ALMA MATER STUDIORUM – UNIVERSITA' DI BOLOGNA

DEPARTMENT OF STATISTICAL SCIENCES

Second Cycle Degree in Statistical Sciences

"Trashgate: a statistical analysis of the 2017 Astros scandal and possible scenarios breakdown"

Presented by:

Alessandro Zilio

0000885265

Supervisor:

Prof. Mario Mazzocchi

Relator:

Prof. Daniele Ritelli

II SESSION

ACADEMIC YEAR 2020/21

Table of contents

Abstract

The 2017 season was surely one to remember for the Houston Astros. After a 101-61 regular season the team, carried by a relentless offense and a rock-solid starting rotation, proceeded to defeat the LA Dodgers on a classic WS battle, earning the Astros their first title in over 50 years of history, but at a steep price.

The Athletic, a specialized site, posted an article on November 2019 where Ken Rosenthal and Evan Drellich, big names on the reporting field, detailed allegations proved by then-Astros pitcher Mike Fiers of an illegal cheating scheme that led Astros hitters to know pitches beforehand.

This essay tries to answer an impossible question: what if the Astros didn't cheat? Would they have won the title or not?

We'll begin this paper with a simple introduction to our data and methods, and after a brief portrayal of the behind the scenes shenanigans we'll get to the meat of the order: by means of an iteration of pairwise encounters with probabilities dependent on real and positive scores we are going to be able to simulate seasons as groups of head to head matches, setting each team's score as the sum of its players' performances as defined by an *omnia ratio*.

Then we'll consider four different scenarios, each trying to account and fix for the cheating, and for each one of them we'll proceed with 1000 simulations, for a grand total of 4000 random seasons to compare to the original cheating seasons in a counterfactual test of sorts.

Finally, we'll explore the results in terms of wins, accolades and trophies, searching for an answer, and, to our utmost surprise, a different one to what we thought: it wasn't worth it, as in the majority of the cases the Astros would have at least won their division and had a shot in the playoffs, whatever kind of manipulation we do on players' contributions.

To close the game, we'll consider then why did they do it and what it means going forward for players, front office and fans of the beautiful game.

Keywords: 2017, Astros, cheating, scenarios, counterfactual

I. A walkup song: introducing our journey into a baseball season

Life is a matter of crossroads.

How many times did we find ourselves in a tough spot, having to choose between right and left? Thinking about it, the possibility and the relative burden of choice-making is what defines a human being as one.

Usually all these what ifs don't have much of a post-opera evaluation, as we are all great when it comes to backtracking our mistakes and make optimal decisions in hindsight. But there comes to help our most loyal friend, one who's normally right 95% of times and wrong the rest: statistics.

Probabilities, scores and distributions allow us to look back, hopefully not in anger, to our checkpoints and to calculate someway somehow what would have happened if we turned the other way instead of keeping our path.

The Houston Astros won the 2017 WS, the ultimate trophy for each baseball team with all honours, glory and perks along the way, and that is a fact. The Houston Astros cheated in the 2017 season, and that is a proven fact too.

So, the aim of our essay is to at least clarify the pinnacle of all questions, the classic what if that makes for some interesting stories and tv shows: what would have happened in that 2017 season if the Astros didn't cheat? Would they have won either way or was it actually all a product of some dirty misconduct?

Giving a definitive answer is not what we are trying to achieve, as we are statisticians and we stay true to our words: we are 100% sure of anything but we can be 95% sure of something, under some constructions and limits.

We also aren't that much into shallow beliefs and as someone in the business used to say: "In God we trust, all others bring data", so we are going to need some stone cold figures to process and analyse in the search for some light at the end of the tunnel.

Our batch of data will be encompassed by two principal sources: the complete dataset regarding the 2017 MLB season, with numbers, accolades and statlines for all teams and players, publicly available on the official website of the league and in other specialized blogs such as FanGraphs, and then what we could call the "cheating dataset", which is actually a subset of the complete one related only to the occasions in which the Astros cheated within that season, a product of a fan as I am, not a professional but an enthusiast with some serious informatic skills, that is also publicly available in his website.

These datasets contain all information about each player's performance in the 2017 season, so it could be natural to have an idea of what we are going to do: if we take the Astros 2017 players' performances from the complete dataset and remove those that are in the cheating one we should obtain a "clean" slate for the Astros players.

That is partially true, but what to do with it? That's where we get bold! As we want to know if cheating really made the Astros win it all we have to do the time walk again, push the big red reset button and start anew: replay the 2017 season!

As cool as it sounds, it is a rather mechanized process that involves a cold point of view on a baseball game, depriving it of all its emotions and mystery: a game is nothing more than a one vs one matchup where the strongest usually wins if that same match is played over and over again. That's why in the playoff from the divisional round up to the final World Series the winner is decided based on a series of games rather than a do or die single elimination: we want to crown the best, not the luckiest (although, as Mike Matusow would say, in life you have to be both the best and the luckiest, at the same time).

If matches are decided by the strength of teams then we need a proxy for it as well, and we already have one! Each player's contribution to the cause has been recorded in the complete dataset so that at the end we have an idea of the value of each team. To our rescue then comes the branch of statistics strictly related to baseball: sabermetrics.

Since the early 2000's evaluating a player's value, not only in a monetary sense but also as in terms of on field proficiency, has been a key target for all sabermetrics and their efforts produced a multitude of stats and algorithms. In our essay we are going to define and use WAR, Wins Above Replacement, a nice round number that encompasses all the qualities of a player and defines his worth with respect to a replacement, a baseline player as one available on the market or in the minors.

After a detailed construction of WAR and an evaluation of the 2017 Astros using the statistic, we are going to focus on how to replay a season in its entirety, and to do that we'll jump back in time, all the way to the 1950's.

In those bygone days two statisticians, Bradley and Terry, developed a quite simple and straightforward model, that we'll refer to as B-T model for the sake of simplicity, to consider a pairwise encounter: the probability of a team, or player, A to win against his opponent B is equal to a ratio between positive and real scores, as proxies of the strength of each counterpart.

The model has always been used to find, through ML estimations of some sorts, those scores having a set of A to B encounters, but what we are going to do is the opposite: by setting the scores for each team equal to the team's WAR, as sum of those of its players, we can then calculate each win probability for every couple of teams.

Finally, we'll put all those ingredients into a single process: using R we'll construct a function that, on a stepwise procedure, builds a season as a series of one to one matches, then we'll play all those matches randomly, on probabilities set by the specifics of the B-T model, and we'll end up with a win-loss record for each team.

After reorganizing the standings by record we'll crown the division and league champions, with the latter going on a standoff in the World Series, as a seven-game battle where the winner will be decided by… a multinomial.

The concept of this paper is simple: while maintaining all other teams' WAR fixed, what happens when we change in negative the Astros original team WAR in the 2017 season? What is the effect of a big reduction in WAR with respect to a smaller one?

To do this we'll introduce four scenarios, each with different implications on Astros players' WAR figures related to the combination of the complete dataset and the "cheating" subset: on one of them we'll erase the subset from the complete dataset, in another one we'll keep only part of it, while in the other two we simply won't consider the subset and reconstruct the 2017 players' statlines based on previous seasons' performances.

For each scenario will correspond a different Astros team WAR, and as such we'll have different probabilities and records. As to have solid results we'll play 1000 seasons per scenario and collect the outcomes for the Astros teams in terms of record, final position division and league-wise, eventual WS appearances and wins.

On a clearer point of view, we'll address the landscape of scenarios and results, trying to answer the oldest question in the book of sports: did cheating really matter?

And, by the numbers, the answer will be a sour "not that much": the Astros were just so much better than all the other teams in their division that winning it was a given also when losing a third of their value in WAR, while we'll explain why playoff results are to be taken with caution due to the simplistic way we built our postseason.

As a closing note we'll then ask ourselves, given those results, why did the Astros cheat? And what is the next page in the book going to write about the protagonists of the biggest sport scandal of the last decade?

It will all come down to human nature, greed, a tainted legacy and some dirty base hits.

II. Trashgate: the 2017 Astros run and how it all went down the drain

As in any other sport, the aim of every single baseball team in the MLB is and has always been that of winning the "last damn game", as to quote a great movie on the subject.

Therefore, it shouldn't come as a surprise that, from its birth, cheating has been an integral part of the game, leading to scandals, rule changes and public hearings aired on national tv.

One of the most famous instances of cheating in the early days came in the 1910's, when the White Sox, led by aptly-named Shoeless Joe Jackson, where accused of having literally sold the World Series games to their opponent, and as a result Jackson and others were banned from the game and are forever remembered as Black Sox, a stain on one of the historic franchises in American sports.

Other illegal methods were common in the 1930-1960's, in great part concerning the equipment: either the bats were corked, adding an insert of metal to have more power, or the ball was doctored with substances such as liquorice, pine tar, dirt or spit. While these intricacies were admitted at first, the fact that they led to unfair advantages led to their banning and to a stricter rulebook.

But the worst was yet to come, and it involved the League directly: between the 1980's and early 2000's a series of massive scandals regarding top players tarnished the reputation of MLB and as a result it is still today remembered as a proper "Dark age" in baseball: the Steroid Era.

In that span of time several among the best players in the game, and future sure-fire Hall of Famers such as Barry Bonds, Roger Clemens, Alex Rodriguez and many others, were involved in an incident regarding a laboratory named BALCO that provided several athletes in major sports with PEDs, Performance Enhancing Drugs, illegal substances as steroids and HGH.

Names were listed in the notorious Mitchell report and hearings became a national matter, were beloved players either lied and committed perjury, as Cubs slugger Sammy Sosa, quietly admitted their guilt, as A's first basemen Mark McGwire, or simply tried to evade questions, as Bonds and others.

Not only that, it was also proved that MLB commissioner Bud Selig knew of their misdoings, kept his mouth shut and then tried to stop the car when it was too late by calling for anonymous tests.

That remained, at least until a year ago, the lowest point in MLB history, with records broken but marked with an asterisk as a sign to never forget what happened behind the scenes, and exceptional players, the Bonds and Clemens, still waiting to be enshrined as greats due to their reputation as cheaters.

And then, on a November of 2019, one of the best kept secrets in the industry came out and sent baseball into a frenzy: the Houston Astros, one of the most exciting and beloved teams in MLB for its young core of talented players and an analytical vanguard in the front office, were accused of having cheated since their title winning 2017 season.

In this chapter we are going into detail on how the Astros cheated, who were the culprits and what have been the consequences, as to then start an analysis of the effects that cheating had on players and whether it was actually needed to win.

a. The run: how they won

If a movie had to be made about a baseball team in the last decade, the 2017 Astros could have been a top choice, with or without the scandal that came out later.

On paper the team was poised to compete for an October berth, with a good chance at winning its Division, the American League West, and then fight some immovable objects such as the Red Sox and the Yankees in the playoffs.

At the All Star break the Astros set atop of their division with an astounding 60-29, a product of both dominant pitching and bashful hitting.

The starting rotation, headed by Dallas Keuchel and held up by a resurgent Charlie Morton, somehow flaming 98 mph fastballs right by hitters after an injury that made him lose a year, and by the explosion of Lance McCullers jr, gave a rather easy job to a modest bullpen, with the sole Ken Giles as a viable late-game option.

Meanwhile the real surge was happening in the bats: the whole lineup experienced unforeseen progresses, power and discipline went through the roof and while Jose Altuve kept on slashing hits at his usual pace, George Springer cemented himself as a devastating leadoff hitter, Carlos Correa showed the talent to be the top shortstop in the league and Marwin Gonzalez went bonkers by adding power to a consistent average bat, while also playing everywhere and every day.

Some numbers and accolades for the pre All Star Astros speak for themselves: two Pitcher of the Month awards in April/June, one for Keuchel (5-0 record, 1.21 ERA) and the other for McCullers (22 consecutive scoreless innings in 4 starts), a Player of the Month award for Correa (.386 AVG, 7 HR, 26

RBI), a comfortable 16.5 games lead on the Division and five players chosen for the All Star Game in Keuchel, McCullers, Altuve, Springer and Gonzalez.

July was more of the same, with Altuve getting Player of the Month thanks to a hilarious .485 for the fifth-highest average in one month since 1961, but then the wheels started to come off in August: following the All Star break the Astros slumped and went 11-16, moreover the city of Houston suffered the consequences of Hurricane Harvey with floods, deaths and devastation that left the Astros as a sole bright spot for all Houstonians.

On the verge of the Trade deadline, after 17 losses in 27 games, the general manager of the Astros Jeff Luhnow swung a trade with literally seconds to spare and brought home Justin Verlander, one of the most accomplished pitchers of the decade and stalwart of the rebuilding Detroit Tigers to bolster the rotation and revive a team that started to crumble.

Once an almost done player, Verlander, helped by pitching savant Brent Strom and by an analytics department like none, showed that he still had a lot to give and carried Houston through the last part of the regular season, winning all his five starts and helping the team on a 22-8 run that ended the regular season for the 2017 Astros at a team record 101 wins.

Being first in the Division granted the Astros a free pass, avoiding the Wild Card game straight to the ALDS, that went on to be the start of a rollercoaster of a postseason run.

In the ALDS the Astros faced a juggernaut, the Boston Red Sox, in what was previewed as a clash of titans and instead went on to be a comfortable 3-1 series win for Houston, thanks to unbelievable pitching performances by Keuchel, Verlander and the McCullers/Morton duo.

Waiting for another battle were the NY Yankees, the team to beat in the American League, that proceeded to sweep the Minnesota Twins in their ALDS.

The Yankees-Astros 2017 ALCS turned out to be a nail-biting seven game series, with an extra inning game, a walk off at home for Houston and a deciding match where the M/M duo held the fort and completed a hometown comeback from 2-3 to 4-3 to send the Astros to the World Series for the first time in more than 10 years.

Asking a random fan, an insider of the game or a player, the 2017 World Series between the Astros and the LA Dodgers were arguably the most exciting in a decade or so, at least before all the behind the scenes shenanigans came out.

A much awaited back and forth bloodbath between two history-caliber offences transformed into a pitching clinic in the first four games, ended with a 2-2 series tie: in game 1 Clayton Kershaw, LA and MLB’s top pitcher at the time, silenced Houston’s lineup, while in game 2 a great performance by Verlander was rescued only in the ninth inning via a Marwin Gonzalez homerun off of LA closer Kenley Jansen. Game 3 and 4 were more one-sided, for Houston and LA respectively, and the series went on in Houston for game 5.

And what a game! One of the most memorable in WS history, a ten-inning mashing party that saw homeruns aplenty, with Houston falling and rising back up thanks to the longball more than once, first a Springer revival, then Altuve to tie and Correa on a skyscraping short porch special. At the end the Astros prevailed in the tenth on a walk off single by Alex Bregman off Jansen, to close the book on a six-hour marathon that led H-Town to a 3-2 series lead.

Game 6 had Verlander going for Houston, but not the bats and so it was back to LA for a winner-takes-all game 7. And really, a good game but not as that game 5: Houston went on an early lead and while McCullers exited early due to pure wildness, the bullpen and a final three shutdown innings by Morton secured the win and the WS, the first for the Astros in their 50+ years history.

Scenes of sheer joy went on and on for more than an hour at Chavez Ravine, with players and staff celebrating each other, Morton and Verlander as their heroes, Springer winning a deserved WS MVP award and manager Hinch, GM Luhnow and owner Crane thanking the fans and the city for its support.

Those were the days for a Houston fan, and as one I remember all too well: the hours spent on trying not to scream at the computer, the highs and lows on wins and losses, an entire sleep pattern gone to dust as games started at 2am (time zones…) and ended at 6.30, or 7.30 for that game 5, so that I went to bed when people woke up, and the complete elation of winning it all.

And it lasted longer than I thought, such a great feeling that neither a subpar 2018 season nor a last-game WS loss in 2019 could break the happiness that derived from looking at the WS winner patch on my baseball cap.

But in November 2019 the dream shattered as glass, no more joy or happiness, just a cruel and cold title that burned down an entire eventful winter: the Astros cheated.

b. The scandal: how they lost

November is usually a month of transition for baseball: the postseason is over and teams start to prepare for the next year, reasoning on what to do with their rosters, with possible free agents, salaries raising due to arbitration and trades that can be made to improve the club on the road to win the WS.

That wasn't the case in 2019: out of left field the Athletic, one of the leading sites regarding baseball, released an article written by Ken Rosenthal and Evan Drellich that would cause an earthquake all throughout the sport titled "The Astros stole signs electronically in 2017 part of a much broader issue for Major League Baseball"(Rosenthal, Drellich, 2019).

The piece detailed allegations proved by then-Astros pitcher Mike Fiers of an illegal cheating scheme that led Astros hitters to know pitches beforehand, to gain an unfair advantage on opposing pitchers and increase the chance to hit a good pitch or let go of a bad one for a ball (Passan, 2019).

To fully grasp how this system came up we need to take a step back: as dutifully explained by notorious statistician Jared Diamond, the foundation of the cheating complex that made the 2017 Astros negatively famous stems on an algorithm developed by the analytical department of the team and called "Codebreaker"(Diamond, 2020).

An interesting analogy to understand how the algorithm worked is to compare it to the Enigma machine that made Turing famous during WW2: he was able to craft a system that could translate encrypted enemy transmissions in real time, giving to his side an enormous advantage.

Baseball is sort of a war when you think about it: pitcher and catcher find an agreement on what pitch to make, type and location of it, using signs that the catcher's hand shows to the pitcher, who can accept or ask for another sequence.

