- pfx_z is the vertical movement in feet from the catcher's perspective of a given pitch i
- age_bat is the age of batter for pitch i

To implement this model, we used the brms package. We were able to specify a Bayesian-style model that would compile Stan code for us, using standard model formulations.

We considered adding random slopes to our model, but we did not have a strong motivation to compare those random slopes, and knew that computation times would increase while confidence intervals would most likely also increase.

We quantified uncertainty for our model in several ways:

[3] Statcast variable definitions: https://baseballsavant.mlb.com/csv-docs; https://www.mlb.com/glossary/pitch-types

- Our fixed effects all have uncertainty, as we have full posterior distributions for our release speed, plate positioning, spin, and age.
- Our random effects have uncertainty, both in the group-level means and in their standard deviations.

Results

After running our model, we wanted to check and see if there were any issues and if the results we got made sense.

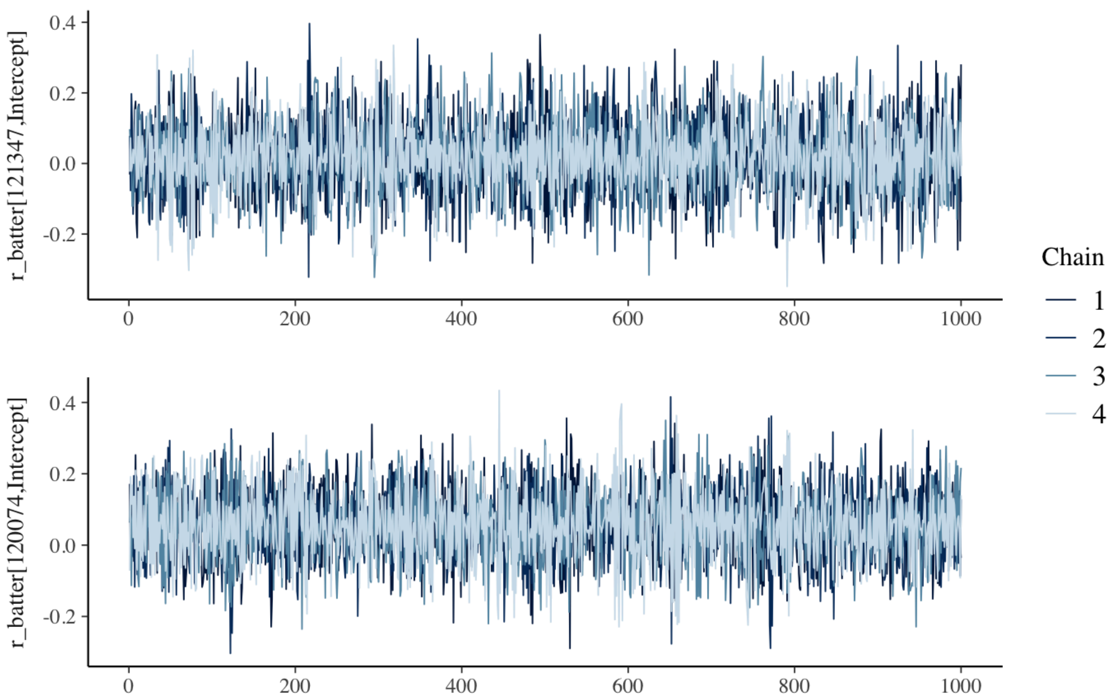
Model evaluation

Below is the output of our model. We used a very large dataset as input and ran 4 chains of 2000 iterations. The Rhat values are all around 1.00, so we know the model converged correctly and mixed well.

```
Family: bernoulli
  Links: mu = logit
Formula: hit ~ release_speed + plate_x + plate_z + pfx_x + pfx_z + age_bat + pitch_type +
(1 | batter) + (1 | pitcher) + (1 | game_year) + (1 | batter_team)
   Data: model_data (Number of observations: 242956)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Multilevel Hyperparameters:
~batter (Number of levels: 2839)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.12      0.01     0.11     0.14 1.00     1274     2277

~batter_team (Number of levels: 30)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.02      0.01     0.00     0.04 1.00      856     1188

~game_year (Number of levels: 10)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.02      0.01     0.00     0.04 1.00      886     1411

~pitcher (Number of levels: 2397)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.10      0.01     0.09     0.12 1.00     1347     2547

Regression Coefficients:
