

A Look at Strikeout and Homerun Rates in 2020

(Current through August 25)

by Dani Treisman

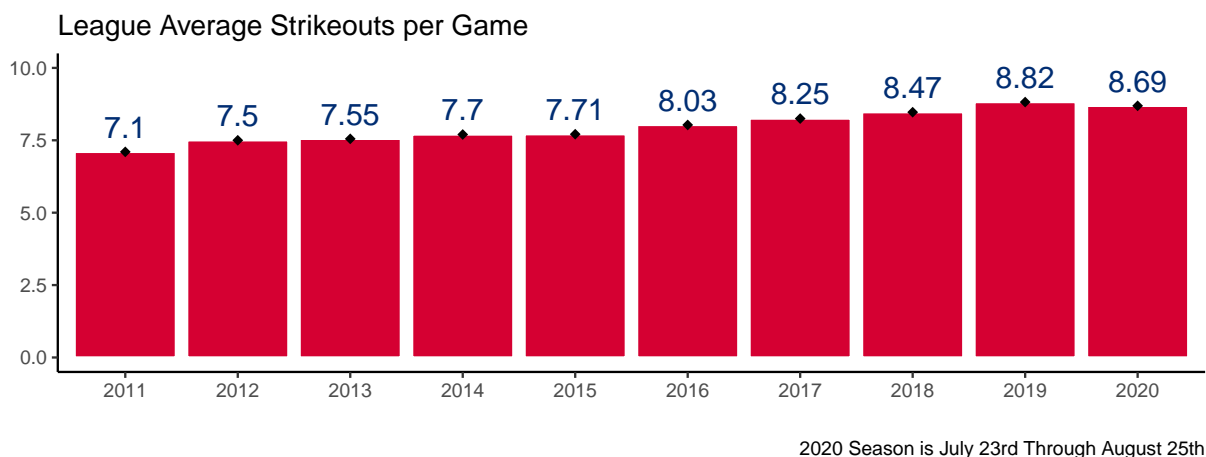
It may be cliché at this point, but the 2020 MLB season has been anything but ordinary. With this season being almost a third the length of a regular season AND starting several months late, there will likely be many deviations from traditional sabermetric models and ideas.

For example, there is an idea that the strikeout rates are correlated with the rate of homeruns. The basic idea is that players who hit more homeruns are swinging harder and being a bit less selective with the hope for a bigger payoff. It's a risk-reward trade off. This phenomenon is explained in depth in [this](#) article by Michael Salfino of FiveThirtyEight in 2019. There is a bit of controversy about whether high-homeruns and high-strikeouts is a good approach. However, strikeouts and homeruns are both seemingly higher than ever. While watching the first few games of the 2020 season however, it seemed (from nothing more than a gut feeling) that this season, strikeout rates were significantly higher than they should be, with homerun rates not rising as well. I decided to take an in-depth look at strikeout and homerun rates this year compared to previous years to see if I was correct or if I was suffering from the effects of sports withdrawals. Here is a short look at some of my findings for this season.

(Data is from [Retrosheet](#) and [Baseball-Reference.com](#))

Fig. 1 shows that the league average of strikeouts per game in 2020 are actually less than 2019 and the only year over year decrease at least as far back as 2011. I figured that it's "early" in the season (even though it's technically halfway) and it is possible that early season strikeout rates are lower as it may take pitchers some time to warm up.