The batter can try to have a peek, although it falls as disrespect of unwritten rules with potentially harmful consequences, or, and this is commonly accepted, if there's a runner on second base then he can try to signal the batter at the plate the pitch if he decoded the message. That's also why catchers and pitchers agree on different sequences and codes for each kind of pitch (fastball, change, curve…) depending on the situation as far as runners on base are concerned.

What the Astros' Codebreaker did was nothing illegal, rather a great idea: using past games' broadcasts the different sequences were decoded, for each pitch type, location and on base situation, so that at the end there was a clear view of the "playbook" of the opposing side (Diamond, 2020). Then it was up to the batters to study the sequences and their corresponding meanings and deliver them to their teammate at the plate if they ever got to second base, also considering the need for an absolute eagle eye to see catcher signs from almost 100 feet away.

If it all stopped there, with these "dark arts" as Tom Koch-Weser, one of the front office minds that created the Codebreaker (Diamond, 2020), defined the algorithm, there wouldn't have been a scandal at all, maybe some curiosity from the industry and some complaints by other teams but nothing major. What happened is what the title of Diamond's article states: "The Astros front office created Codebreaker; the players took it from there".

As recounted by numerous articles and by the official report of the commissioner regarding the scandal and its repercussions (Manfred,2020) what happened later, the real cause of all troubles, was a coach/player driven scheme led by bench coach Alex Cora (Browne, 2020) and with veteran player Carlos Beltran at the forefront (Rosenthal and Drellich, 2020).

No rule was in fact broken with Codebreaker, it was the step further that went beyond the limits of legality: a camera was installed in center field zooming directly on the opposing catcher's signs, then the camera feed was displayed in a monitor installed on the tunnel leading to the dugout as to be visible in real time on the neighborhood of the action.

Then how come the batters got an advantage if the monitor was far away from their sights? That's when a technological affair became ludicrously trivial: a member of the staff was given the part of the signal transmitter from the tunnel to the field via banging a trash can with a bat…yes that's true!

The number of bangs was all the information the batter needed: using the Codebreaker playbook a bang could mean a fastball or a breaking ball in different situations, but the perk this rudimental system gave is pretty clear. Knowing what's going to come, whether the ball is going to be straight, changing or curving is a great piece of intel in a game that is all about inches and small differences: it's like being aware of what the other side does in a prisoner's dilemma, so you can always make the best choice.

Take for example a pitcher that throws a fastball and a devastating slider: a ML batter can time up on any fast and straight pitch, from 80 mph to north of 100 mph (160+ kmh), but a sharp breaking ball is always harder to hit and leads to more swings and misses, strikeouts and unpleasant experiences for the batter.

Now imagine that this pitcher knows that batters are helpless against his slider, so he doesn't even need to risk throwing it in the strike zone and people are still swinging at pitches away or in the dirt. The plan for him is easy: get ahead with some breaking stuff and early count fastballs and proceed to send guys packing with some sliders away from the zone.

Nice work, but not if I know what you're going to throw! Imagine a classic at bat: the pitcher tries to get ahead 0-1 with a slider away but here comes the bang! The batter knows it's going to be a slider, and usually out of the zone, so he doesn't swing, and the count goes 1-0, with big differences in the run production outcomes (Clemens, 2020).

This is what the 2017 Astros did for the majority of the season at home games, at least until an incident that really changed the complexion of this issue, such that without it there probably wouldn't have been a thorough investigation, a national debate, and also this thesis for what it's worth.

It was late in the season, a meaningless game between Houston and Chicago: the Astros already had the Division in hand and their minds on the playoff, while the White Sox had nothing to fight for in another season of rebuild. The real clue is that there were really few spectators at Minute Maid Park in Houston that day, which meant no loud noises, chants and heckles.

In the final third of the game Chicago pitcher Danny Farquhar, who by the way almost died in the same ballpark due to a brain aneurysm during a game, stepped on the mound trying to shut down the Astros side but noted a strange fact: whenever his catcher signaled for a changeup, a bang could be heard coming from the Astros dugout.

Something was fishy so he called for a mound visit to change the sign sequence, in a pretty deliberate way, and from there on no more bangs were heard. Small problem: Farquhar wasn't the only one who smelled something.

Noted baseball fan and content creator Jomboy, Jimmy O'Brien in real life, posted on his YouTube channel a clip of the at bat that went viral and started fires of protest and resentment from all other teams: the bang, thanks to the absence of other significant noises, was crystal clear, repeated at each changeup thrown by Farquhar (Jomboy media channel, 2019).

The veil came off and the witch hunt started galore: fans, statisticians and opposing players started digging on past games to see if they could hear those bangs, if there was a scheme, if the Astros were really cheating, a tangible and proven fact, not some shady paranoia.

And, spoiler alert, they were: evidence started to pile up and commissioner Rob Manfred couldn't really escape the heat coming from the press, the other teams and the community and called for a massive investigation, with thousands of interviews with players and computers and documents seized from the Astros front office.

That's how the trash-banging scandal, or what I would call "Trashgate" in a classic American spirit, started, on a road that would change the history of a franchise, the career of several players and the perception of the game itself: baseball wasn't as clean, as pure, as above us mortals as we thought.

Baseball and all its community had to wait until February 2020, when the aforementioned report on the investigation came out, to see the repercussions of the Astros scandal.

The punishment was among the harshest ever inflicted not only in baseball, but in the history of American sports, and yet many criticized commissioner Rob Manfred's decisions as a slap on the wrist, not hard enough to prevent future issues of the same kind.

According to the report (Manfred, 2020) the penalties for the franchise were: the maximum fee as per MLB's policies, around 5 million $, the loss of the first and second round draft picks in the 2020 draft and the depletion of the international money pool. That means not only a direct financial hit, but also a great loss in terms of talent, both national coming from college and HS and extra-USA through signing bonuses.

Furthermore, for their inability to act and put a stop to the matter, GM Jeff Luhnow and manager AJ Hinch were suspended and banned from the field for a year, while punishment on bench coach Alex Cora was deferred after a similar investigation on the 2018 Red Sox. On the aftermath both Luhnow and Hinch were fired by Astros owner Jim Crane, on a sort of cleaning house process (Passan, 2020).

What really ticked off the great part of the industry was the absence of any kind of punishment to those who gained most from the scheme: players. Due to the fact they collaborated in the investigation, rather to assure they would do so, they were exempted from fines, suspensions and other bothers, which caused a non-trivial internal fight in the MLBPA, the Players' Association, between Astros and other teams' players.

Some of the best in the business, Mr. Baseball himself Mike Trout, were waiting for a hit straight onto the middle of the problem and instead they saw their Astros counterparts getting pitched around, to that followed heavy criticism and irrecuperable cracks on friendships.

Discussions went on and are still coming out each day after a year, to get the magnitude of this scandal. Astros players and staff tried to stem the bleeding either apologizing and admitting their misdoings, Hinch first on national TV and then some ex-Astros such as Evan Gattis, Morton and Collin McHugh, or hiding behind some corporate prepared speeches, first and foremost a cringing press conference with Altuve and Bregman.

The heights of the debate were reached after Carlos Correa, who was an original apologizer, went on a heated exchange with both LA Dodger Cody Bellinger, that asked for the restitution of the 2017 WS and gave all credits for it to the cheating scheme, and NY Yankee Aaron Judge, beaten by Altuve on the 2017 MVP award, that called the accolade as stained and asked for it to be given back.

The Astros shortstop went hard on them, saying that, while guilty, the Astros won the 2017 with merit and so had Altuve on the MVP, not only that, the latter was also against the scheme of trash-can-banging and asked for it to not be employed on his at bats (Rosenthal, 2020).

Nevertheless, the scandal added another tail early in 2020: now Red Sox manager Alex Cora was given the same yearly ban as Hinch and Luhnow, although investigation on 2018 Boston didn't find the same issues, and NY Mets new manager Carlos Beltran was dismissed by the team shortly after the Manfred report came out, making him an almost unique case of manager fired after zero games at the helm (Waldstein, 2020).

Now what did Astros fans do while their team was being demolished by fair critics? Some of them tried defending the players on silly remarks such as “everyone is doing it, we were just unlucky to be caught”, others, like me, waited for it to end and then make some evaluations.

And then there’s a little cluster of Astros fans that was fed up with the team and decided to dig deeper on the scandal and muster up some hard facts, not to support but to bury their former beloved franchise for its behavior. Tony Adams was one of them and his research came out to be a national sensation and a cornerstone for many studies and this thesis.

III. Bangs: a fan’s effort on the pursuit of truth

In the forthcoming chapter we are going to discuss and consider one of the databases we are going to use in our analysis, the Bangs dataset.

This frame was not a product of a statistician per se, rather a fan of the game and the Astros, as myself, but that doesn’t mean it isn’t well structured and built, on the other hand it is so particular and detailed that only a fraction of it will be employed in our work.

The work by Adams is related to audio recognition of singular audio traces (our bangs) among the plethora of sounds in a stadium and can be compared to searching for bird screams in a forest (Ulloa et al., 2016). The process itself is akin to several works on SED (Sound Event Detection), using YouTube soundtracks of games and labeling the bangs as in the point-labeling of data per (Kim and Pardo, 2019), although without involving a deep learning model as usual, i.e. (Mandel et al., 2020), but just stopping at the construction of the dataset.

If we consider the whole data regarding the 2017 Astros season that can be found in the official MLB website and also in specialized libraries such as FanGraphs and Baseball Reference, the Bangs dataset is none other than a subset of the complete dataset where a dummy, equal to 1 if cheating was employed and otherwise 0, was added.

The Bangs dataset is comprised of 60 of the total 162 games on a season, as it considers only home games on a certain time interval, and contains information about players, state of play and outcomes on that we’ll work and manipulate later on our scenarios.

After introducing the dataset we’ll consider what kind of findings came out from other authors’ analyses on the same data, and as expected they’ll point out to some kind of advantage for Astros players, but not as big as we could have thought: hitting a baseball is still damn hard also when you know what’s coming!

Finally, we’ll move to our own analysis, abandoning player specific considerations for a rather brazen adventure: simulating an entire baseball season.

a. Adams' dataset: figures and meanings

There's a massive amount of sheer patience and work behind the construction of the Bangs dataset.

While it only encompasses home games, and only almost 60 of them, the procedure Adams had to follow to extrapolate what we'll employ later is not that straightforward. Think about it, in the Farquhar case it was the perfect storm, no fans at the stadium and no noises, but how can we detect the trash-banging among the standard ballpark sounds?

What Adams did is analyzing the spectrum of sounds as sort of audio tracks for each game, compare it to clear cases of trash-bangs (as the Farquhar one) and recognize the number and position of bangs for each of the 50+ games in his dataset. In his words: "I wrote an application that downloaded the pitch data from MLB's Statcast. This data has a timestamp for every pitch. I then downloaded the videos from YouTube and, using the timestamp, created a spectrogram for every pitch. A spectrogram is a visual representation of the spectrum of frequencies in an audio file. I could then playback the video of the pitches and, helped by the visual of the spectrogram, determine if there was any banging before the pitch." (Adams, 2020)

It may seem like we are dealing with a small scale phenomenon, but it isn't so: there were cases of 2+ bangs per at bat, games in which the total number exceeded 20, meaning that the dataset, while containing only a third of the games the Astros played in the 2017 season, is pretty big and apt to deeper analyses, that have and are being done nowadays.

Adams' dataset is available both in the csv and Excel format, which I'm going to refer to as per his structure that makes for an easier visual explanation.

Let's have a look at a snippet of the dataset, and then see if we got the gist of it:

game_id	game_pk	game_date	oppo	final_a	final_h	inning	top_bot	batter	at_bat_event	pitcl	pitch	has_	bangs	call_code	description	on_1b	on_2b	on_3b	youtube	pitch_	youtube	pitch_d	game_pi	event	pitch_play	atbat_play	away_	home_
2017_04_	490111	2017-04-03	SEA	0	3	1	bottom	George Springer	Home Run	SI	FB	n		B	Ball	f	f	f	af5e55C(	1473	https://	2017-04	1704040(	32	ca9ed282-	1421aabe-	136	117
2017_04_	490111	2017-04-03	SEA	0	3	1	bottom	George Springer	Home Run	SI	FB	n		F	Foul	f	f	f	af5e55C(	1489	https://	2017-04	1704040(	33	7f89900c-f	1421aabe-	136	117
2017_04_	490111	2017-04-03	SEA	0	3	1	bottom	George Springer	Home Run	SL	BR	n		B	Ball	f	f	f	af5e55C(	1512	https://	2017-04	1704040(	34	6c875b0c-l	1421aabe-	136	117
2017_04_	490111	2017-04-03	SEA	0	3	1	bottom	George Springer	Home Run	SI	FB	n		E	In play, run(s)	f	f	f	af5e55C(	1529	https://	2017-04	1704040(	35	1421aabe-	1421aabe-	136	117

We had to chop off some space here and there to fit it in, but that's how the dataset is built. Yes, it seems quite complex but to be fair we'll use just a modicum of the information provided.

To clarify and get things straight here's some context on what we're dealing with:

- *Game id*: the id code for each game as stored by the MLB website. Ex 2017_04_03_seamlb_houmlb_1 means Seattle vs Houston on 03/04/2017
- *Game pk*: a counting number for games as in the whole league
- *Game date*: the date on which the game was played
- *Opponent*: the name of the opposing team in its 3 letters abbreviation Ex. SEA = Seattle Mariners
- *Final away runs*: the number of runs scored by the away team. NB: in American sports the match is scheduled as away vs home team, so that SEA vs HOU is played at Houston
- *Final home runs*: the number of runs scored by the home team, in our case it will always be the Astros
- *Inning*: the inning in which the at bat takes place, 1-9 with possible extras

- *Top bottom*: a dummy to identify whether it's the top or bottom of the inning (not really useful here as home team bats always at the bottom)
- *Batter*: the name of the batter at the plate
- *At bat event*: the result of the at bat. Ex. Home run, walk, strike out and others
- *Pitch type*: the kind of pitch thrown at the batter, on a 2-digit form. Ex SI = sinker, SL = slider, FF = four seam fastball (These are all common standards in baseball analysis, although we won't dwell on them in this thesis)
- *Pitch category*: as before, just a classification of the pitch as either a fastball, FB, a breaking ball, BR, or a changeup, CH. Easier to use then types, can lead to some misunderstandings
- *Has bangs*: a dummy, y = yes if the at bat has bangs, n = no if not
- *Bangs*: counts the number of heard bangs in the event. Ex. 1B is a sole bang, 2B means two bangs and so on.
- *Call code*: a 1-letter coding of event result. Related to game-scoring, not that useful for us
- *Description*: a written resume of event result
- *On 1b/2b/3b*: a dummy, signals the presence (t = true) or absence (f = false) of a runner on first, second and/or third base. NB: it is different from the bang counter as 1B and on.
- *YouTube id*: the id for each game video on YT
- *Pitch YouTube seconds*: counts the number of seconds from the start of the video to when the pitch is thrown
- *YouTube url*: link to YT video
- *Pitch date time*: time frame of the thrown pitch
- *Game pitch id*: an id code for each thrown pitch in a game
- *Event number*: a counting number of events. NB: it never starts at 1 as the home team bats at the bottom of the inning
- *Pitch play id*: an id code for each pitch result
- *At bat play id*: an id code for each at bat
- *Away team id*: an id code for the away team
- *Home team id*: an id code for the home team, in our case always 117 = Houston Astros

Now, we will only consider a few of the columns for the analysis that will be made so there's no need to panic. Before getting in on the act, let's consider what others already did using this dataset.

b. The road so far: recap of other studies and analysis on the dataset

Just after the dataset was made available to everyone through Adams' website statisticians and baseball fans went on a spree and started delving on the numbers to see the effects that cheating had on the Astros performance.

Many of those studies can be found on the same website, that is also a compiler of analysis done on the subject, but our point of focus will be the work done on two of the most famous specialized oasis in the net as far as baseball is concerned: FanGraphs and The Ringer.

Starting from the latter, Ben Lindbergh wrote an article in which he explores the effects of the trash-banging scheme in terms of team-wide plate discipline, as ability to draw base on balls and avoid strikeouts by not swinging out of the zone and making contact into it, and then looked at the possibility of the Astros having cheated in the postseason.

While he doesn't use the dataset, this article states the point of focus of many studies done after the dataset came out: understanding how the scheme affected the Astros' plate discipline.

What he discovered is that the improvements the Astros made in terms of avoiding the K were staggering:" The Astros slashed their strikeout rate dramatically between 2016 and 2017. Astros nonpitchers struck out in 23.4 percent of their plate appearances in 2016, the league's fourth-most-frequent rate. In 2017, the league-wide nonpitcher strikeout rate rose by 0.6 percentage points, but the Astros' K-rate dropped precipitously to 17.2 percent, the lowest in the league" (Lindbergh,2019).

But then, there are some questions on whether that gain was only due to cheating: as stated by both the author and Jeff Sullivan, now working in the Tampa Bay front office, there's also to consider that in the 2016 offseason the Astros made a conscious effort to drop some strikeout prone guys in favour of more contact (Sullivan, 2017).

The interesting point Lindbergh makes is related to the 2017 postseason: the Astros had a 230 points difference in OPS (On Base plus Slugging Percentage) from home, .862, to road, .632, and this with Minute Maid Park being somewhat of pitcher friendly.