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept         1.16      0.15     0.86     1.45 1.00     3756     3555
release_speed    -0.03      0.00    -0.03    -0.03 1.00     3427     3599
plate_x          -0.07      0.01    -0.09    -0.05 1.00     5171     2986
plate_z           0.10      0.01     0.08     0.11 1.00     6501     3028
pfx_x             0.01      0.01    -0.00     0.02 1.00     4994     3397
pfx_z             0.02      0.01    -0.01     0.05 1.00     2875     2316
age_bat          -0.00      0.00    -0.01     0.00 1.00     3766     3005
pitch_typeCS     -0.36      0.42    -1.21     0.41 1.00     5778     2935
pitch_typeCU     -0.23      0.03    -0.29    -0.17 1.00     2586     2891
pitch_typeEP     -0.89      0.23    -1.36    -0.45 1.00     6360     2778
pitch_typeFA      0.52      0.18     0.15     0.87 1.00     5006     2205
pitch_typeFC      0.20      0.03     0.15     0.25 1.00     2816     2692
pitch_typeFF      0.33      0.02     0.28     0.38 1.00     2401     2626
pitch_typeFO     -0.29      0.30    -0.91     0.30 1.00     5955     2933
pitch_typeFS     -0.21      0.04    -0.30    -0.14 1.00     4601     2823
pitch_typeKC     -0.16      0.04    -0.24    -0.08 1.00     2652     1576
pitch_typeKN     -0.30      0.15    -0.59    -0.02 1.00     4305     3195
pitch_typeSC      0.02      0.50    -1.00     0.93 1.00     5581     2589
pitch_typeSI      0.47      0.02     0.42     0.51 1.00     2300     1937
pitch_typeSL     -0.11      0.02    -0.15    -0.07 1.00     2308     2499
pitch_typeST     -0.33      0.04    -0.41    -0.26 1.00     3925     2781
pitch_typeSV     -0.46      0.11    -0.67    -0.25 1.00     5004     2912
```

Below are trace plots of chain mixing for two players from our dataset, Alex Rodriguez and David Ortiz. The trace plots indicate that the chains did mix well, and there are no upward, downward, or flat trends.

Figure 4: Trace plot diagnostics



Interpretation

Looking at how everything fits together within the context of our initial question, we are able to determine the impact of each of our parameters in evaluating hit probability and thus, which players are more likely to continue hitting even if batting average declines. We account for uncertainty in our interpretations by including credible intervals to show a distribution of potential values instead of a single output.