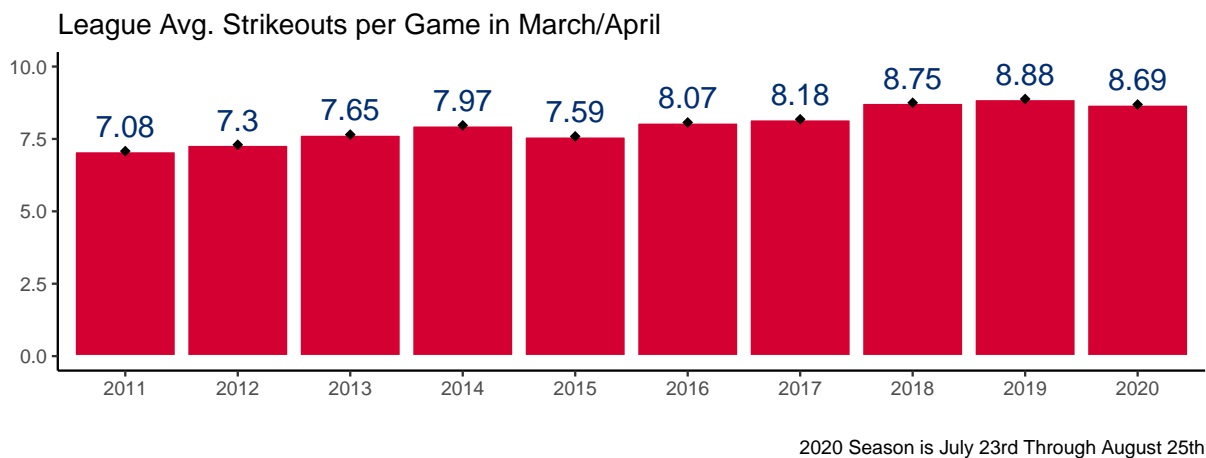
Fig. 1



To verify this I looked at Ks per game in 2020 (July/August) versus March/April of previous season (Fig. 2).

(Through August 25th, each team has played around 30-35 games, and previous seasons looked at were closer to 25-30).