While, as he states "the statistical case is less compelling than it would be if sign-stealing made hitting as simple as it seems like it should" (Lindbergh, 2019), the fact that some of these numbers popped out without the access to the dataset led to further studies and to some deeper understandings of the issues at hand.

Enter Jake Mailhot, one of the young guns at FanGraphs, arguably the best site dedicated entirely to the statistical side of baseball in the net.

In a series of articles from November 2019 to February 2020 he first widened the scope of Lindbergh's findings on team plate discipline and then, exploiting Adams' dataset, focused on the effect of the scheme on single plays, games, and players.

On his first piece the FanGraphs resident started from the results of Lindbergh and analysed in detail the assumption that the Astros had massive improvements in plate discipline at home due to the cheating.

The results, as shown in the upcoming tables, are clear on that front:

Table 1: plate discipline of Houston Astros batters at home

Astros Plate Discipline At Home			
Year	O-Swing%	Z-Contact%	SwStr%
2016	28.60%	79.15%	13.20%
2017	27.80%	85.85%	9.50%
Change	-0.80%	6.70%	-3.70%
			SOURCE: Baseball Savant

O-Swing% is the percentage of swings to pitches outside the strike zone, while Z-Contact% is the percentage of swings resulting in contact with balls thrown inside the strike zone. SwStr% is the percentage of swings resulting in balls missed by the batter.

Table II: pitch dependent production of Houston batters at home in run values

Astros Pitch Type Run Values At Home			
Year	Fastball	Breaking	Offspeed
Swing Runs			
2016	-83.52	-61.36	-15.12
2017	-44.15	-24.81	-17.2
Change	39.37	36.55	-2.08
Take Runs			
2016	79.96	30.01	23.66
2017	75.23	43.46	21.25
Change	-4.73	13.45	-2.41

Note that run values are calculated as on the possible 24 base/out and count states on the RE288 formula by noted sabermetric Tom Tango (Tango, 2019).

Run value, as defined by Tango in his table related to at bat count and on base situation, is the worth in runs of a specific pitch-result combination, our event for that we collect data. Swing Runs are related to the occurrences in which the batter swings at the pitch: given that baseball is a sport of failures there are many more occasions in which a swing has a bad outcome (miss, out) then a good one (hit), therefore the negative run value total.

On the other hand, Take Runs are related to the occurrences in which the batter doesn't swing, rather gets the ball go by. This usually happens when the pitch is not in the strike zone, so that the count situation (balls and strikes) improves for the batter and increases the possibility of scoring, gaining therefore run value, and that can lead to runs or free bases (BB, our bases on balls).

What this table says is that, through the cheating scheme, the Astros became much more aggressive on fastballs in general, improving greatly on Swing Runs while taking less, fared better in all facets against breaking balls, as they swung more when in the zone and let them go when outside, and had some backlash in the offspeed category. Knowing what was going to come led to different swing choices and patterns.

Table III: zone dependent production of Houston batters at home in run values

Astros Zone Run Values At Home		
Year	In Zone	Out of Zone
Swing Runs		
2016	-53.8	-77.82
2017	20.8	-75.08
Change	74.6	2.74
Take Runs		
2016	-114.56	214.33
2017	-107.37	214.99
Change	7.19	0.66
SOURCE: Baseball Savant		

This table enforces the aforementioned arguments: the Astros greatly improved their results by swinging more on pitches in the zone with positive outcomes, causing also less strikes taken in and better out of the zone discipline.

There are great progresses in overall discipline from 2016 to 2017: less swings at pitches out of the strike zone (O-Swing% decrease) and more swings at pitches inside it (Z-Swing% increase) all while cutting the missed balls on swings in total (SwStr% decrease).

Then the focus shifts on pitch category, as we had defined on the Adams dataset as either fastball, breaking ball or offspeed (change): from 2016 to 2017 the Astros improved severely on their performance versus FBs and BRs, while getting slightly worse against CHs.

This advanced proficiency on hitting fastballs and breaking balls and swinging at more pitches in the zone is confirmed on the third table: a 75 point increase in run value on pitches in the zone pairs almost perfectly with the sum of increases in run values for FBs and BRs, a clear sign that maybe the real upside of the cheating has been the ability to adjust. As Mailhot writes: "Knowing whether a pitch is a fastball, or a breaking ball, is far more important because of the way those types of pitches move. If a fastball is signalled, the batter can gear up for heat and make sure they put a good swing on a relatively straight pitch. If a breaking ball is signalled, they can focus on the movement of the pitch to remove the deception it would otherwise create" (Mailhot, 2019).

Then the author tried to perform the same analysis on road stats and, while not so prominent, some parallels came up on the swing tendencies and run values on zone/pitch categories, so that the assumption of Sullivan, that roster changes made a difference for the Astros' plate discipline, is not debunked.

On his second essay, Mailhot took again the run values on pitch categories and analysed home/away differences on several players in the Astros roster.

While not finding a lot of conclusive evidence, a pattern was somewhat predominant: take for example the run values on pitch categories for then 2017 MVP Jose Altuve:

Table IV: Jose Altuve's production in terms of run values for pitch type on set time intervals

Jose Altuvé Pitch Type Run Values 2017

José Altuve	Fastball	Breaking	Offspeed
Swing Runs			
Home - Pre 5/19	-1.34	-5.61	-6.5
Home - Post 5/19	-1.23	1.07	4.24
Away - Post 5/19	0.83	2.15	3.5
Take Runs			
Home - Pre 5/19	0.11	-0.19	1.05
Home - Post 5/19	1.53	0.77	3.16
Away - Post 5/19	2.05	0.94	0.92

SOURCE: Baseball Savant

Pitch Values scaled to per 100 pitches.

From Mailhot, 2019

After the 5/19, a timeframe that Mailhot took from the audio scans done by Baseball Prospectus writer Rob Arthur (Arthur, 2019) that saw that date as starting point of the cheating scheme, there are stark improvements in the prowess on both swing runs and take runs against BRs and CHs, and that is the same for the majority of Astros hitters.

Knowing if a ball is going to come straight, zoom away or dive down had quite the effect on swing choices, with players ready to punish pitches in the zone and lay off bad offerings out of the strike zone.

What remains true is that those progresses present themselves also on the road, where the scheme couldn't be set in motion, so while the possibility that the majority of the differences was due the cheating the same could be said about a drastic change of mentality for all hitters with sounding results.

On his last effort on the matter, dated February 2020, Mailhot had finally the chance of employing the Adams dataset on his findings and found out that, while the accuracy of bangs-pitch category was a good 93% the remaining 7% of mismatches was really detrimental to the cause (Mailhot, 2020).

It seems counterintuitive but the real gains were made on FBs that usually had no bang as a signal: Mailhot concours that maybe cut-fastballs, tailing in on the batter and not straight but classified as FBs, were signalled as breakers to explain the fact.

He then tried to look at the weight of the bangs in terms of relation to the outcome of the game, by using Adams' information on single pitches as events and relating to each its own Leverage Index (LI), a proxy of the importance of an event on the final result of the game, but what he found was mind-boggling:

Table V: banging at bats as per leverage index and run value

Astros Banging Scheme Leverage Index			
Leverage	At-bats	%	Cumulative Run Value
Low Leverage	308	49%	1.94
Medium Leverage	262	42%	1.97
High Leverage	53	9%	-2.00

From Mailhot, 2020

Cumulative run value is none other than the sum of all run values of each outcome related to the at bat situation. Leverage is a measure of the specific weight of an at bat in regards to the "destiny" of a game: a batting turn on a 15-0 game is almost not consequential to the sorts of the match, while a bases loaded at bat on a tied game in the last inning is quite the deciding moment. More on LI (Leverage Index) on (Slowinski, 2010) and correlated.

Somehow the Astros had more success with bangs in low and medium leverage situations than in big spots, and no player showed particular tendencies, so that while Yuli Gurriel simply loved hearing bangs no matter the situations, Marwin Gonzalez on the other side found no help whatsoever (Mailhot, 2020). Moreover, just a single game, versus the Toronto Blue Jays, was won thanks to a bang on a determinant at bat.

The real fact that Mailhot points out is that the Astros banged the can no matter what, was it a useless out or a key matchup. The cheating was there, no matter what.

c. Our game: what else can be found on the data?

Previous studies have shown some clear results: the Astros experienced great progresses in all metrics related on plate discipline, either contact on zone or lay off out of it and better damaging capacity on both fast and breaking pitches.

Questions remain on what these progresses are due to: is it only the banging scheme or there's also the different roster construction to account for?

Then again, what else can be done with the Adams dataset?

The real riddle bothering fans and the whole baseball community is whether the Astros could have won it all without employing the scheme or if it was really that productive and profitable that it turned out to be indispensable for the Houston side on its journey to the WS.

Now, as per baseball nature itself as a sequence of random events, there's no clear answer: declaring that yes, the Astros deserved or no, the cheating won it for them, would be nonsensical. Think about it: what guarantees you that without an event there wouldn't have been the same consequences? Who tells you that, if Will Harris doesn't throw that cutter to Howie Kendrick then the Astros win again in 2019? Maybe they do, or maybe Harris throws a fastball that gets hit the same and Washington wins!

The only thing that can be done is rely on some data and then trust what happens on a majestic number of iterations of the same procedure: if a coin is thrown three times and always comes tails, then is the probability of getting tails really 100%? No, it's the basics: if we throw that coin, mind you a good one not a one-dollar magic trick one, more times then we'll get a close 50/50 split as we increase the number of throws.

Good news is we have the data to make some assumption and, while not getting the answer itself, we can explore some possible occurrences on the Astros run to the WS: first there's the player stats for the "tainted" 2017 season, that will act as the starting point of the analysis, and then we have the Adams dataset, which can be considered as a subset of those stats with the added regressor called Bangs.

What we'll do next is first filtering the 2017 stats through the Bangs dataset to cleanse the player performances of the cheating stain, and then we'll throw the coin thousands of times. We are going to do what every fan of a non-WS winner wants to do at the end of the road: replay the season.

IV. What if we didn't? Exploring possible scenarios for the 2017 Astros

Anything can happen if we consider one and only baseball season.

Projections are done before every season starts, records are simulated and player statlines are spit out of some algorithms as to account for past successes and future declines. What happens every single year is that, while some teams slump horribly, ending under 20 wins from their projected ranking, others outperform and find themselves in the playoffs. The same can be said for players: stars could finally come down given the aches of experience and injuries, whereas new young talents could burst out of the gates and carry unexpected franchises to the promised land.

Doing an analysis on a single season such as the 2017 one has the potential to be an enormous waste of time: leaving fate, and randomness, to decide who wins, who loses and what players go up, down or sideways on trades is giving too much to obtain no actual useful information.

In the forthcoming chapter the 2017 Astros season is going to happen again, thousands of times, versus the same opponents but accounting for changes in the Houston side as to penalize them for the trash-banging system.

The key to the whole simulation method is related to the nature of a baseball season: it is nothing more than a series of 1 vs 1 matches between teams, for 162 games in the regular season and more in the eventual playoffs. Now think about each team not as a group, rather a single entity: each match is a pairwise encounter, and there are many ways in which we can find out the probability of one team beating the other!

The chosen one for this essay is the Bradley-Terry model. Introduced in 1952 by the two statisticians in the journal *Biometrika* (Bradley and Terry, 1952) and based on previous studies by Zermelo (Zermelo, 1929), it considers each probability corresponding to pairwise encounters as a ratio between scores, which are real and positive values usually found through ML estimation (Hunter, 2004) that sum up each team/player/entity's ability and strength.

What we are going to do is the opposite of what is usually done with the B-T model: instead of estimating the scores we are going to fix them, one for each team, and then run a simulation of a season where each match has a probability of either team winning equal to the ratio as defined by Bradley and Terry.

The whole process is nothing fancy: a rather usual Monte Carlo simulation will be performed on four different scenarios corresponding to four different scores for the Astros team, each one linked to the datasets in use and to some kind of punishment to give to the original Houston strength of the team. A thousand simulated seasons for every scenario are to be set in motion, considering all other teams' scores as fixed and equal for all scenarios.

This is not the first time the B-T model finds its way into people's pastimes: Bradley himself tried to use his model on sports (Bradley, 1976). More recently the B-T model has had several applications in the betting environment, as it has been used to provide forecasts of tennis matches (McHale and Morton, 2011) or to model specific tournament seasons in team sports such as basketball and football (Cattelan et al., 2012).

In the baseball field the first application of the B-T model was done in the early 90's (Barry and Hartigan, 1993), although the most important progresses on the matter have been achieved by noted sport statistician Jim Albert. He provided the B-T and other model's possible relationships with sports (Albert et al., 2005), and along fellow Italian Max Marchi, who studied in Bologna and worked for MLB's own Cleveland Indians, wrote what is a must read for baseball programming (Albert and Marchi, 2013). We'll follow the model as in (Albert and Marchi, 2018), making some subtle changes to R functions on calendar of matches.

What about the scores? While in other sports it's rather difficult to conceive a figure that explains the strength of a player, and of a team as a sum, this is ancient history for baseball.

In our paper each team's score will be the sum of its players' performances as defined by the king of all diamond statistics, the *omnia ratio* that dominated the scene for a decade: Wins Above Replacement, or WAR.

In layman's terms: baseball is about scoring runs, similar to goals in football or shots in basketball. Each player contributes a certain amount of runs to his team by playing the field, and he has many ways to accrue value in terms of runs: he can either use his bat to score them, impede the opponent to score by not letting him get hits on the mound or in the field through defence, and take some advantages on the opponent by running the bases.

Summing up each player's value in runs we obtain a team's value in runs. That gets then compared to the value in runs of a benchmark, a team of replacement level players, which are the ones without a contract or in the minor leagues. The remaining value, so called runs above replacement, gets divided by the average number of runs needed to win a game in that season to obtain WAR.

Think about WAR as one of the thousands of test scores used in psychology (Lyman, 1998): instead of studying responses to questions and stimuli to infer on a subject's personality we'll use on field responses as specific results to infer on each player's strength with respect to an hypothetic replacement level player.

The actual procedure is much more complicated, and we'll get through it, but as a reminder we'll only focus on the hitting side of baseball, on those players that handle the bats and try to outhit the opposition, as they are the ones to that the scandal is related to.

In the first part of the Chapter we'll introduce the method in more detail, from the intricacies of relating the B-T model to baseball to the hardships of calculating WAR and employing it as cornerstone of all simulations.

Only then we'll go through the scenarios: in each one of them the Astros team value has been hurt by using the Bangs dataset in conjunction with the complete season results, taking away the whole subset, or just a part of it, from the lot, or by considering previous seasons' tendencies and outcomes and projecting them to the 2017 season as to avoid the interference of cheating.

To each scenario will correspond different players' values, that by reflex are going to change the team value and all the probabilities as defined by B-T.

For every landscape we then proceed with the simulation of 1000 seasons, ending our marathon with results related to Win-Loss records, regular season performances and playoff appearances, and as we'll see, the differences are going to be stark but the consequences not as much so: the 2017 Houston Astros were simply a particularly strong team, and yet, to our soulful regret, they cheated.

a. Let's play two: a method for madness

This essay is trying to answer an almost impossible question: could the Astros 2017 season have gone the same way without them cheating? While this point is hard to prove, either accepting it or refusing it are not the right thing to do, what can be done is replay the same season accounting for some changes and then crunch some numbers and search for some light at the end of the tunnel.

This journey goes through two steps: evaluating and performing modifications on player performances and only then reimagine a 2017 season given the aforementioned shenanigans.

The proxy this study will use to evaluate players performance is WAR, Wins Above Replacement, maybe the most used and abused omnia ratio in sabermetrics. While its construction will be discussed later on, what is going to be the main focus is how the WAR value modifications on each player add up given different scenarios, that we'll be somewhat harmful on the Astros team WAR, to then end up with a number of team WAR for each considered possibility.

Now, WAR would have been the crème of the crop in the early 2010s: those were the days after the Moneyball, were different dimensions on each player's contribution were only starting to be explored.

Defense, baserunning and other shades of the game previously left to be valued by hard round numbers were incorporated and added to hitting and pitching prowess to form the foundations of WAR.

Years have gone by, thousands of different WARs have seen the light of the day, but mostly technology has made its way to the diamond in a rather bold manner: nowadays we are able to know specifics for each ball thrown and hit, for each movement in the field of play and for all possible situations baseball-wise.

Research is more and more specific, related to minimal traits, tendencies, abilities that we previously ignored that even existed, so that a statistic that bunches up and returns a sole figure is somewhat archaic.

Nevertheless it's still employed as of today to consider both a player's contribution to a team, and retains some importance on the evaluation when it's time to sign contracts or trade for young prospects, so it's going to be the choice in terms of the Astros team strength, by adding up the WAR produced by all its players.

Then this team WAR is going to be cut and sliced as to punish Houston for its mischievous ways: a set of scenarios is going to be considered, each with its own repercussions on team WAR given that some negative elements are to be added to statlines or vice versa, some positive numbers are to be lowered.

But how is WAR tied to season simulations? That's where a time leap back to the 50s is needed.

In 1952 two statisticians, namely R. A. Bradley and M. E. Terry, elaborated a homonymous model to predict the outcome of comparisons between two elements, and well, it suits baseball and sports in general really well (Bradley and Terry, 1952). Thinking about it a baseball season is nothing more than a bunch of pairwise comparisons, of encounters between two teams, so this model is apt to our target.