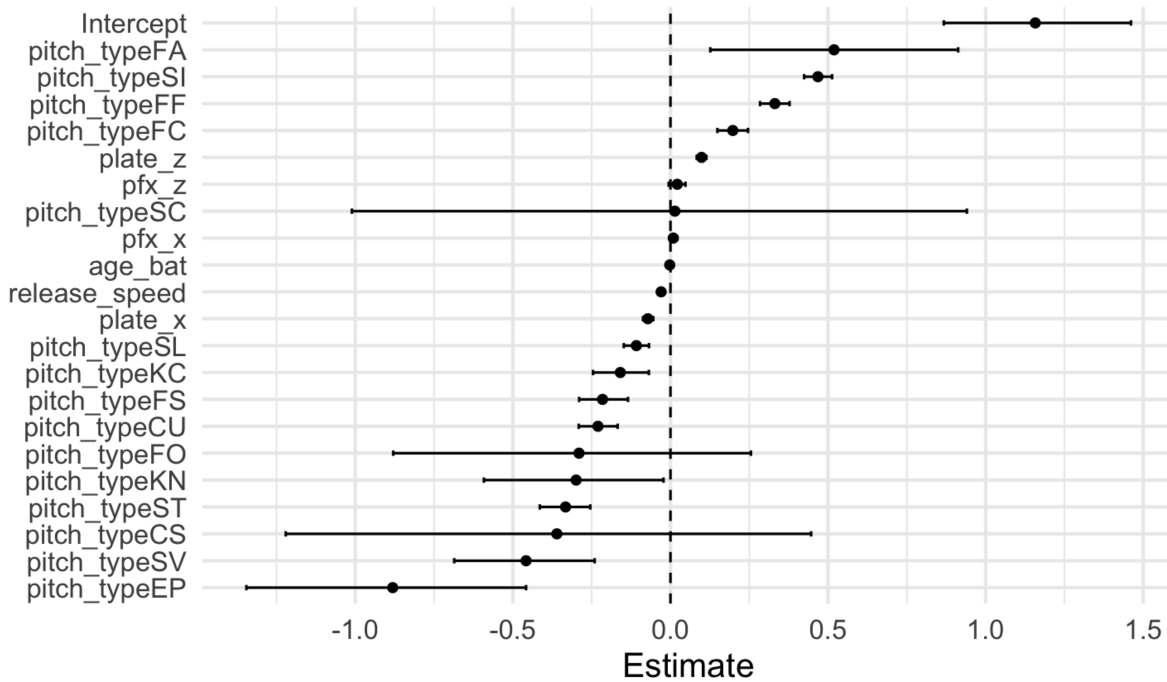
We begin by looking more closely at the coefficients of our random and fixed effects.

Fixed effects

Figure 5 shows our fixed effect uncertainties. The results suggest that certain pitch types are important contributors to hit probability. Batters facing fastballs, sinkers, and cutters are more likely to get a hit, while batters are less likely to get a hit when facing curveballs, sliders, and some rarely thrown pitches (eephus, slurve, etc.). For some rarely thrown pitches, like slow curveball, forkball, and screwball, we do not have enough data to make an informed estimate. Batter age, pitch release speed, and ball movement (pfx_x, pfx_z) appear to have minimal effect on hit probability.

The intercept term is the largest positive influence on hit probability. The intercept being the most predictive term tells us that our model is not capturing certain elements of hit probability.