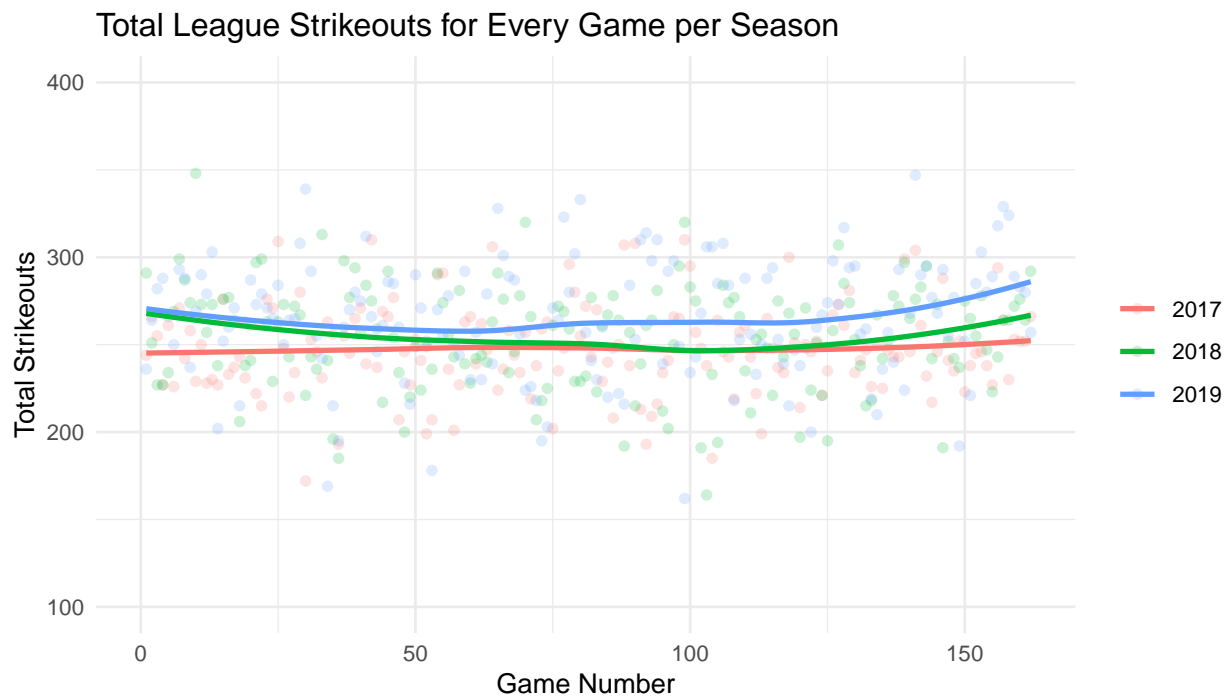
Fig. 2



2020 is the second year over year decrease in strikeout rate since 2011, the first being in 2015. It is possible that throughout the season, the strikeout rate will change, but it does seem as though it may end up lower than previous years.

To find out if strikeout rates might recover, I've plotted total strikeouts in every game for 2017 through 2019 with a smoothing line for each year (Fig.3). It looks like strikeouts remain mostly constant through most of the season, with an increase in the last 35-40 games. It will be interesting to see if this trend is the same when we look back at the 2020 season.