The Bradley-Terry model is built on the fact that, given two elements from the same populations, we can estimate the probabilities of the pairwise comparisons for that two elements as the ratio between an element's score and the sum of the two elements' scores. Considering elements i and j and p(i) and p(j) as positive real-valued scores we can write the probability P (i > j) as

$$P(i > j) = \frac{p_i}{p_i + p_j}$$

How can the scores be found? There's more than a single avenue: if we have the results of the comparisons but not the scores it is possible to proceed using those as weights for a ML estimation on given comparison (ex. i > j), but that's not our case.

What we are going to do is the exact opposite: scores are going to be set by us and then comparisons, and seasons as sums of them, are going to be performed in an iterative algorithm. As much as I want to be a trailblazer, this idea was first introduced and then refined by Jim Albert and fellow countryman Max Marchi back in 2013 (Albert and Marchi, 2013).

They took the Bradley-Terry model and set as scores a Talent variable which was a random normal on set standard deviation as to give each team an equal shot to greatness, and then wrote an algorithm that simulated one and more seasons given the talent vector on a sequence of Bradley-Terry iterations. The result was a series of season where to each teams corresponded a random talent, and where a sequence of coin tosses weighted as per the probability defined by B-T were performed.

That means that more talented teams had greater chances of winning games, division and WS, but the talent of each team was a spin of the wheel.

This assumption, while good in a sense that each baseball season is its own world, led me to think that there must be a solid measure of talent that doesn't sway much year to year and points to the fact that team A is better than team B, and that sent me to WAR, in a good sense.

While it isn't to say that the team with the highest WAR always wins, the fact is that most great teams on paper end up at least to the playoffs, then in October it's anybody's game. Keep in mind, great teams will always have high WAR totals as the stat is performance-related, but a roster of players with previous high WAR doesn't equal sure-fire wins: injuries, slumps and outliers are real, so that predictive models nowadays are based on WAR and other peripherals but can account for unforeseen events.

Each team's WAR is going to be set as its score, so that probabilities as B-T model are defined and seasons replayed.

At the end we'll have a series of scenarios for possible 2017 Astros situations, and to each scenario will correspond a different Astros team WAR that will lead to changes in the probabilities for each game as pairwise encounter. Finally, we'll check for significant changes in the paths to the WS, in terms of regular season records, division and league wins and World Series triumphs.

Before getting in on the act, it's customary to provide a brief introduction to WAR, consider the statistics that are going to change due to different scenarios and the underlying matters that the employment of WAR brings to the table.

b. Stats: changes and considerations on players' performances

The key component on the first step of our analysis is the evaluation of players' performances and how potential changes of them are reflected in a numerical way.

To do so we'll employ what has been the statistic for a decade, a well-known construct that made its way onto front offices, diamonds and stands: Wins Above Replacement, commonly referred to as WAR.

Born thanks to the minds of Tom Tango, a statistician involved into all kinds of sports but baseball foremost, and Sean Smith, a professional in the sabermetric field working for Baseball Reference at the time, WAR saw its foundation in 2007 and has been constantly rewritten, redeveloped and modified as to keep up with the technological wave that crashed on baseball (Tango, 2007; Smith, 2010).

Nowadays each team and professional website has its own formula for WAR: Fangraphs has fWAR, Baseball Reference has bWAR and Baseball Prospectus relies on VORP, Value Over Replacement Player.

That's to say that as a statistic, WAR can be dealt in so many ways given its composition into a multitude of different pieces, adding up as those of a puzzle to show the contribution of a player in each facet of the game. To understand how WAR works there are two avenues to admire.

First, we need to define the contribution of the player, or his *Player Runs* as the authors called it. As we previewed, changes are going to be made on the hitter side, so we'll refer to the classic WAR for hitters while taking that of pitchers as a constant on each iteration/season.

Being a hitter in MLB is not easy at all, not only because hitting a baseball coming at north of 90 mph is arguably the hardest task in all of sports, but also because there's much more to do in the field: a hitter needs to have a good bat, then he must be able to defend his position on the diamond, whether it means catching flyballs on the outfield or snagging groundballs in the infield or else, but he needs also to provide threat on the bases, gaining an extra 90 feet or stealing on opposing pitchers.

WAR considers all of this nooks and crannies of the game while also weighting the overall performances for the league in which the hitter plays and his position, giving a bonus in case of hard tasks (catcher, shortstop, center fielder) or shedding away some value for easier roles (first base, left field, designated hitter).

Therefore, we can consider:

Player Runs = Batting Runs + Fielding Runs + Baserunning Runs + Positional/League Adjustment

And now a disclaimer: the only aspect we'll change are the *Batting Runs*. For each scenario will correspond a different value that will then impact *Player Runs* and WAR. Other components, those related to fielding and baserunning and the adjustments, are to be considered as fixed: in layman's terms we'll consider a number of seasons in which each player runs and defends the same position in the same league and with the same prowess, only having differences in his performance at the plate.

Batting Runs are based on each player's wOBA, weighted On Base Average, a measure of his aptitude at the plate that weights each event for the number of bases it creates: for example if a single is weighted 1 then a homerun is 4 times more valuable (but also much harder to get in real life if you're not named Joey Gallo). All it's needed to calculate wOBA is a player's standard stat line: At Bats (AB), Plate Appearances (PA), singles (1B), doubles (2B), triples (3B), homeruns (HR), bases on balls (BB, of which IBB if intentional), hits by pitches (HBP) and sacrifice flies (SF). Then each round number gets multiplied for a season-specific constant and the sum of them gets divided by the number of "pure" ABs

For example, on the 2013 season:

$$wOBA = (0.690 \times uBB + 0.722 \times HBP + 0.888 \times 1B + 1.271 \times 2B + 1.616 \times 3B + 2.101 \times HR) / (AB + BB - IBB + SF + HBP)$$

From Slowinski,2010

Obviously, there will be different constants for the 2017 season, although the procedure is the same. The next step is to go from wOBA to wRAA, weighted Runs Above Average, a simplified version of the *Batting Runs*:

$$((wOBA - League\ wOBA)/wOBA\ Scale)*PA = wRAA$$

wRAA is none other than a scaled version of the residual wOBA of a player with respect to the league figure, multiplied then by the number of PA for that player. To get to the Batting Runs there's also to consider the league (AL or NL) and the park in which the hitter plays, so:

$$Batting\ Runs = wRAA + (lgR/PA - (PF*lgR/PA))*PA + (lgR/PA - (AL\ or\ NL\ non\text{-}pitcher\ wRC/PA))*PA$$

Where lgR/PA is the ratio runs/PA for the league, PF is the Park Factor on base 100 and wRC/PA is similar to R/PA considering a weighted measure of runs created.

Getting to our *Batting Runs* is the first big chunk of work, considering that all other run components are set in stone, fielding and baserunning from the "tainted" Astros 2017 and adjustments as same for each team in that season. Then we need to finally get to the WAR

WAR = (Player Runs + Replacement Level Runs)/RPW

Our player gets then compared with a replacement, a proxy of either a minor league free agent or a bench player in MLB, that has a certain number of runs at his credit:

$$Replacement\ Level\ Runs = (570 * (MLB\ Games/2{,}430)) * (Runs\ Per\ Win/lgPA) * PA$$

As per standard, the replacement level is set at 1000 WAR per 2430 games, 57% of it to hitters. As we can see it is also related to the denominator of WAR, so called RPW or Runs Per Win, which is the season-specific figure of runs needed to win a game:

$$RPW = 9*(MLB\ Runs\ Scored\ /\ MLB\ Innings\ Pitched)*1.5 + 3$$

And so, we finally got to WAR…but what are we going to do with it? It's going to be a straightforward operation of subtle changes: given different scenarios that involve bangs and the cheating scheme, we are going to perform slight modifications on players' statlines, working on the number of ABs and/or PAs and on the other basic round stats we encountered before, 1/2/3B and HR foremost.

Those little subtractions will trickle down as drops of rain on a window: first, wOBA is changing and therefore wRAA and Batting Runs, consequentially Player Runs are to be hurt and WAR at the end of the road.

It seems like a real pain to calculate, but thankfully there are some well-equipped WAR calculators on the net. For this analysis we will use a retooled version of the Weinberg calculator, adding information on Park Factors and league ratios (Weinberg, 2013).

Let's see for example Jose Altuve's 2017 season, for that the Houston player won the AL MVP over NY Yankee Aaron Judge.

Some notes on the changes with respect to the original Weinberg version:

- Constants have been updated to the 2017 ones,
- Both the Rep level and Batting have been considered first as Runs and at the end as wins dividing them by RPW in the WAR formula,
- For the Rep level we consider the unified formula as proposed by FanGraphs and Baseball Reference (Cameron, 2013),
- For the Bat we consider the complete formula as the FanGraphs one, by adding a park component related to PF (Park Factor, the 5-year running as basic) and a league component based on the wRC/PA ratio for the American League (AL),
- Positional adjustment is done with respect to the classic FanGraphs (Slowinski, 2010) edition but by considering games and not innings played, although the final discrepancies are not that heavy,
- Fielding and Baserunning runs are straight from FanGraphs,
- We added the League Adjustment, based on the ratio between the sum of AL run components and AL PAs,

Table VI: WAR calculator

WAR calculator, 2017 edition						
Player Statline						
	Inputs		**2017 constants**			
BB	58		BB	0,693	Lg wOBA	0,321
IBB	3		IBB		wOBA Scale	1,185
HBP	9		HBP	0,723	R/W	10,048
1B	137		1B	0,877	Lg PA	185295
2B	39		2B	1,232	Lg R	22582
3B	4		3B	1,552	PF	0,97
HR	24		HR	1,98	AL WRC	11374
AB	590				AL PA	92622
SF	4					
PA	662		**League Runs 2017**		**Pos.Adjustments 2017**	
G	153		Bat	-89,1	C	11,8
			Bsr	-8,3	1B	-11,8
Rep Runs	20,46		Fld	20,9	2B	2,4
wOBA	0,405		Pos	-226,1	SS	7,1
wRAA	47,0				3B	2,4
Bat Runs	48,8				LF	-7,1
Pos	2,4				CF	2,4
Fld Runs	-1,1				RF	-7,1
BsR	4				DH	-16,5
League	2,2					
WAR	**7,6**					

The only input we needed was the basic statline of Altuve, along with his Baserunning and Fielding Runs.

We will repeat the same for every player considering different changes, so that on every scenario will correspond a different line and at the end different WAR values.

The sum of those values for hitters in the Astros 2017 roster and of the WAR provided by the pitching side, considered as fixed so that Houston always trades for Verlander and such, will give us the Astros 2017 team WAR for the specific scenario, that, in conjunction with the opposing teams' WAR, will be the cornerstone of the probabilities as defined by the B-T model and with them we are going to run thousands of 2017 seasons to see if winning was really a question of cheating.

c. Scenarios: a breakdown

Our analysis needs a starting point, that is obviously the tainted 2017 Astros season comprehensive of all the occurrences of the trash-banging cheating scheme.

Proceeding by steps, we first have to consider each player's performance in the aforementioned season, to then transform it in a player-specific WAR that added with all others returns the team WAR that will be the cornerstone of the simulated season.

On the original 2017 season the Astros roster performed as follows:

Table VII: 2017 Astros hitters' statlines and WAR values, cheating season

Batter	G	PA	AB	1B	2B	3B	HR	BB	SF	HBP	IBB	WAR
Jose Altuve	153	662	590	137	39	4	24	58	4	9	3	7,6
Alex Bregman	155	626	556	95	39	5	19	55	7	7	2	3,5
Carlos Correa	109	481	422	83	25	1	24	53	4	2	5	5,1
George Springer	140	629	548	92	29	0	34	64	4	11	1	4,5
Marwin Gonzalez	134	515	455	81	34	0	23	49	2	6	4	4,0
Yuli Gurriel	139	564	529	96	43	1	18	22	6	7	1	1,7
Josh Reddick	134	540	477	99	34	4	13	43	12	0	1	3,4
Carlos Beltran	129	509	467	65	29	0	14	33	6	3	3	-1,1
Brian McCann	97	399	349	53	12	1	18	38	5	7	3	2,6
Jake Marisnick	106	259	230	30	10	0	16	20	1	6	1	1,4
Evan Gattis	84	325	300	45	22	0	12	18	3	4	0	1,4
Derek Fisher	53	166	146	21	4	1	5	17	0	3	1	0,0
Tyler White	22	67	61	8	6	0	3	4	1	1	0	0,2
JD Davis	25	68	62	6	4	0	4	4	1	1	0	0,1
Cameron Maybin	21	63	59	5	1	1	4	3	0	0	0	0,0
Nori Aoki	71	224	202	40	12	1	2	15	4	2	1	-0,2
Juan Centeno	22	57	52	10	0	0	2	4	0	0	1	-0,3
Max Stassi	14	31	24	1	1	0	2	6	1	0	0	0,2

Some important notes: WAR is considered as fWAR as data is taken from FanGraphs, moreover only players with a feasible number of PAs and bangs are considered, so that the roster is not complete (some exceptions are Colin Moran, Tony Kemp and some PAs by pitchers) and the team batting WAR is not the actual figure but really close to it.

All the abbreviations have already been introduced, the only important number in this table is the player WAR: Altuve won the MVP that season with a 7.6 WAR that denotes a superstar-like performance (Slowinski, 2010), while young talents such as Bregman and Correa brought 3+ WAR contributions and resurgent old timers as Reddick and McCann saw their best season in ages. The team WAR for the original 2017 Astros season stands at a considerate 53,1: how was it in comparison with the other teams?

Table VIII: 2017 teams' value in terms of WAR, cheating season

NL East	**WAR**		**AL East**	**WAR**
ATL	24,8		BAL	19,9
MIA	33,4		BOS	41,4
NYM	31,3		NYA	52,6
PHI	24,7		TBR	36,7
WAS	47,4		TOR	27,3
NL Cent			**AL Cent**	
CHC	43,5		CHW	18,5
CIN	24,7		CLE	57,1
MIL	34,4		DET	25,7
PIT	26,6		KCR	24,8
SLN	42,0		MIN	34,9
NL West			**AL West**	
ARI	42,6		OAK	27,1
COL	32,3		HOU	53,1
LAD	55,0		TEX	21,3
SDP	14,9		LAA	29,8
SFG	21,9		SEA	30,9

East, Central and West are the three divisions in each league. 40+ WAR is considered to be a proxy of a playoff bound team.

Pretty good! The AL West, Houston's division, is notably weak, with Seattle as closest competitor on a 20+ WAR difference, which is to say that the Mariners needed two Mike Trouts, the best player in the game, to catch up. An interesting detail: the Astros didn't have the biggest WAR in the game, rather they were third behind Cleveland and the LA Dodgers, to say that in a sport as long and as streaky as baseball the best team doesn't always win.

Now we have all we need to start our simulation. The procedure will be performed in a similar way as that of Albert and Marchi's book, adapted for the 2017 season with some additions done mirroring Albert's work on the abbreviated 2020 pandemic season (Albert, 2020).

At first, we'll proceed by building a calendar of matches for the regular season, as team vs team encounters on a structure similar to the real life one so that each team plays 162 games, with small changes due to the complexity of the actual MLB schedule.

Then we'll calculate the foundation of each game, the basics that decide who wins and who loses: pairwise encounters win probabilities based on the B-T model and on the vector of WAR team values, with the one for the Astros strictly related to the scenario we'll be considering.

Only with those we can let the wheel spin and play the season, treating each game as a coin toss on set probabilities. After a lot of pennies launched in artificial air, we'll have some regular season standings with win-loss records to deal with.

By rearranging the standings based on those records we'll assign some accolades, such as division and league winners, and set the postseason scenery and ultimately the WS matchup, the final series of games between the best teams in both leagues.

Lastly, the trophy for that particular season we'll be assigned by one of statistician's best friends, a multinomial on seven tries using the aforementioned probabilities.

This whole process is going to be done on a loop for a thousand repetitions per scenario, and at the end we are going to focus on the results achieved by the Houston Astros: how many games did they win on average in the hypothetical 1000 season? How many times did they win their division? Were they also successful in the American League? Did they make it to the WS? Did they win it all?

Step 1 is the composition of a schedule of games that combined form a season:

```
> library(dplyr)
> library(tidyverse)
> make.schedule<-function(teams,k){
   n.teams <- length(teams)
   Home <- rep(rep(teams, each=n.teams), k)
   Visitor <- rep(rep(teams, n.teams), k)
   schedule <- tibble(Home = Home,
                      Visitor = Visitor)
   dplyr::filter(schedule, Home != Visitor)
 }
> NL_East <- c("ATL", "MIA", "NYM", "PHI", "WAS")
> NL_Cent <- c("CHC", "CIN", "MIL", "PIT", "SLN")
> NL_West <- c("ARI", "COL", "LAD", "SDP", "SFG")
> AL_East <- c("BAL", "BOS", "NYY", "TBR", "TOR")
> AL_Cent <- c("CHW", "CLE", "DET", "KCR", "MIN")
> AL_West <- c("OAK", "HOU", "TEX", "LAA", "SEA")
> teams <- c(NL_East, NL_Cent, NL_West,
           AL_East, AL_Cent, AL_West)
> league <- c(rep("NL", 15), rep("AL", 15))
> division <- c(rep("NL_East", 5), rep("NL_Cent", 5),
              rep("NL_West", 5), rep("AL_East", 5),
              rep("AL_Cent", 5), rep("AL_West", 5))
> Team_info <- data.frame(Team = teams,
                         League = league,
                         Division = division)
> s1<-make.schedule(NL_East,18)
> s2<-make.schedule(AL_East,18)
> s3<-make.schedule(c(NL_East,AL_East),1)
> s4<-make.schedule(NL_Cent,18)
> s5<-make.schedule(AL_Cent,18)
> s6<-make.schedule(c(NL_Cent,AL_Cent),1)
> s7<-make.schedule(NL_West,18)
> s8<-make.schedule(AL_West,18)
> s9<-make.schedule(c(NL_West,AL_West),1)
> schedule <- rbind(s1, s2, s3, s4, s5, s6,
```

```
              s7, s8, s9)
```

We resorted to the second edition of the book, that introduces the *tidyverse* package on R as a way to make it more like SQL, the famous query language program. On the other hand the function *make.schedule* is exactly the same as Albert's, and so is the construction of division, leagues and team info based on his last work on the matter.