Figure 5: Fixed Effects 95% Credible Intervals



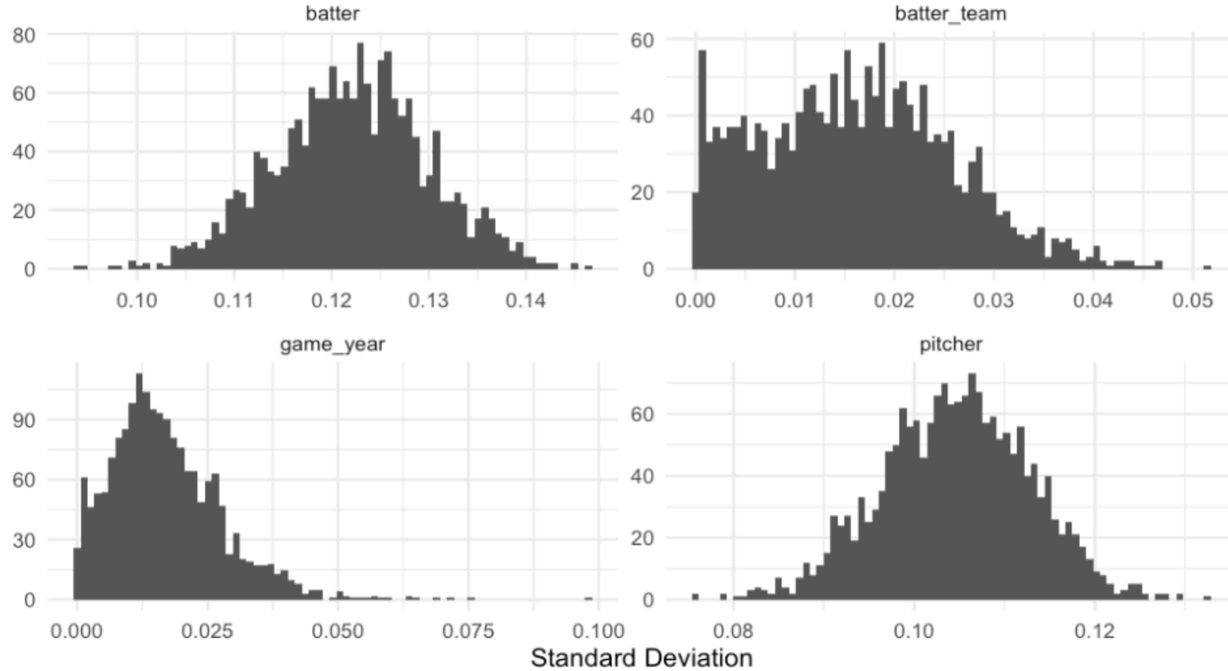
Random effects

Table 1 shows our random effect estimates, and Figure 6 displays the distributions they are drawn from. Batter, game year, and pitcher all appear to follow relatively Normal or truncated Normal distributions. The batter themself contributes most to variation in hit probability, followed by pitcher. Game year seems to have a very small effect, especially when compared to batter, showing that changes within a batter matter far less than the composition of batters on team batting averages.

Table 1: Random Effects Credible Intervals

Group <chr>	**mean_sd** <dbl>	**lower_95** <dbl>	**upper_95** <dbl>
batter	0.12200518	0.1064824840	0.13731140
batter_team	0.01609658	0.0006636082	0.03651651
game_year	0.01705795	0.0011100545	0.04106675
pitcher	0.10457011	0.0882018094	0.11989414

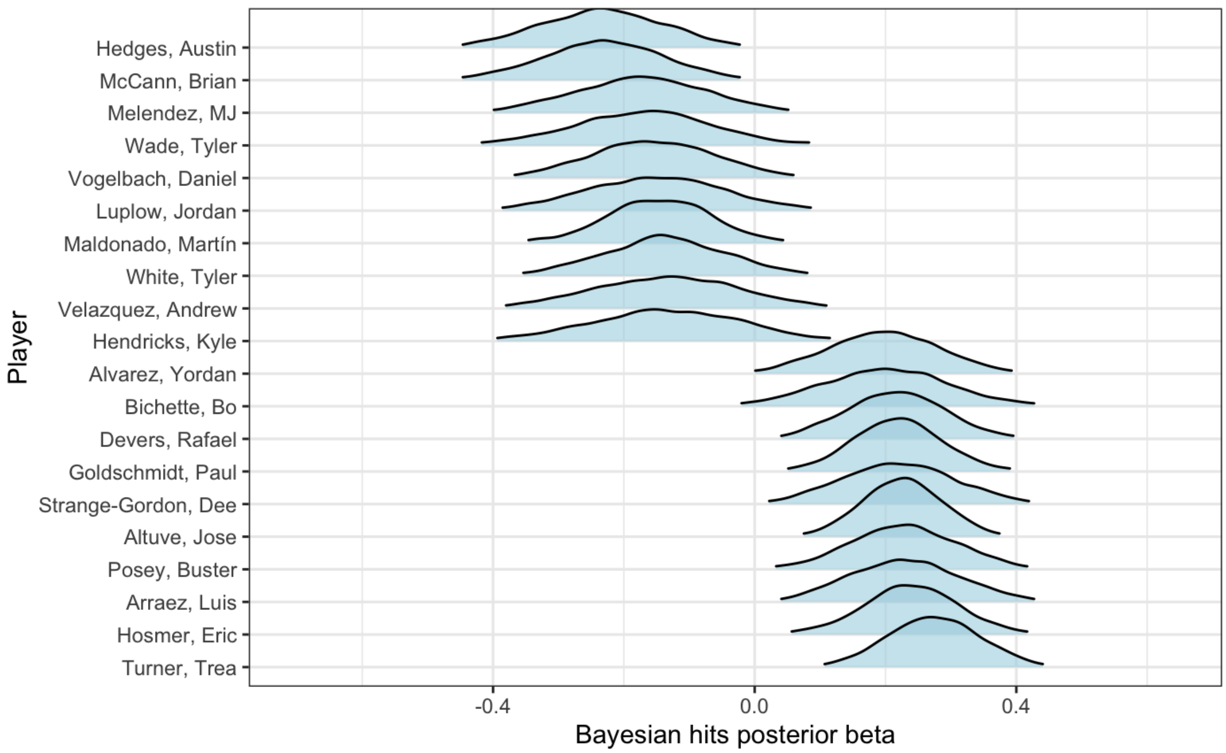
Figure 6: Distributions of Random Effects