Fig.3



As I mentioned earlier, there is [evidence](#) to suggest that homeruns are positively correlated with strikeouts, the idea being that batters who swing harder hit more homeruns but may swing and miss more often. Fig. 4 shows this trend. With so many years it is difficult to see the year axis labels, but around 1960, strikeouts per season makes a dramatic increase, so I removed years before 1961 (see Fig.5).

Fig.4

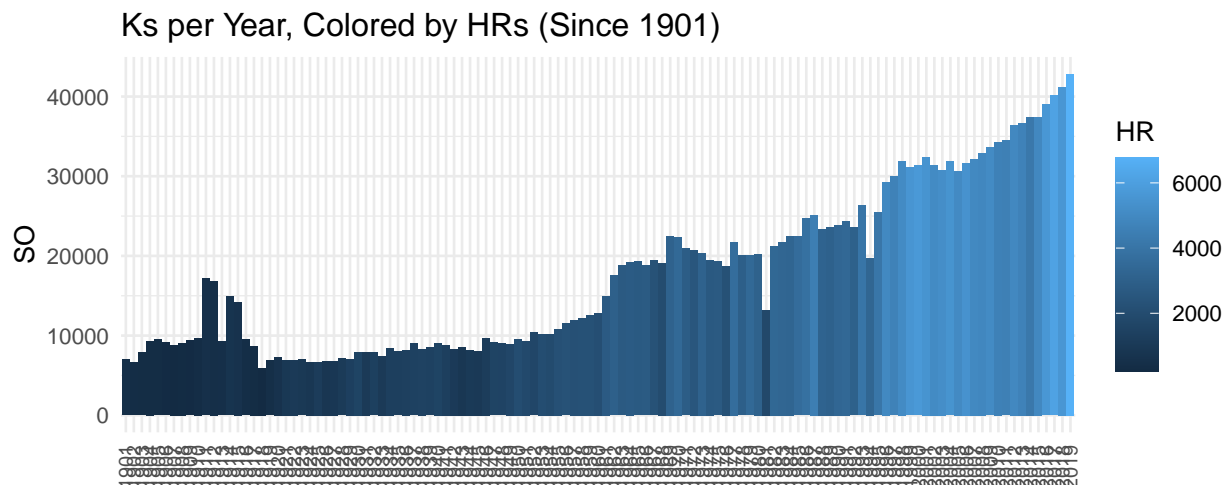
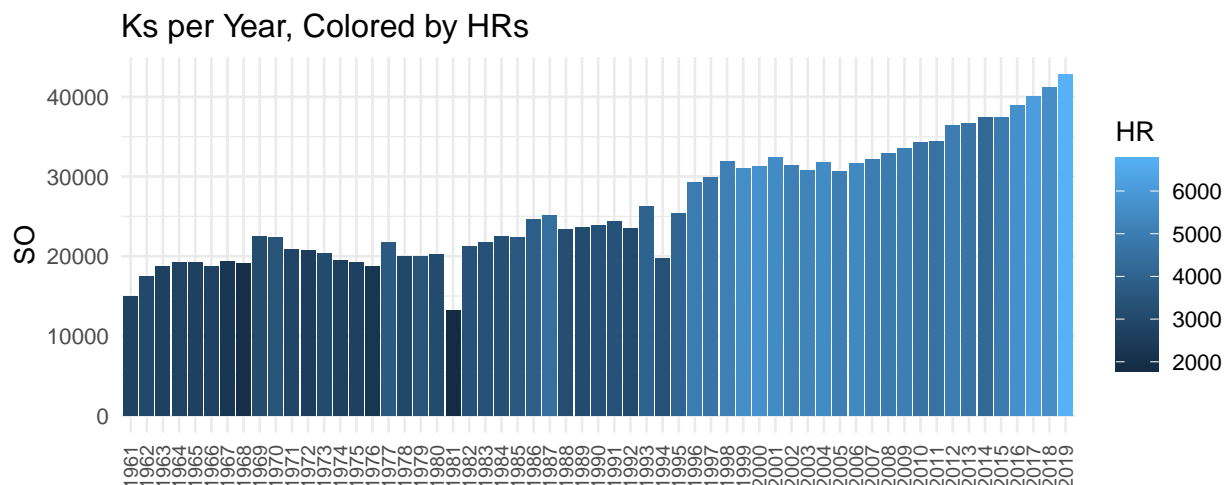