The only notable change we performed involves the actual schedule: as it is composed of 162 games, we needed to simplify the procedure so that of those we have 152 intra-division games and 10 games with the same division on the other league. That's obviously not the real schedule, rather a quick and immediate fix to avoid complications, but it comes at a price: teams such as Houston, Cleveland and the Dodgers will have and astounding number of wins, in the realm of 120+, given that they'll play the majority of their games against much worse opponents.

To make the results as close as possible to a real-life baseball season we can't just take team WAR and use it, as the difference between teams is so big that round numbers would just exasperate the win totals on an almost grotesque way. Therefore, here comes the trick:

```
> war17cheat<-c(24.8,33.4,31.3,24.7,47.4,43.5,24.7,34.4,26.6,42,42.6,32.3,55,14.9,21
.9,19.9,41.4,52.6,36.7,27.3,18.5,57.1,25.7,24.8,34.9,27.1,53.1,21.3,29.8,30.9)
> wars<-war17cheat/mean(war17cheat)
> WAR<-tibble(Team = teams,
             War = wars)
```

First, we built the vector with team WARs but then we transformed each value with respect to the mean WAR of the league. Note that this is not a normalizing procedure: talent is not normally distributed in our humblest opinion, but either way neither the sole WAR figures are enough. Dividing each team WAR by the mean WAR for all baseball returns all positive values on a 0-2 range that is much more malleable and closer to the real world we live in.

From here on the path we followed is that of Albert and Marchi:

```
> SCH <- schedule %>%
   inner_join(WAR, by = c("Home" = "Team")) %>%
   rename(War.Home = War ) %>%
   inner_join(WAR, by = c("Visitor" = "Team")) %>%
   rename(War.Visitor = War)
> SCH %>%
   mutate(prob.Home = exp(War.Home) /
             (exp(War.Home) + exp(War.Visitor))) -> SCH
```

After computing all WARs we built the *schedule* data frame by joining the schedule table to the WAR values as transformed for each team, is it playing at home or on the road. Then we constructed the probabilities as defined by the B-T model, so that each home team has a different probability for each pairwise encounter which is team dependent (note that prob.Visitor is just 1- prob.Home so we don't need it). Note that we chose the classic exponential version of the B-T probabilities.

Now let's play a season:

```
> SCH %>%
   mutate(outcome = rbinom(nrow(.), 1, prob.Home),
          winner = ifelse(outcome,
                          Home, Visitor)) -> SCH
> SCH %>%
```

```
  group_by(winner) %>%
  summarize(Wins = n(), .groups = "drop") %>%
  inner_join(WAR, by = c("winner" = "Team")) ->
  RESULTS
> RESULTS <- inner_join(RESULTS, Team_info,
                        by = c("winner" = "Team"))
```

Each game is nothing other than a flip of a coin in which the probability of the home team to win is defined as we saw before, and that of the visiting team as its reciprocate adding to 1. After all games are performed, we save the results in a homonymous data frame in which all records are grouped by team, to that then we added team information on division and league:

	winner	Wins	War	League	Division
1	ARI	87	1.2772337	NL	NL_West
2	ATL	65	0.7435539	NL	NL_East
3	BAL	53	0.5966420	AL	AL_East
4	BOS	93	1.2412552	AL	AL_East
5	CHC	91	1.3042175	NL	NL_Cent

This is just a snippet of what the *results* table looks like. Note that R allows to reorder the table for whatever column we want and takes the first one, names in our case, as default.

Now that we played the season it's time to focus on the playoff., but first let's award division champions and wild cards:

```
> div <- RESULTS %>%
  mutate(Winner.Div = 0,
         prob = exp(War),
         outcome = sample(nrow(.), prob = prob)) %>%
  arrange(Division, desc(Wins), outcome) %>%
  dplyr::select(-outcome)
> div[c(1, 6, 11, 16, 21, 26), "Winner.Div"] <- 1

> wcard<- div %>%
  mutate(Wild.Card = 0,
         outcome = sample(nrow(.), prob = prob)) %>%
  slice(- c(1, 6, 11, 16, 21, 26)) %>%
  arrange(League, desc(Wins), outcome) %>%
  dplyr::select(-outcome)
> wcard[c(1, 2, 13, 14), "Wild.Card"] <- 1

> div$Wild.Card <- 0
> standings <- rbind(div[c(1, 6, 11, 16, 21, 26), ],
                     wcard) %>%
  arrange(Division, desc(Wins)) %>%
  mutate(Team = winner) %>%
  dplyr::select(Team, League, Division,
                War, Wins,prob,
                Winner.Div, Wild.Card)
> standings
```

While it seems quite the code, it's nothing but a rearrangement: from the *results* data frame we arranged records for division and number of wins, then for all teams in first place for each division, numbered 1,6,11,16,21 and 26 given 5-team divisions, we assigned value one to the dummy "Winner.Div".

After that we took the *div* table, as the *results* one plus the dummy, and did the same procedure as before giving the remaining first two teams for each league the value one to the dummy "Wild.Card" (note that we sliced away the teams that already won the division and only then rearranged the div table), originating a new *wcard* table.

By binding the results of the *wcard* table to the sliced teams on the *div* table, and reordering values for division and wins, we arrive at a *standing* table that, through a classic SQL select command, contains everything we need to know about the regular season we simulated:

	Team	League	Division	War	Wins	prob	Winner.Div	Wild.Card
1	CLE	AL	AL_Cent	1.7119728	112	5.539880	1	0
2	MIN	AL	AL_Cent	1.0463722	86	2.847303	0	1
3	DET	AL	AL_Cent	0.7705377	79	2.160928	0	0
4	KCR	AL	AL_Cent	0.7435539	71	2.103397	0	0
5	CHW	AL	AL_Cent	0.5546672	60	1.741361	0	0

The *standing* table will have 30 rows for the 5 divisions, with information on the team WAR on its transformed value, the number of wins, the probability of winning and the results in terms of division and wild card winner.

So, it's now October, where everybody can win the pennant, but not this time. As everyone knows and loves, playoffs are a world apart from the regular season, where heroes are born, and a Cinderella can go on a run to the World Series. As much as we wanted to give everybody a shot to greatness, we decided to follow the regular season results and award league winners based on RS record, and then make them play a WS on best of seven matchups.

This choice stems for two reasons: simplicity and target. If we had to consider the actual playoff structure, we would have had to build multiple schedules and perform a series of multinomial random encounters to decide the winner for each round up to the WS. That's good, but not as needed in our case: we used regular season WAR as a proxy, and all those beautiful stats can be thrown out of the window in the postseason.

What we are interested is the difference in win-loss record due to different performances as per scenarios, therefore we stuck to this aim and went with it up to the last match. Not only that, we'll end up with a rather straightforward and user-friendly code that can be employed in many other sports and aspects.

On our simulation there are no Wild Card games nor ALDS and NLDS, rather we crown League Champions the two teams with the most wins in baseball, one from the American League and the other from the National League:

```
> win_league <- function(RR) {
     out <- RR %>%
       mutate(Winner.Lg = 0,
              prob = exp(War),
              outcome = sample(nrow(.), prob = prob)) %>%
       arrange(League, desc(Wins), outcome) %>%
       select(-outcome)
     out[1 + c(0, nrow(RR) / 2), "Winner.Lg"] <- 1
     out
   }
> champs<-win_league(standings)
```

The *win.league* function introduced by Albert is nothing but an expansion of the *standings* table to our *champs* table in which we added a further dummy Winner.Lg that is set equal to 1, after rearranging teams for league and record, to the first team in each league. Our reordered *champs* table looks like this:

	Team	League	Division	War	Wins	prob	Winner.Div	Wild.Card	Winner.Lg
1	HOU	AL	AL_West	1.5920448	115	4.913786	1	0	1
2	CLE	AL	AL_Cent	1.7119728	112	5.539880	1	0	0
3	NYY	AL	AL_East	1.5770538	106	4.840673	1	0	0
4	BOS	AL	AL_East	1.2412552	93	3.459954	0	1	0
5	MIN	AL	AL_Cent	1.0463722	86	2.847303	0	1	0

Houston wins the AL! A coincidence indeed, but as we can see there's just an additional column with respect to the previous table for the new dummy. At last let's go to the World Series:

```
> ws_winner <- champs %>%
     filter(Winner.Lg == 1) %>%
     mutate(outcome = rmultinom(1, 7, prob),
            Winner.WS = ifelse(outcome > 3, 1, 0)) %>%
     filter(outcome > 3) %>%
     select(Team, Winner.WS)
> finalstand<-champs %>%
     left_join(ws_winner) %>%
     replace_na(list(Winner.WS = 0)) %>%
     select(- prob)
Joining, by = "Team"
```

All the thrill of a WS match is none other than a multinomial for us…a little disheartening but that's how numbers work. The *ws_winner* table is created by filtering the *champs* table for the league winners and then the result of the WS is decided by a random multinomial on 7 tries for set probabilities where the winner, as Winner.WS, is the team that has an outcome of at least 4 out of 7 tries/games.

Our final standings are then computed on a custom-made table that is just a left join between the *champs* one and the *ws_winner:*

	Team	League	Division	War	Wins	Winner.Div	Wild.Card	Winner.Lg	Winner.WS
30	SLN	NL	NL_Cent	1.2592445	112	1	0	1	1
1	HOU	AL	AL_West	1.5920448	115	1	0	1	0
2	CLE	AL	AL_Cent	1.7119728	112	1	0	0	0
3	NYY	AL	AL_East	1.5770538	106	1	0	0	0
4	BOS	AL	AL_East	1.2412552	93	0	1	0	0
5	MIN	AL	AL_Cent	1.0463722	86	0	1	0	0

On this strange and magnificent simulation the WS matchup is a classic postseason clash between Houston and St.Louis (Pujols on Lidge for those who remember an October moment in baseball history), and the Cardinals win the trophy.

And there we see all the issues in trusting a single simulation: in the scope of a sole season anything (almost) is possible: while Houston winning the AL is understandable given its high team WAR, St.Louis started on the NL with 42 WAR, more than 13 less than the LA Dodgers! It's like having a better record than a team with three more aces on the mound or All Stars on the field, so quite the upset.

That's why we decided to simulate not one, but a thousand 2017 seasons for each scenario, as to have somewhat solid results that mirror the differences in WAR between teams in the same and/or opposite leagues and divisions. To do that we comprised all the previous lines of code in a single function which is WAR dependent:

```
> one.simulation.17 <- function(war){
  require(dplyr)
  make.schedule<-function(teams,k){
    n.teams <- length(teams)
    Home <- rep(rep(teams, each=n.teams), k)
    Visitor <- rep(rep(teams, n.teams), k)
    schedule <- tibble(Home = Home,
                       Visitor = Visitor)
    dplyr::filter(schedule, Home != Visitor)
  }
(…)
  wars<-war/mean(war)
  WAR<-tibble(Team = teams,
              War = wars)
(…)
> finalstand<-champs %>%
     left_join(ws_winner) %>%
     replace_na(list(Winner.WS = 0)) %>%
     select(- prob)
}
```

The *one.simulation.17* function does all the previous steps in automatic, as we only need to provide the vector of WARs as defined by the chosen scenario. To compute 1000 simulations then we only need to loop the function and save the results on a specific data frame:

```
> seasonscheat <- NULL
for(j in 1:1000){
  cheats <- one.simulation.17(war17cheat)
  cheats$Simulation <- j
  seasonscheat <- rbind(seasonscheat, cheats)
}
```

In our case the WAR vector is the one related to the 2017 Astros cheating season, with a HOU team WAR of 53.1. Closing the book on our analysis is a look at the results of our thousand seasons for Houston, in terms of placements, wins and accolades:

```
> counts<- seasonscheat%>%
   group_by(Team)%>%
   count(Winner.Div, Wild.Card, Winner.Lg, Winner.WS)
```

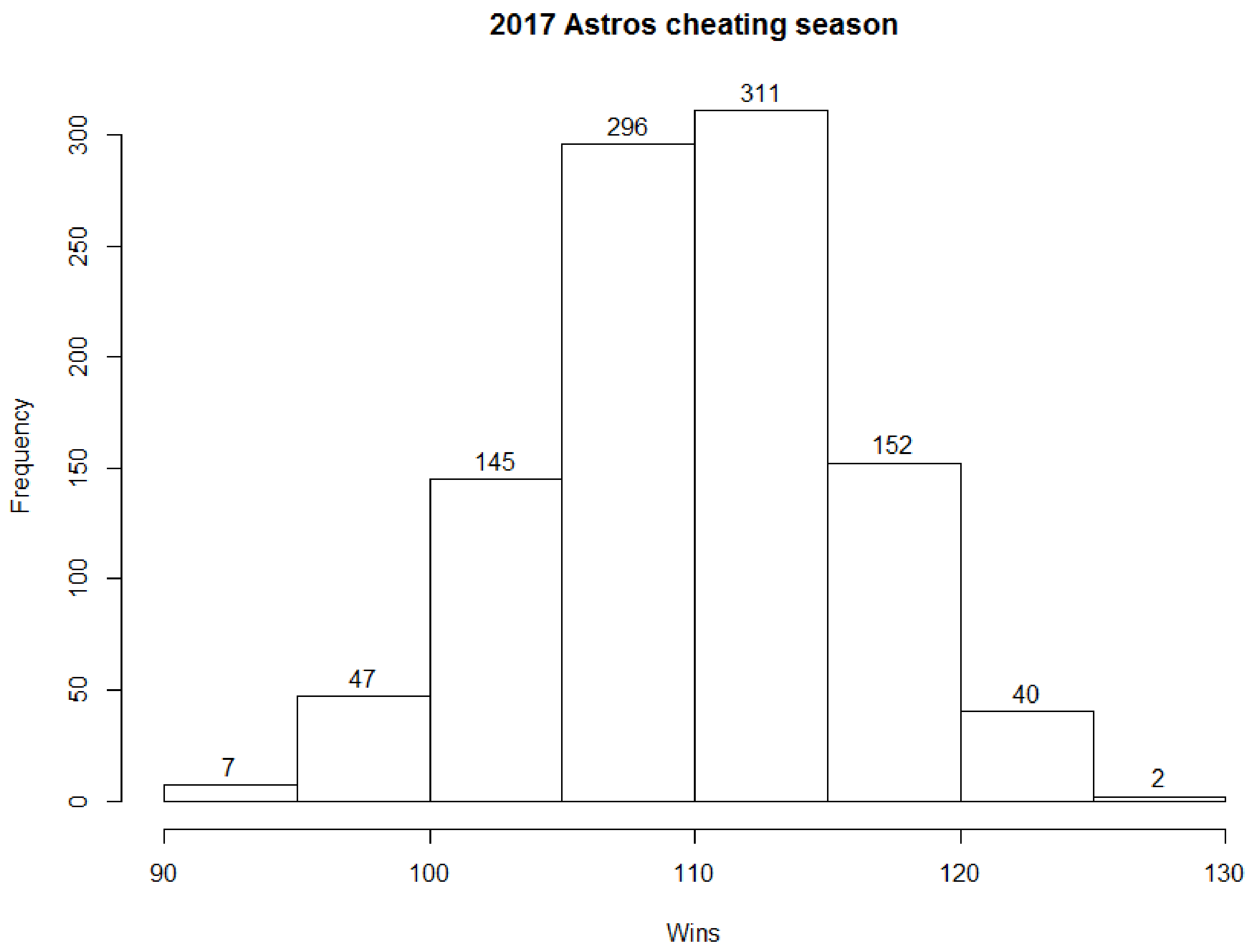
That results in a table of counts as this one:

	Team	Winner.Div	Wild.Card	Winner.Lg	Winner.WS	n
27	HOU	0	1	0	0	1
28	HOU	1	0	0	0	738
29	HOU	1	0	1	0	129
30	HOU	1	0	1	1	132

After 1000 simulation of 2017 Astros cheating seasons we have that Houston wins its division every time but one, in which they set for a wild card placement, moreover they are only division champions 74% of the times, while in the remaining 26% they get the best record in the AL onto the WS, that see an almost even split between wins and losses. We can also look at RS results:

```
avg<-seasonscheat%>%
  filter(Team == "HOU")
> mean(avg$Wins)
[1] 110.38
> hist(avg$Wins, xlab = "Wins", main = "2017 Astros cheating season",labels = T,brea
ks = c(90,95,100,105,110,115,120,125,130))
```

Table IX: 2017 Astros seasons' win distribution, cheating season



Labels are related to the number of seasons out of the 1000 simulated ones in which the Astros win a total of games in the described interval (left limit, such that 90-94 is an interval and so on until a cutoff at 120+ wins).