Player evaluation

Another application for this model is player evaluation by determining which players are more likely to get hits, even as batting average declines. We used the model's posterior distributions for each player acquired to determine each player's 80% credible interval. We used the ggridges library to plot the top- and bottom-10 hitters' hit probability distributions (Figure 7).

Figure 7: Top- and Bottom-10 hitters



The “best hitters” contain some familiar names when it comes to who one might expect to see on a hitting talent list, while the “worst hitters” are comprised of catchers and players who have historically been poorer at hitting. There is some overlap between the best and worst hitters but not a large amount, which suggests that our model can differentiate between the best and worst hitters.

Discussion

Takeaways

Previous literature had suggested that some systemic league effects have contributed to declining batting averages. Our results do not support that conclusion. Factors such as defensive shifting or changes in hitting philosophy over time (captured via year- or team-level effects) were not statistically significant. Batter and pitchers effects were higher and statistically significant, indicating they are more responsible for declining batting averages.

According to our model, fastballs have a large positive impact on hit probability, so why do pitchers continue to throw them? Pitching well is a matter of keeping the batter off balance. So fastballs, particularly those with high velocity, help to set up the other pitches in a portfolio a pitcher has at their disposal. Fastballs are easier to locate for strikes since they have little movement, and having a bigger portfolio makes it harder for the batter to anticipate what pitch is coming. Changing pitches, changing speeds, and changing locations makes it very difficult for a batter to hit the ball well, though a full analysis of pitching is beyond the scope of this current model.

Limitations

While declining batting average is an interesting trend, it doesn't tell the whole story about a player's hitting ability. Power hitters, for instance, are not really accounted for properly by batting average, home runs count the same as singles. Nor are batters that work a count, take a walk, and then steal second at a high effective rate. There are most likely some batting effects not measured by batting average that could make batting appear worse than it is. Our model also does not take these factors into account for the sake of simplicity. That being said, batting average is a key component of other statistics that baseball analysts commonly cite as key measures of hitting talent. Batting average is a worthwhile statistic to model because it isolates one element of hitting talent.

Additionally, in our model, the intercept term is the largest positive influence on hit probability. The intercept value at the player level currently reflects their "intrinsic ability" to hit a baseball that is not reflected in the regression coefficients. We would need to look at other possible parameters to try and capture that ability.

Future Work

Statcast has many more variables we could add to our model, and fixed and random effects could be added or dropped. We could also look at separate models to zoom in on pitching effects, batting effects, and individual matchups between specific pitchers and batters. We could also try to solve the problem of the model not capturing the different "values" of hit types (single vs. home run) by weighting hit outcomes differently or a multi-classification model with options for each hit type. To address the limited scope of batting average on hitting talent, we could use different outcome variables, like on-base percentage, slugging percentage or batting average on balls in play.