Fig.5



Zooming into the last 10 years, Homeruns per game have increased since 2011, although not as consistently between years. Homeruns per game are down compared to last year (Fig.6) and about equal when looking at just the first month of the season (Fig.7).

Fig.6

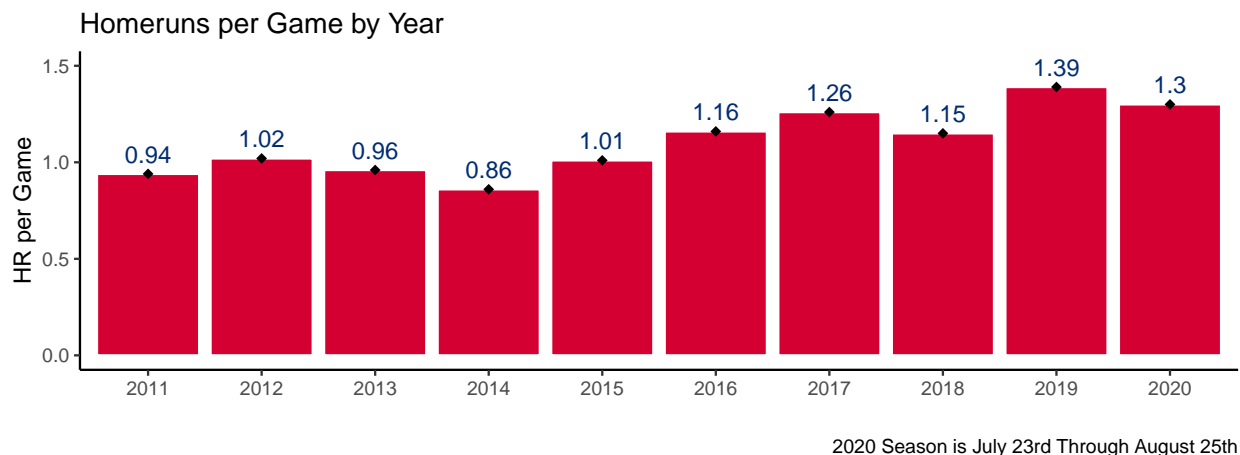
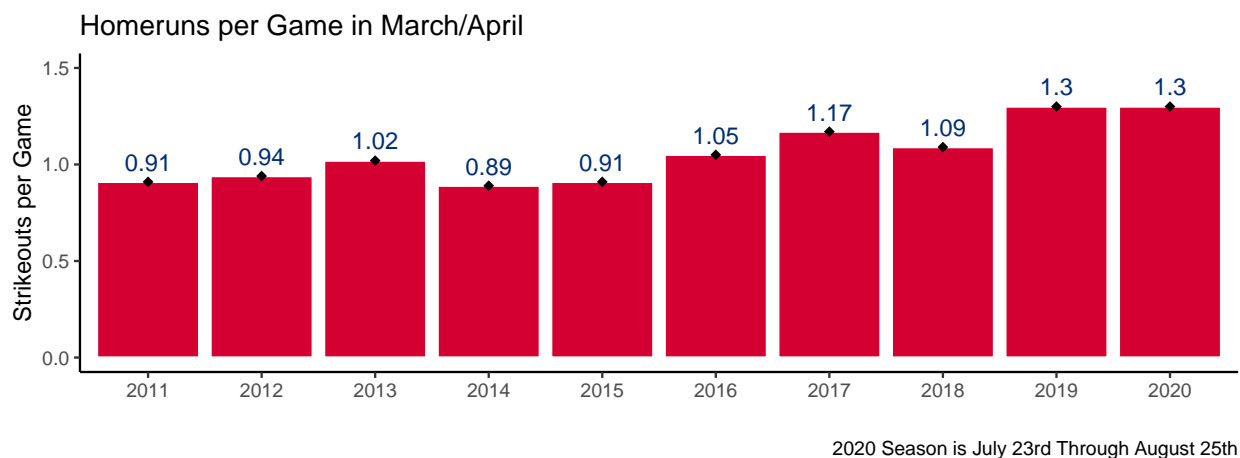
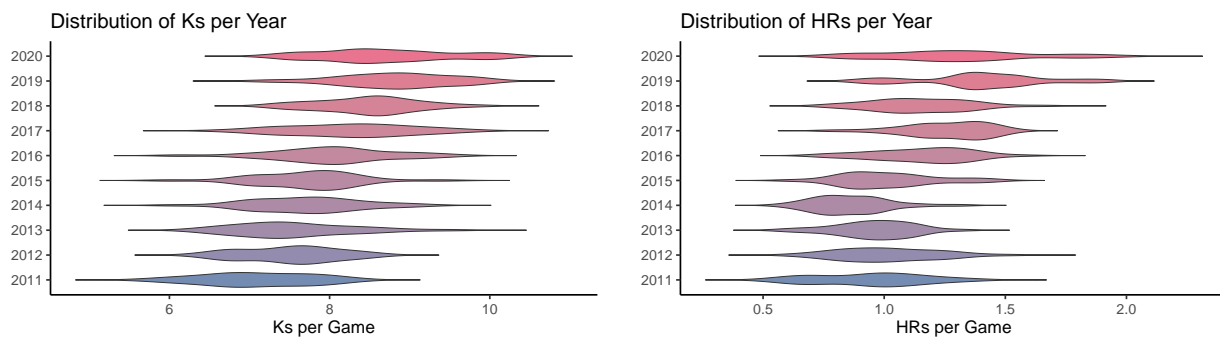


Fig.7



It makes sense that homeruns have more variability than strikeouts since they can often depend on weather, ballpark, and other factors. Also, homeruns and strikeouts are not perfectly correlated, but I am not sure why the strikeout rate is consistently increasing every year while the homerun rate, and likely many other measures, have higher variability. Fig.8 shows the distributions of HRs and Ks per season.