That is the whole lotta lot we are going to search for at the end of our simulations: how did the Astros fare given respective scenarios and changes in player performance and WAR? Do they always have the division in lock? Is there anything to say about playoff chances and regular season results?

i. Scene 1: no PAs

The first scenario we are going to explore is the most immediate one, and the most appropriate for the 2020 we are living in.

In the era of cancel culture the fastest way to approach our problem is to consider things as it didn't happen at all, a 2017 season where all the occasions in which the banging scheme was used are simply vanished from the history books. Seems only fair to consider the Astros performance without the cheating, and it would be easy to think of this fix as a rather blunt method to extirpate all gains generated by the bangs, but is it?

Let's dive in on the number of plate appearances (PAs) that saw one or more occurrences of a bang:

Table X: bangs dataset table, results in players' statlines

Batter	PA	AB	1B	2B	3B	HR	BB	SF/H	HBP
Jose Altuve	22	20	1	2	0	2	1	0	1
Alex Bregman	80	71	13	4	1	4	7	2	0
Carlos Correa	49	43	6	3	0	2	5	1	0
George Springer	77	60	8	5	0	1	15	1	1
Marwin Gonzalez	69	63	6	1	0	2	5	1	0
Yuli Gurriel	69	63	11	7	0	2	5	0	1
Josh Reddick	19	19	5	1	1	0	2	0	0
Carlos Beltran	77	71	9	5	0	1	6	0	0
Brian McCann	30	25	3	1	0	1	4	1	0
Jake Marisnick	41	37	6	1	0	4	3	0	1
Evan Gattis	36	32	2	5	0	0	4	0	0
Derek Fisher	10	8	1	0	0	0	2	0	0
Tyler White	13	13	2	2	0	2	0	0	0
JD Davis	8	8	1	0	0	1	0	0	0
Cameron Maybin	6	6	0	0	0	0	0	0	0
Nori Aoki	8	8	1	2	0	0	0	0	0
Juan Centeno	10	9	1	0	0	0	0	1	0
Max Stassi	8	6	0	0	0	0	2	0	0

We decided to consider a sort of podium structure for the "cheat sheet": a bronze medal to players who had a limited number of PAs (<= 10) with bangs, a silver medal to those players who had a fair amount of banged PAs (>10 but <45) and a gold medal to the lucky ones who had bangs on a notable chunk of their PAs (>= 45).

In scenario 1 what we'll do is simply slice away all banged PAs and their results from the previous 2017 totals and then recover player and team WAR given the changes on the statlines. After a series of pedantic calculations, the new statlines are as follows:

Table XI: 2017 Astros hitters' statlines and WAR values, noPAs season

2017 no PAs												
Batter	**G**	**PA**	**AB**	**1B**	**2B**	**3B**	**HR**	**BB**	**SF/H**	**HBP**	**IBB**	**WAR**
Jose Altuve	153	640	570	136	37	4	22	57	4	8	3	7,3
Alex Bregman	155	546	485	82	35	4	15	48	5	7	2	2,8
Carlos Correa	109	432	379	77	22	1	22	48	3	2	5	4,9
George Springer	140	552	488	84	24	0	33	49	3	10	1	4,1
Marwin Gonzalez	134	446	392	75	33	0	21	44	1	6	4	4,4
Yuli Gurriel	139	495	466	85	36	1	16	17	6	6	1	1,1
Josh Reddick	134	521	458	94	33	3	13	41	12	0	1	3,2
Carlos Beltran	129	432	396	56	24	0	13	27	6	3	3	-1,0
Brian McCann	97	369	324	50	11	1	17	34	4	7	3	2,6
Jake Marisnick	106	218	193	24	9	0	12	17	1	5	1	0,9
Evan Gattis	84	289	268	43	17	0	12	14	3	4	0	1,3
Derek Fisher	53	156	138	20	4	1	5	15	0	3	1	0,0
Tyler White	22	54	48	6	4	0	1	4	1	1	0	-0,2
JD Davis	25	60	54	5	4	0	3	4	1	1	0	0,0
Cameron Maybin	21	57	53	5	1	1	4	3	0	0	0	0,1
Nori Aoki	71	216	194	39	10	1	2	15	4	2	1	-0,3
Juan Centeno	22	47	43	9	0	0	2	4	0	0	1	-0,1
Max Stassi	14	23	18	1	1	0	2	4	1	0	0	0,2

The Astros team WAR in a 2017 season without banged PAs checks out at 50,3 WAR, a 2.8 points deficit with respect to the original team WAR, that is as the Astros lost a man on the starting rotation or a good bat in the lineup. While the number is still towering over the opposition in the AL West, it is interesting to see such a small malus on a consistent reduction of PAs: why is that?

What it's easy to forget is that taking out PAs means not only erasing singles, home runs and extra base hits from a player's statline, but also a decent number of outs or unproductive plate appearances: look at the curious case of Marwin Gonzalez! In the original cheating season he scores 4 WAR but after erasing his banged PAs he gets a 0.4 boost instead of losing performance, and that is because in almost 70 PAs he loses a pair of bombs, six singles and five walks but also an enormous amount of outs, almost 10% of his seasonal PAs!

That said we have all we need to simulate a thousand 2017 seasons in which all cheating PAs for Astros hitters are gone. First set the new WAR vector:

```
war17noPAs<c(24.8,33.4,31.3,24.7,47.4,43.5,24.7,34.4,26.6,42,42.6,32.3,55,14.9,21.9,
19.9,41.4,52.6,36.7,27.3,18.5,57.1,25.7,24.8,34.9,27.1,50.3,21.3,29.8,30.9)
```

Then start the simulation through the *one.simulation* loop:

```
> seasonsnoPAs <- NULL
for(j in 1:1000){
  noPAs <- one.simulation.17(war17noPAs)
  noPAs$Simulation <- j
  seasonsnoPAs <- rbind(seasonsnoPAs, noPAs)
}
```

After a little while we'll have or 1000 seasons in storage. Let's check how the Astros fared with a small reduction of their team WAR due to some PA erasures:

```
> countsnoPAs<- seasonsnoPAs%>%
   group_by(Team)%>%
   count(Winner.Div, Wild.Card, Winner.Lg, Winner.WS)
```
On the table for Houston we had:

	Team	Winner.Div	Wild.Card	Winner.Lg	Winner.WS	n
28	HOU	0	1	0	0	3
29	HOU	1	0	0	0	861
30	HOU	1	0	1	0	77
31	HOU	1	0	1	1	59

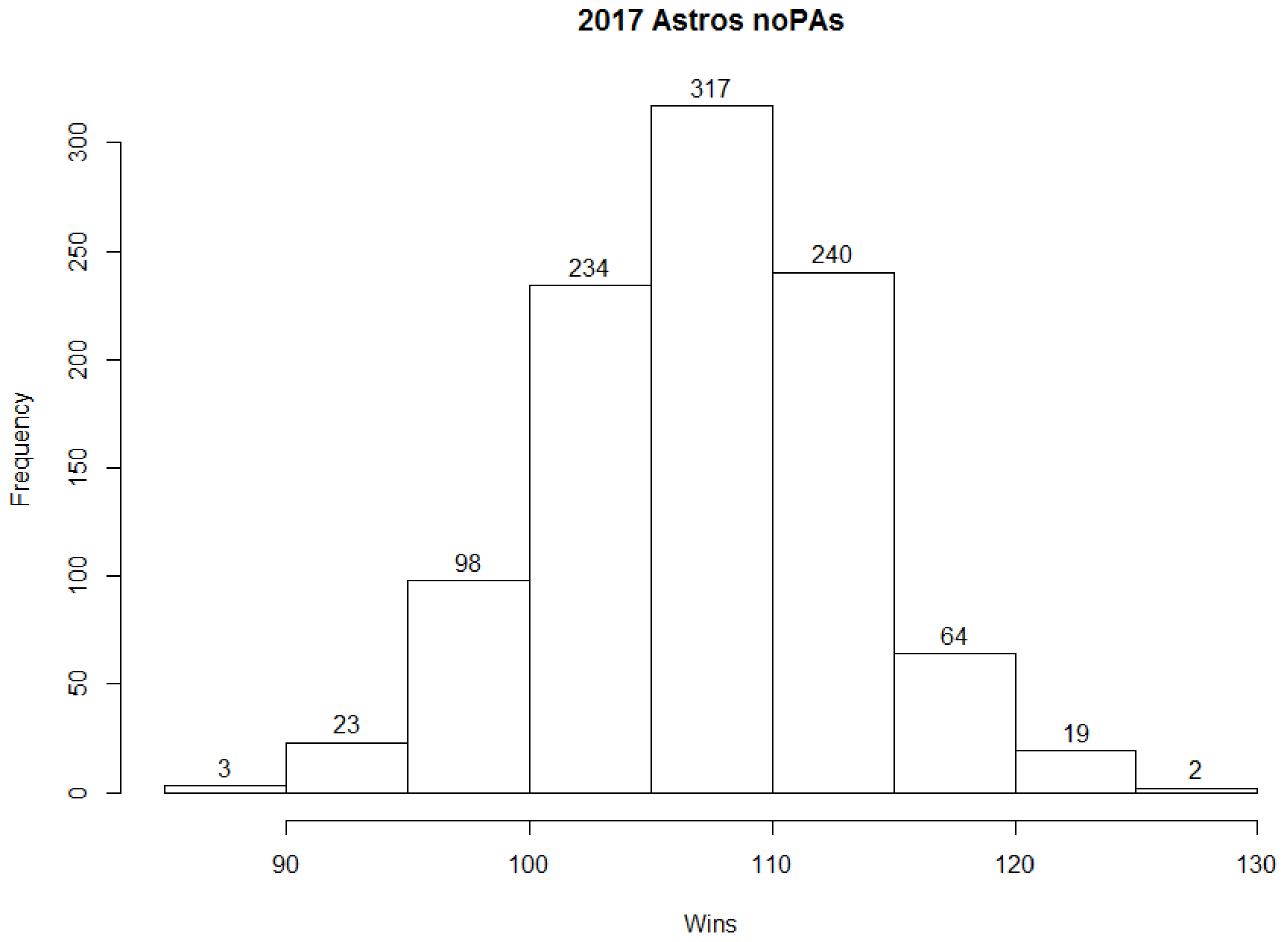
Not a lot of significant changes: the Astros win their division almost every time, apart from 3 seasons where they only get a Wild Card spot. This means that a deficit of little more than 2.5 WAR is not enough to close the gap between Houston and the other teams in the AL West. The Astros stop at that 86% of the times, while they accomplish the best record in the AL on the remaining 14% of seasons, getting a WS trophy in 6% of their tries and losing the final series on the other 8%.

A 2.8 team WAR loss generated a 12% swing in the chances of getting to the World Series and a 5% drop in the probability of winning rings, trophies and glory. And remember, playoffs in real life are a shot in the dark so that everyone can win, while we built a structure that grants the team with the best record a secure path to the WS, so postseason chances are to be taken with a grain of salt.

What didn't change in scenario 1 is the utter and complete dominance of the Astros in their division, so that while we cut out all cheating PAs, we got back a similar result in terms of regular season standings. Let's check them in more detail:

```
> avgnoPAs<-seasonsnoPAs%>%
   filter(Team == "HOU")
> mean(avgnoPAs$Wins)
[1] 107.604
> hist(avgnoPAs$Wins, xlab = "Wins", main = "2017 Astros noPAs",labels = T,breaks =
c(85,90,95,100,105,110,115,120,125,130))
```

Table XII: 2017 Astros seasons' win distribution, noPAs season



As we can see there aren't many feasible changes in the win distribution for the 1000 simulated season on the Astros side, also the mean number of wins still looms around 105+ attesting the strength of the Houston team against much weaker opponents in the American League West.

Well, maybe we have been too kind! While thinking that simply slicing away all bang occurrences from existence could have proven significative, the outcome left a lot to desire.

Let's get angry then and punish the cheating some more!

ii. Scene 2: no bases

In scenario 1 we tried to seek for some meaningful changes in the possibility of regular and postseason outcomes for a 2017 Astros season where all the PAs involving the banging cheating scheme were taken out of the equation.

As it turned out there wasn't much movement in Houston's chances of winning the division, solidly at almost 100%, achieving the best record in the AL almost 14% of all simulated season, and winning the WS at half the shot at the AL crown.

How can we penalize the Astros more? A clear path is given us by the consideration we made before on the nature of erased PAs: a lot of them where outs and that is true for both bad hitters and all-time greats. Think about it: in the last season such feat happened, Ted Williams, the legendary Boston Red Sox player who etched his name into baseball history on a career that saw WW2 get in-between, was able to hit for an AVG of 0.400.

That is unheard of, as nowadays great contact hitters such as Altuve in Houston barely make it to 0.350, but it's also a testament to how hard the game of baseball is: Ted Williams was an out 60% of the times in one of the best seasons ever by a batter! By cutting all PAs with a bang we are shedding a consistent number of outs from Astros players' statlines, therefore boosting their contributions.

So, what about cutting all PAs with bangs that have led to some sort of gain in terms of bases? Not only base hits, also walks, hit by pitches and all extra base occurrences. It is somewhat of a step back from before, as we are maintaining all banged PAs that resulted in outs, but will it change something for players' performances?

Table XIII: 2017 Astros hitters' statlines and WAR values, nobases season

2017 no bases												
Batter	**G**	**PA**	**AB**	**1B**	**2B**	**3B**	**HR**	**BB**	**SF/H**	**HBP**	**IBB**	**WAR**
Jose Altuve	153	655	588	136	37	4	22	57	4	8	3	6,9
Alex Bregman	155	595	547	82	35	4	15	48	5	7	2	1,3
Carlos Correa	109	464	416	77	22	1	22	48	3	2	5	4,0
George Springer	140	598	531	84	24	0	33	49	3	10	1	3,1
Marwin Gonzalez	134	500	449	75	33	0	21	44	1	6	4	3,0
Yuli Gurriel	139	538	523	85	36	1	16	17	6	6	1	-0,3
Josh Reddick	134	531	475	94	33	3	13	41	12	0	1	2,7
Carlos Beltran	129	488	461	56	24	0	13	27	6	3	3	-2,5
Brian McCann	97	389	344	50	11	1	17	34	4	7	3	2,1
Jake Marisnick	106	244	226	24	9	0	12	17	1	5	1	0,1
Evan Gattis	84	314	296	43	17	0	12	14	3	4	0	0,7
Derek Fisher	53	163	144	20	4	1	5	15	0	3	1	-0,1
Tyler White	22	61	61	6	4	0	1	4	1	1	0	-0,5
JD Davis	25	66	62	5	4	0	3	4	1	1	0	-0,2
Cameron Maybin	21	63	59	5	1	1	4	3	0	0	0	0,0
Nori Aoki	71	221	202	39	10	1	2	15	4	2	1	-0,5
Juan Centeno	22	55	51	9	0	0	2	4	0	0	1	-0,3
Max Stassi	14	29	22	1	1	0	2	4	1	0	0	0,1

Ouch! This is a big-time turnaround, as a lot of hitters lose 1+ WAR, with Gurriel now in the negative WAR zone, Bregman gone from solid contributor to reserve level and Beltran checking in one of the worst seasons ever by a bat.

All cuts and bruises to the statlines return a total Astros team WAR for the 2017 no bases season of 38.6, which is a yikes of epic proportions. From the original season Houston lost a Trout and a number two starter worth of production following the changes, a hard knock as it brings the Astros closer to the competition in their division but much further from other great AL teams such as Cleveland.