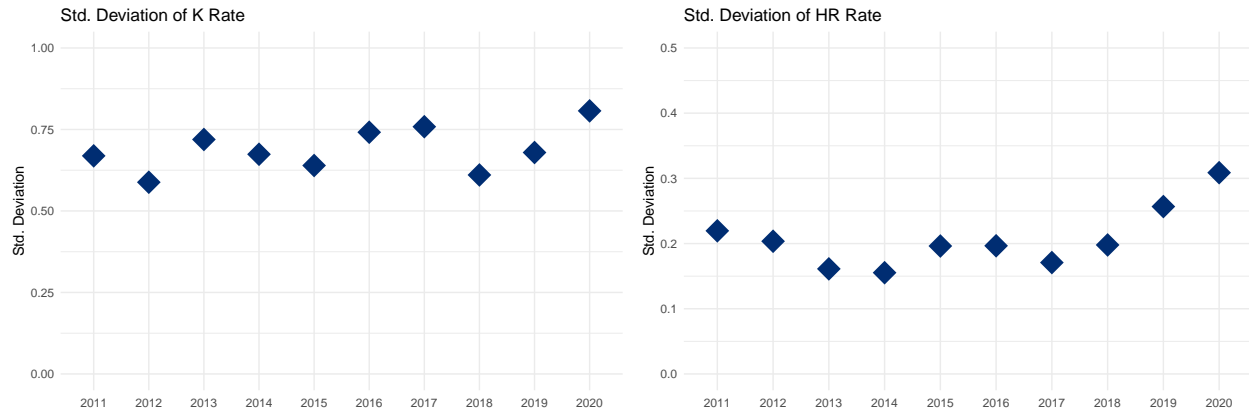
Fig.8



While I am not exactly sure what to make of these visualizations, it is interesting to see that in 2020 so

far, there is more variance in the distribution of homeruns and strikeouts than in most other seasons. It is possible that due to many teams having to place players on the injured list (IL) and replacing them with minor league call-ups there is a more variability between teams, but the standard deviation of HR rate and K rate for each season has increased in the last three seasons (Fig.9), so it is unclear from these charts.

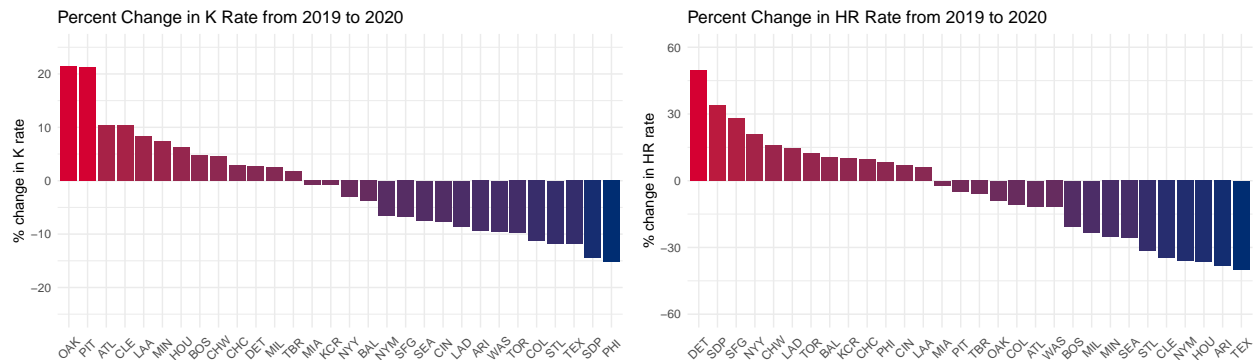
Fig.9



Let's dive deeper and see if we can find additional insights by looking at strikeout and homerun rates by team.

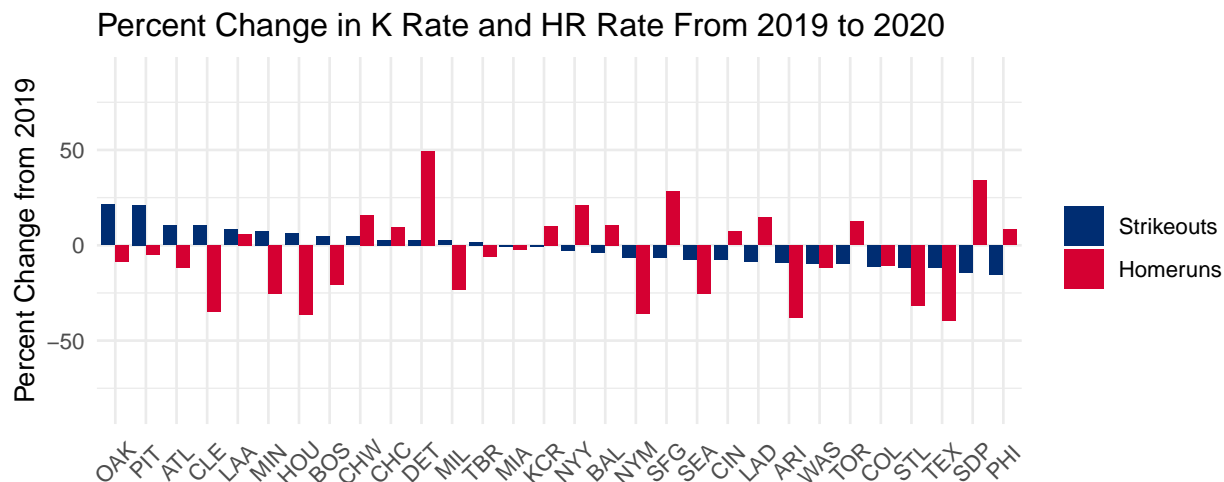
First, let's take a look at the percent change in K rate and HR rate from 2019 to 2020.

Fig.10



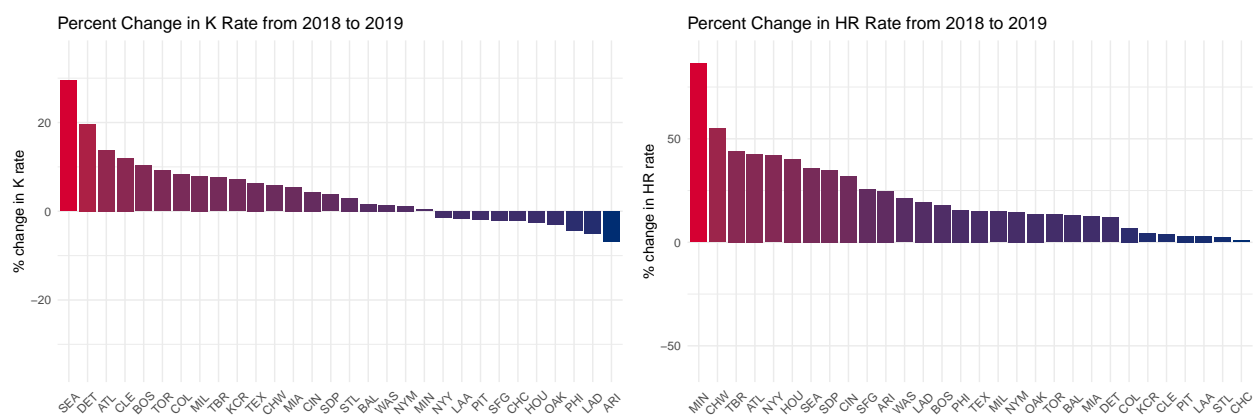
We can see from Fig.10 the high variability of the HR rate among all teams with Detroit at a 49.46% increase and Texas with a 39.86% decrease in homeruns per game. Strikeout rates seem slightly more normalized with but two teams standing out at each extreme. Oakland and Pittsburgh stand out with a 21.43% and 21.22% increase in K rate, respectively. Atlanta was the next closest team at a 10.38% increase. Fig.11 shows these numbers side by side for each team.