We are expecting big losses in terms of chances of success, so let's spin the wheel by starting from a new WAR vector:

```
>war17nobases<-c(24.8,33.4,31.3,24.7,47.4,43.5,24.7,34.4,26.6,42,42.6,32.3,55,14.9,21.9,19.9,41.4,52.6,36.7,27.3,18.5,57.1,25.7,24.8,34.9,27.1,38.6,21.3,29.8,30.9)
```

Then simulate the usual 1000 seasons with the updated WARs and probabilities as per B-T model:

```
>seasonsnobases <- NULL
for(j in 1:1000){
  nobases <- one.simulation.17(war17nobases)
  nobases$Simulation <- j
  seasonsnobases <- rbind(seasonsnobases, nobases)
}
```

Now it's the time to see what happened to the Astros on the aftershock of our changes:

```
> countsnobases<- seasonsnobases%>%
   group_by(Team)%>%
   count(Winner.Div, Wild.Card, Winner.Lg, Winner.WS)
```

and on the table as before:

	Team	Winner.Div	Wild.Card	Winner.Lg	Winner.WS	n
29	HOU	0	0	0	0	96
30	HOU	0	1	0	0	96
31	HOU	1	0	0	0	806
32	HOU	1	0	1	0	1
33	HOU	1	0	1	1	1

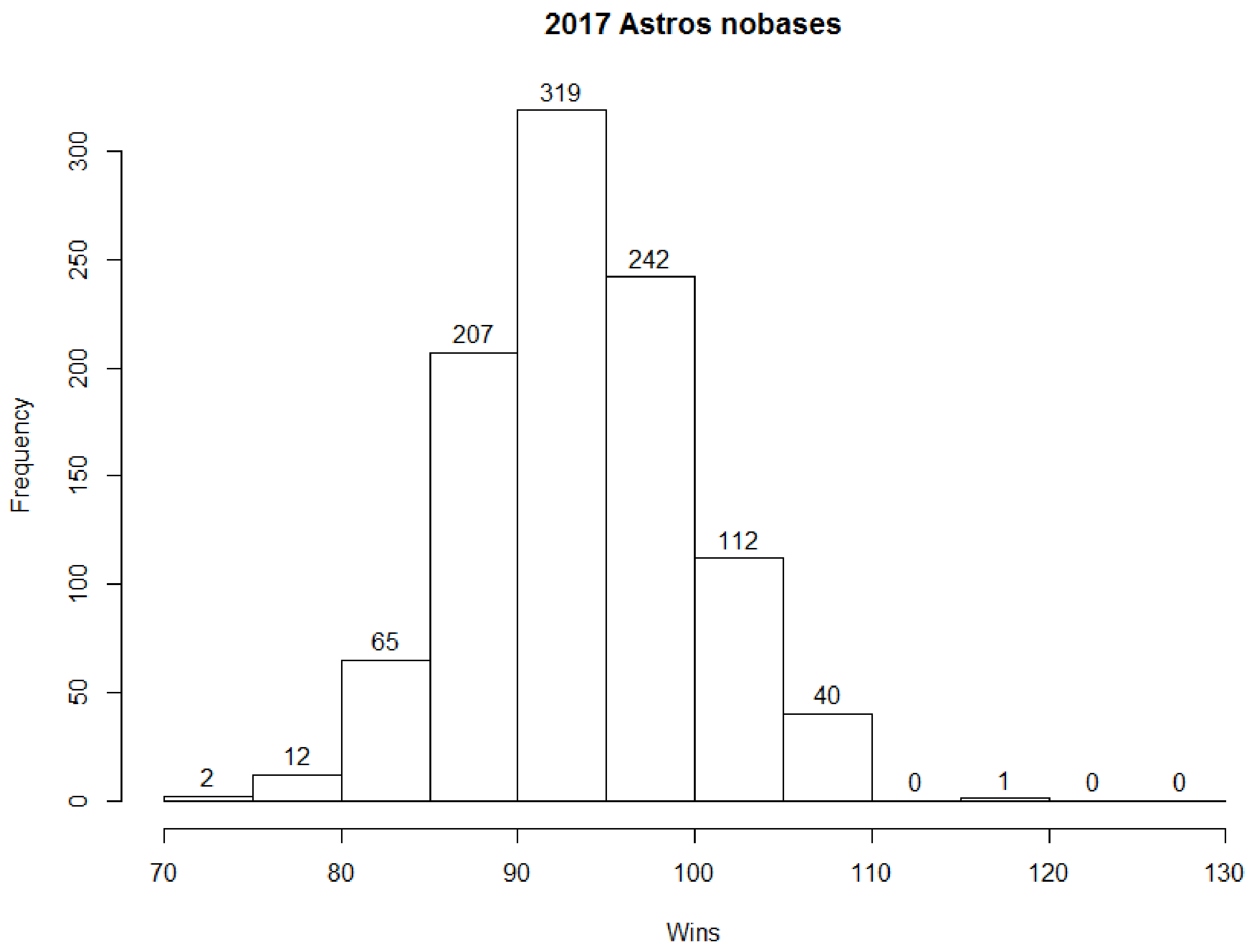
That's what happens when you lose 15 WAR worth of value: from a sure-fire first place in the division the Astros fall to a 81% shot at the AL West, while reaching a Wild Card spot or actually not even that in equal probabilities on the remaining 19%. Also, they are not going much forward after their division! Only in 2 out of 1000 season do the Astros have the best record in the AL, and a sole time they raise a WS banner.

What really jumps out is the loss of dominance in the division. Let's dig deeper on the regular season results:

```
> avgnobases<-seasonsnobases%>%
   filter(Team == "HOU")
> mean(avgnobases$Wins)
[1] 94.118
```

That was a called regression, as losing 15 WAR in a simulation where it is the foundation of winning probabilities returns a mean loss of 15 wins in the regular season:

Table XIV: 2017 Astros seasons' win distribution, nobases season



Winning upwards of 110 games is utopic now! Houston had to settle for a lot of 95-win seasons, not bad at all, and endure some atrocious <80 wins displays, something that was unconceivable under previous circumstances.

If we wanted to bury Houston, we dug a grave as deep as any. Is it fair and square though? We only kept the bad results and discarded all positive outcomes, but why should those outs repeat every time? Sometimes the hitter, bang or not, could have been able to get something out of the PA rather than an out, and we are simply not conceding this chance.

We need to find a middle ground, and all studies we introduced before can lead us to a new avenue.

iii. Scene 3: plate discipline

In Chapter II we looked at what kind of studies have been done on the 2017 Astros season up to now, whether they used or not the Adams' dataset and what are the results in terms of player and team performances.

As noted by the articles of Mailhot and Lindbergh the clear advantage Houston took from the banging scheme is a number of progresses in terms of plate discipline: less swings and misses, lower chase % on pitches outside the strike zone, more contact and damage on offerings inside the zone.

While working on contact is a hard task due to the infinite possibilities that can be considered by changing a player's statline in terms of hits, a simple fix to maybe hinder the positive results due to cheating can be that of dealing with walks, our bases on balls.

If the perk of the banging scheme works, it would be customary to see an increase in BBs as hitters know which kind of pitch is coming and therefore make better swing decisions, laying off all those pesky breaking balls in the dirt to end up with a free pass at first.

What we have done is a quick change on the number of BBs a player achieved in the 2017 season: instead of the actual figure, we took the previous 3-year average in terms of BB% and calculated the number of BBs by multiplying all PAs without bangs (but including original walks) for the average walk percentage as defined before. Note that for rookies and those with careers < 3 years we employed the BB% as predicted by ZiPS, an algorithm developed by Dan Szymborski at FanGraphs (Cistulli, 2016).

The consequences should be a stark decrease of bases per balls and their relative impact in terms of batting runs and WAR. Let's check for ourselves:

Table XV: 2017 Astros hitters' statlines and WAR values, plate discipline season

2017 plate discipline													
Batter	**G**	**PA**	**AB**	**1B**	**2B**	**3B**	**HR**	**BB**	**SF/H**	**HBP**	**IBB**	**BB %**	**WAR**
Jose Altuve	153	640	592	136	37	4	22	36	4	8	3	5,6	6,1
Alex Bregman	155	546	496	82	35	4	15	38	5	7	2	6,9	2,1
Carlos Correa	109	432	382	77	22	1	22	45	3	2	5	10,5	4,7
George Springer	140	552	476	84	24	0	33	63	3	10	1	11,5	4,9
Marwin Gonzalez	134	446	418	75	33	0	21	21	1	6	4	4,7	3,0
Yuli Gurriel	139	495	465	85	36	1	16	18	6	6	1	3,6	1,1
Josh Reddick	134	521	466	94	33	3	13	43	12	0	1	8,2	3,0
Carlos Beltran	129	432	380	56	24	0	13	43	6	3	3	10,0	-0,1
Brian McCann	97	369	324	50	11	1	17	34	4	7	3	9,3	2,6
Jake Marisnick	106	218	202	24	9	0	12	10	1	5	1	4,5	0,4
Evan Gattis	84	289	263	43	17	0	12	19	3	4	0	6,6	1,6
Derek Fisher	53	156	137	20	4	1	5	16	0	3	1	10,1	0,1
Tyler White	22	54	48	6	4	0	1	4	1	1	0	8,3	-0,2
JD Davis	25	60	54	5	4	0	3	4	1	1	0	6,8	0,0
Cameron Maybin	21	57	52	5	1	1	4	5	0	0	0	7,9	0,2
Nori Aoki	71	216	194	39	10	1	2	16	4	2	1	7,5	-0,3
Juan Centeno	22	47	44	9	0	0	2	3	0	0	1	6,6	-0,2
Max Stassi	14	23	22	1	1	0	2	0	1	0	0	1,7	0,0

There we have it! Some players saw a clear regression in their performance, with Altuve losing 1.5 WAR due to lost discipline and many others missing half a win here and there. All the small losses add up to a total team WAR for the 2017 Astros plate discipline season of 48 WAR, much better than scenario 2 but 5 WAR less than the original, which means losing a great hitter or a rotation stalwart.

Now it's time to simulate. We expect the usual dominance in the division and maybe slight worse results in the playoffs with respect to scenario 1, but a much better performance overall if compared to scenario 2:

```
>war17platedisc<-c(24.8,33.4,31.3,24.7,47.4,43.5,24.7,34.4,26.6,42,42.6,32.3,55,14.9
,21.9,19.9,41.4,52.6,36.7,27.3,18.5,57.1,25.7,24.8,34.9,27.1,48,21.3,29.8,30.9)
```

And then the simulations:

```
>seasonsplatedisc <- NULL
for(j in 1:1000){
  platedisc <- one.simulation.17(war17platedisc)
  platedisc$Simulation <- j
  seasonsplatedisc <- rbind(seasonsplatedisc, platedisc)
}
```

Counting our results for 1000 seasons:

```
> countsplatedisc<- seasonsplatedisc%>%
   group_by(Team)%>%
   count(Winner.Div, Wild.Card, Winner.Lg, Winner.WS)
```

in table form:

	Team	Winner.Div	Wild.Card	Winner.Lg	Winner.WS	n
29	HOU	0	0	0	0	1
30	HOU	0	1	0	0	5
31	HOU	1	0	0	0	901
32	HOU	1	0	1	0	51
33	HOU	1	0	1	1	42

As we expected the situation is much more pleasant than the previous scenario: Houston still leads the division almost every season, apart from 5 in which it only gets the Wild Card and a sole disastrous display where they don't reach the playoffs. In terms of record the Astros achieve the best win total in the AL almost 10% of the times, winning the WS in 4% of all simulations.

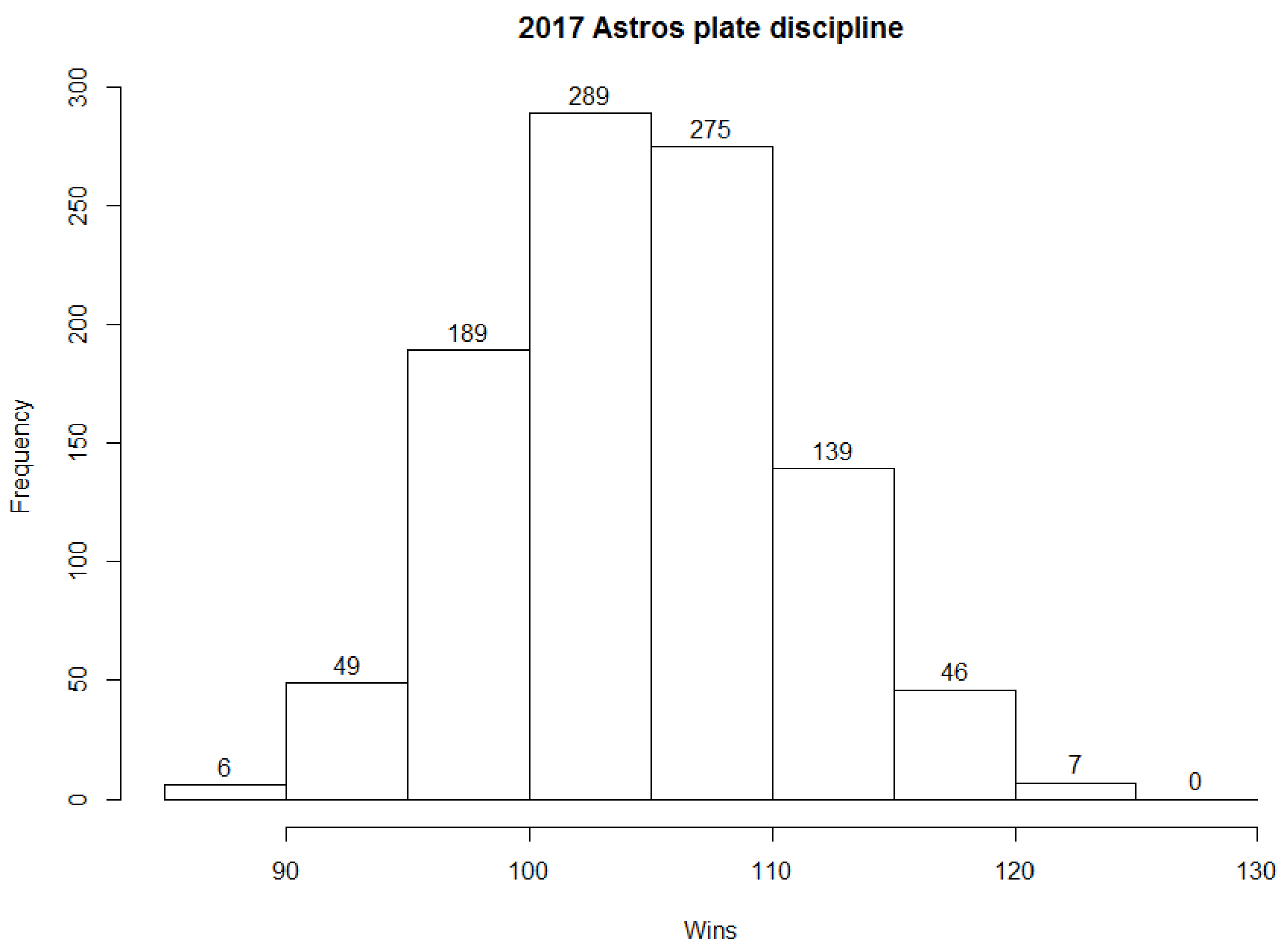
If compared to scenario 1, the added fix of bases on balls counts for little to no change in the chance of winning the AL West, still a comfortable task, and for a 4.5% decrease in the shot of having the best record in the AL, with 3% less chances of taking home the trophy.

What does it mean looking at regular season results?

```
 > avgplatedisc<-seasonsplatedisc%>%
   filter(Team == "HOU")
> mean(avgplatedisc$Wins)
[1] 105.059
```

The mean record is much better than before and 2-3 wins less then scenario 1 and the original season as expected from the team WAR. Graphically:

Table XVI: 2017 Astros seasons' win distribution, plate discipline season



Rather similar results to those of scenario 1, with a slight edge of 100-105-win seasons instead of 105+, although the meaning is the same, that the Astros are still one of the teams to beat in the AL.

One point we couldn't attack directly were the gains on contact and batting prowess related to the newfound "ability" of knowing pitches beforehand, but there's another way we can look at the same argument from a different point of view: what if all progresses by Astros hitters in 2017 were due to the cheating? How can we erase them and evaluate the matter?

iv. Scene 4: back to 2016

A consideration that many could do is that stopping at the mere bases on balls is an understatement of the advantages provided by the banging scheme, and they would have a point. Let's exaggerate this statement and say that all kinds of improvements that happened for each player in 2017 are due to cheating and that only.

Therefore, we need to recalibrate each performance to the 2016 season for each player. As easy as it seems, this method comes with some warnings: simply taking the 2016 statlines and WARs is not enough, we need to somehow project those values for 2017; then again how can we say for sure that the statlines in terms of hits and all other stuff hold up?

What we decided for is a simple adjustment, that avoids using each player's 2016 statline but focuses on that year's WAR: we only need 2016 WAR and PAs for each player and then, having the 2017 PAs (total plate appearances, no matter the bangs) we just use a proportion:

2016 WAR : 2016 PA = 2017 WAR : 2017 PA

It's quite the chop on, a plain and easy way to recalculate WAR for our 2017 season given the previous one, and so it comes at a price: we are simply putting all our eggs in a basket, considering each progress related to WAR as cheating-driven, also those on baserunning and defense that shouldn't be impacted, so that the run values of 2016 are projected on 2017 PAs to give our player WARs.

How does it all add up to?

Table XVII: 2017 Astros hitters' WAR values, 2016 data season

Batter	WAR
Jose Altuve	6,3
Alex Bregman	2,9
Carlos Correa	3,8
George Springer	4,2
Marwin Gonzalez	0,4
Yuli Gurriel	0,8
Josh Reddick	1,6
Carlos Beltran	2,0
Brian McCann	1,4
Jake Marisnick	0,5
Evan Gattis	2,0
Derek Fisher	0,1
Tyler White	0,0
JD Davis	0,1
Cameron Maybin	0,4
Nori Aoki	1,3
Juan Centeno	-0,3
Max Stassi	-0,2

One thing you'll notice is that, while some players seems to gain in terms of performance, as Beltran jumps up from slums to actual good role player and Aoki from bad to serviceable, other players, young ones as Bregman and Correa and resurgent performers such as Reddick and McCann, see sharp decreases in WAR, no one more than the always weird Marwin Gonzalez, that loses almost all of his 4 WAR of breakout campaign down to his precedent replacement value.

Note that the same considerations on rookies and short-career players hold onto this scenario, as we used the 2016 WAR projections as per ZiPS on Davis, White and Fisher (Cistulli, 2016).

After all the shakes the Astros team WAR for the 2017 season on 2016 data is 46.3, a slight decrease from scenario 3 but much better than the pit of doom that was scenario 2.