Fig.11



It seems that the percentage change for Ks and HRs since 2019 do not seem to follow the general idea that strikeout and homerun rates are correlated. On a per-player level, Ks and HRs are highly correlated but this positive correlation does not show here. It is possible that this is due to personnel changes or various players hitting more homeruns/striking out more than they usually do. However, since there are many more double headers due to the shortened season and games postponed due to covid outbreaks, it is possible that starting rotations are not as strong due to the need for extra pitchers in the rotation. It's also possible that some teams are benefiting from playing against teams with long layoffs, double headers, or crazy travel schedules and are hitting against them better than they would on a normal season. Lastly, with expanded rosters, players who are put on the Injured List are being replaced with players who are much less experienced. Most seasons, if a player ends up the IL, there is an MLB quality replacement on the bench. This season, with so many players on the IL for COVID, replacement players are much less skilled and some teams have significantly more players infected with COVID than others.

Fig.12

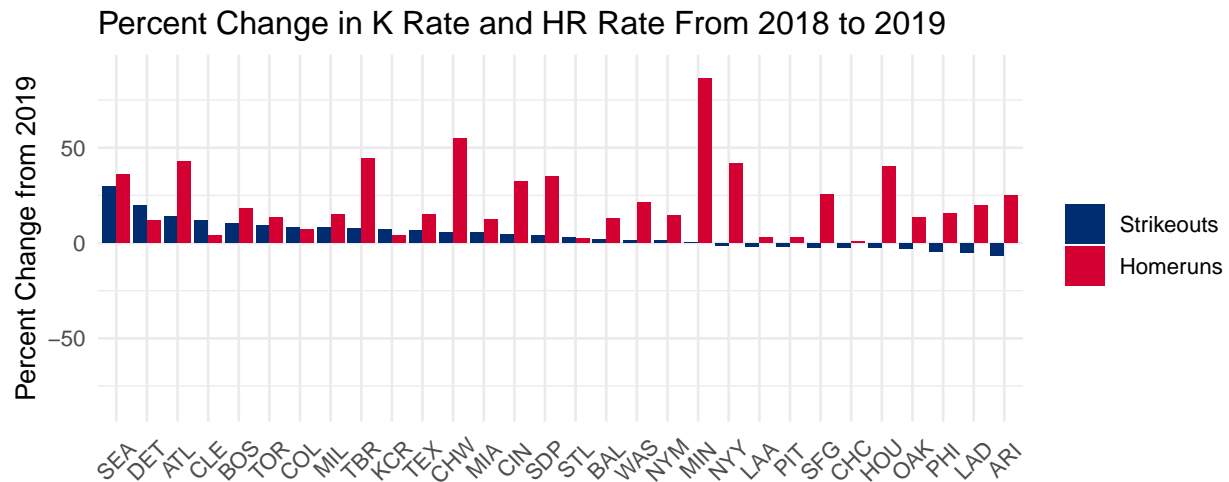


Looking at the percent change from 2018 to 2019 (Fig.12) one can see how different this year is compared to previous years. Not a single team hit less homeruns per game in 2019 than 2018 and only 10 teams had less strikeouts. from 2019 to 2020, 17 teams had a lower strikeout rate. Looking at the K and HR rates side

by side in Fig.13, the comparative trends between HR and K rates seems a bit more reasonable. As shown earlier, homerun rates have higher variability than strikeout rates, but despite that, the HR rates vs. K rates seem to make a bit more sense.

Side note: The four teams with the highest percent decrease from 2018 to 2019, all had an increase in homerun rate inversely proportional to their respective strikeout rates. This confirms that HR rates are not solely dependent on strikeout rates and that teams can simply improve and just hit better.

Fig.13



The 2020 MLB season has already shown signs of its uniqueness in many ways. The homerun-strikeout trade-off will likely continue and the rates for both will likely continue to increase over the next several years unless a new approach is discovered. Sabermetrics is constantly evolving with the growing availability and specificity of game data and it will be interesting to see how hitting approaches change as the number of tools available for analysis continue to increase. I am excited at the insights I have found regarding strikeouts and homeruns this season and look forward to analyzing more aspects of the 2020 season.

Socials:
[LinkedIn](#)
[Twitter](#)
[GitHub](#)