For the last time, let’s replay our season again and again, starting from the updated WAR vector:

```
> war17data16<-c(24.8,33.4,31.3,24.7,47.4,43.5,24.7,34.4,26.6,42,42.6,32.3,55,14.9,21.9,19.9,41.4,52.6,36.7,27.3,18.5,57.1,25.7,24.8,34.9,27.1,46.3,21.3,29.8,30.9)
```

Onto the simulations:

```
>seasonsdata16 <- NULL
for(j in 1:1000){
  data16 <- one.simulation.17(war17data16)
  data16$Simulation <- j
  seasonsdata16 <- rbind(seasonsdata16, data16)
}
```

After the simulation set up the usual count data frame:

```
> countsdata16<- seasonsdata16%>%
   group_by(Team)%>%
   count(Winner.Div, Wild.Card, Winner.Lg, Winner.WS)
```

that returns the following results:

	Team	Winner.Div	Wild.Card	Winner.Lg	Winner.WS	n
31	HOU	0	0	0	0	2
32	HOU	0	1	0	0	12
33	HOU	1	0	0	0	931
34	HOU	1	0	1	0	31
35	HOU	1	0	1	1	24

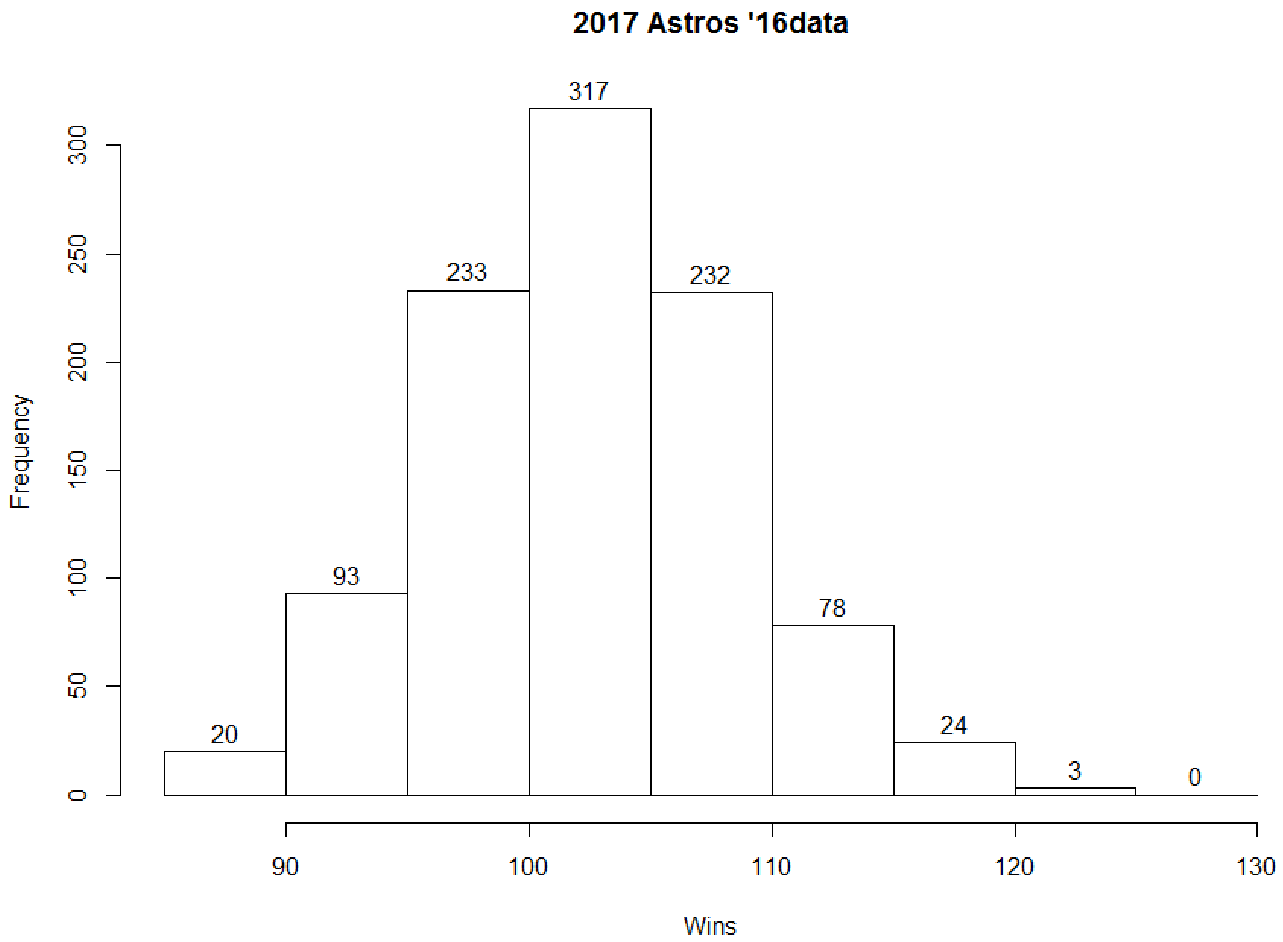
It’s a little step back from scenario 3 but still miles better than the misery of scenario 2: the Astros win their division a comfortable 98.5% of the times, missing the playoff on 2 out of 1000 simulation and stopping at a Wild Card spot on 1% chances. Houston achieves the best record in the AL 5% of all simulations, gaining a path to the WS and winning the trophy on a 2.5% shot.

These are rather small drops from the previous scenario: a 1.7 WAR decrease meant a 1% shift on chances from winning the division to just a WC, and on postseason matters it caused a 4% drop in the probability of having the best AL win record and a subsequent 1.7% decrease in WS chances.

In terms of regular season, we expect similar results to that of scenario 3 as for mean wins and win distribution:

```
> avgdata16<-seasonsdata16%>%
   filter(Team == "HOU")
> mean(avgdata16$Wins)
[1] 102.943
> hist(avgdata16$Wins, xlab = "Wins", main = "2017 Astros '16data",labels = T,breaks = c(85,90,95,100,105,110,115,120,125,130))
```

Table XVIII: 2017 Astros seasons' win distribution, 2016 data season



We actually have a steeper drop in the mean number of wins per season, with an almost 3 win decrease with respect to the previous scenario, and the same percentage of seasons with 95-100 wins and 105-110 wins, both really good results but at the same time a clear sign that as we close the WAR gap between teams the competition is enhanced and leads to costly losses.

V. Sliding doors: analyzing the landscape of results as per scenarios

In our exploration of possible scenarios for the 2017 Astros season we tried to cover a multitude of arguments on the benefits of cheating and on how to operate on players' performances to cut away all their fruitful implications.

From shelving all positive banged results, or all occurrences of bangs without discrimination, to adjusting statlines on a particular or general way, we constructed a series of possible shenanigans and we attached to them a correspondent effect on team results in the form of WAR.

Using it as proxy for a simulation of thousands of season as series of tosses of weighted coins we arrived at different stops, each with specific characteristics regarding records, win distributions, postseason accolades and individual contributions.

A clear-cut method, maybe a little too naive and epicurean, gave either way a reasonable conclusion: changing results means dealing with performances and values that can directly determine the outcome of a season.

It's time to sum up all the results we got out of 4000 seasons in 4 different scenarios and to compare them with the original 2017 Astros cheating season, to see if and how much our changes have affected the complexion of Houston's run at a World Series.

a. The endgame: team projections

Thousands of seasons are enough to give us an idea of what we are going to see considering each scenario and the correspondent modifications we applied for each one of them in the calculation of players' performances.

A decrease in WAR for a certain number of players is a guarantee of lower team WAR, that shrinks the probability of winning a match whoever the opponent is and therefore leads to less wins and lower chances of good results, both in the regular season and in the playoffs.

What we need to observe is a spectrum of all those chances on our five hypothesis of season, as to quantify in raw numbers and percentages the different lay of the land in terms of winning games and titles.

To simplify all things, we resort to a table similar to those of players' statlines:

Table XIX: scenarios' results in terms of team WAR, wins, accolades and win distributions

Season results	2017 cheat		2017 no PAs		2017 no bases		2017 plate discipline		2016 data	
	figure	season %	figure	season %	figure	season %	figure	season %	figure	season %
Astros team WAR	53,1	/	50,3	/	38,6	/	48,0	/	46,3	/
Average wins	110,38	/	107,60	/	94,12	/	105,06	/	102,94	/
Division wins	999	99,90%	997	99,70%	808	80,80%	994	99,40%	986	98,60%
League wins	261	26,10%	136	13,60%	2	0,20%	93	9,30%	55	5,50%
WS wins	132	13,20%	59	5,90%	1	0,10%	42	4,20%	24	2,40%
WC spots	1	0,10%	3	0,30%	96	9,60%	5	0,50%	12	1,20%
No playoff seasons	0	0,00%	0	0,00%	96	9,60%	1	0,10%	2	0,20%
<80 wins seasons	0	0,00%	0	0,00%	14	1,40%	0	0,00%	0	0,00%
80-84 wins seasons	0	0,00%	0	0,00%	65	6,50%	0	0,00%	0	0,00%
85-89 wins seasons	0	0,00%	3	0,30%	207	20,70%	6	0,60%	20	2,00%
90-94 wins seasons	7	0,70%	23	2,30%	319	31,90%	49	4,90%	93	9,30%
95-99 wins seasons	47	4,70%	98	9,80%	242	24,20%	189	18,90%	233	23,30%
100-104 wins seasons	145	14,50%	234	23,40%	112	11,20%	289	28,90%	317	31,70%
105-109 wins seasons	296	29,60%	317	31,70%	40	4,00%	275	27,50%	232	23,20%
110-114 wins seasons	311	31,10%	240	24,00%	0	0,00%	139	13,90%	78	7,80%
115-119 wins seasons	152	15,20%	64	6,40%	1	0,10%	46	4,60%	24	2,40%
120+ wins seasons	42	4,20%	21	2,10%	0	0,00%	7	0,70%	3	0,30%

Season% is related to the percentage of seasons out of the 1000 simulations in which the Astros achieve a certain classification and win games on set intervals.

As we can see this structure contains both round numbers, in WAR, wins and counts, and percentages calculated as parts of our 1000 season simulation per scenario.

Some results were to be expected: the cheating season retains the higher team WAR and as a consequence the higher average wins and title wins, both in division and AL; while the no bases season, that penalizes Houston's team of 15 WAR, returns the worst landscape of the lotto as it introduces a bunch of negative results in 200 of the 1000 simulated season, cutting from division wins to WC spots and no playoff seasons.

A pillar on all scenarios is the high probability of the Astros to win their division, the AL West: in the worst case scenario it is almost 81%, bleaker then all other 98%+ outcomes but still a consistent shot at gaining a place in the postseason without having to sweat a do or die WC game.

But this also leads to a mortifying consideration: no matter the penalty we inflicted Houston, and scenario 2 is a harsh one, the team always had the upper hand on the opposition in the division…but then why cheating? Was that 10% more chances at skipping a playoff round really worth disgracing the history of a team and the career of great players?

That's why it is so mind-numbing for a fan only trying to understand the reasons behind such an act: the 2017 Astros were a powerhouse no matter what, their division was arguably the weakest in all of baseball, the roster was poised to be a destructive force and playoffs were a clinch almost every time, no matter the possible incidents along the way.

The only motive we found for such a choice is highlighted in yellow and it involves win distribution: what we did is none other than point out the "fat" central part of each scenario's seasons counted in win groups, taking all the previous graphs and visually represent them on the table.

As we can see in the cheating season the big chunk, more than 60% of all simulations, returns a win total between 105 and 114 wins, which would be historic given that the last team to win more than 115 games where the 2001 Seattle Mariners, with Ichiro, The Kid and Edgar, that went 116-46 (and yet didn't win it all).

More than that: in 20% of the seasons the Astros rewrite baseball history, winning upwards of 115 games, and maybe that willingness to be remembered as regular season conquerors led an already stacked team to its cheating ways, otherwise such a rash and nonsense decision can only be attributed to greed and disrespect to the game, and so all criticisms and penalties are deserved, much heavier than they have been given.

By looking at each scenario's win distribution it is clear that the most probable outcome out of all in all configurations considered was a dominating season, with a win record between 100 and 110 wins, which is actually the interval where the original 2017 Astros cheating season lies, as to say that if we wanted to search for a feasible impact of cheating on the regular season landscape, well we found nothing that says so.

Playoffs are a different beast, with all its problems and variance: as we considered them, they are more a regular season part 2, were RS record matters and the final WS is a coin toss most of the times. In real life this is much different, and it is where cheating could have had a serious impact on home games, as a win in a 7-game series is much more valuable than one on a 162-game season.

One final point to bring up: a WS winning shot between 2,5 and 6% may seem low but it’s actually a great projection. On other algorithms much more complex than our poor-men WAR based one, as are ZiPS and others, where playoffs are much more fluctuating and a series of possible mishaps and miracles are considered, a team having a 5% chance at the trophy is a warning sign for all others, as it shows a steady roster whatever the casualties.

The 2017 Astros were great, yet they broke the rules to get more, and for that there’s no number that can give a plausible explanation, rather just a reminder that baseball players are humans as us, with all the sins and mistakes we do.

b. To each its own: player projections

The Astros cheating scandal, as it was painted by the media and set in stone by Manfred on his report, was a player-driven scheme for that the manager and GM took the hit, on their inability to stop it rather than their involvement, whereas players themselves were granted total immunity by the league to have their testimonies.

Blames are to be given to those who deserve them, and by looking at the numbers of each player’s statline as per scenario we’ll try to check who gained more from the scheme, as to understand who was up for it and if there were players that didn’t really wanted it but kept their mouth shut as it’s done in different kinds of organizations.

As usual, a table will tell us a story that we are going to explore:

Table XX: 2017 Astros hitters' production in terms of wOBA, Batting Runs and WAR, scenario dependent

Batter	2017 cheat			2017 no PAs			2017 no bases			2017 plate discipline			2016 data		
	wOBA	BatR	WAR	wOBA	BatR	WAR	wOBA	BatR	WAR	wOBA	BatR	WAR	wOBA	BatR	WAR
Jose Altuve	0,405	48,7	7,6	0,405	47,3	7,3	0,394	42,3	6,9	0,382	34,6	6,1	0,391	41,3	6,3
Alex Bregman	0,351	17,4	3,5	0,346	13,0	2,8	0,311	-3,6	1,3	0,333	6,8	2,1	0,336	10,7	2,9
Carlos Correa	0,394	30,8	5,1	0,400	30,1	4,9	0,369	19,9	4,0	0,396	28,4	4,7	0,349	13,5	3,8
George Springer	0,376	30,8	4,5	0,381	29,4	4,1	0,353	17,9	3,1	0,397	36,9	4,9	0,353	19,6	4,2
Marwin Gonzalez	0,382	27,8	4,0	0,410	34,8	4,4	0,363	19,1	3,0	0,371	20,2	3,0	0,298	-7,2	0,4
Yuli Gurriel	0,344	12,7	1,7	0,339	8,9	1,1	0,304	-6,2	-0,3	0,341	9,5	1,1	0,292	-10,7	0,8
Josh Reddick	0,357	18,0	3,4	0,355	16,5	3,2	0,344	11,7	2,7	0,351	14,7	3,0	0,324	2,6	1,6
Carlos Beltran	0,283	-14,8	-1,1	0,287	-11,1	-1,0	0,249	-28,1	-2,5	0,313	-1,7	-0,1	0,358	14,7	2,0
Brian McCann	0,323	1,6	2,6	0,326	2,4	2,6	0,309	-3,0	2,1	0,326	2,4	2,6	0,326	1,1	1,4
Jake Marisnick	0,343	5,5	1,4	0,328	1,9	0,9	0,285	-6,8	0,1	0,303	-2,7	0,4	0,257	-12,5	0,5
Evan Gattis	0,325	2,0	1,4	0,329	2,7	1,3	0,300	-4,8	0,7	0,341	5,6	1,6	0,345	8,0	2,0
Derek Fisher	0,291	-3,7	0,0	0,295	-2,9	0,0	0,284	-4,6	-0,1	0,300	-2,4	0,1	0,295	0,0	0,1
Tyler White	0,356	2,1	0,2	0,290	-1,3	-0,2	0,234	-4,3	-0,5	0,290	-1,3	-0,2	0,287	-1,5	0,0
JD Davis	0,318	0,0	0,1	0,312	-0,3	0,0	0,276	-2,3	-0,2	0,312	-0,3	0,0	0,290	0,0	0,1
Cameron Maybin	0,277	-2,2	0,0	0,307	-0,5	0,1	0,277	-2,2	0,0	0,326	0,4	0,2	0,352	1,6	0,4
Nori Aoki	0,300	-3,4	-0,2	0,295	-4,1	-0,3	0,285	-6,2	-0,5	0,297	-3,8	-0,3	0,325	2,0	1,3
Juan Centeno	0,269	-2,3	-0,3	0,303	-0,6	-0,1	0,258	-2,8	-0,3	0,288	-1,2	-0,2	0,306	-0,8	-0,3
Max Stassi	0,330	0,3	0,2	0,384	1,3	0,2	0,327	0,2	0,1	0,264	-1,0	0,0	0,068	-6,4	-0,2
TOTAL batting WAR			34,1			31,3			19,6			29,0			27,3
Astros pitching WAR			19,0			19,0			19,0			19,0			19,0
TOTAL Astros WAR			53,1			50,3			38,6			48,0			46,3

Considering the team, it's worth noting that while the worst results for the most part are related to scenario 2 as it was bound to be, the best returns in terms of WAR, for scenarios different than cheating, are divided among the noPAs fix and the 16'data solution.

The measure in which single players 'performances are hurt is also an interesting point of view: in the first two scenarios, dealing straight with banged PAs, some players get barely scratched and others are turned to dust.

These results fit the narrative that Correa blatantly exposed: Altuve as one wasn't too fond of the trash-banging, losing a tenth of his value on pure cheated plate appearances, while other batters, particularly those with a lot of swing and miss in their game, saw noticeable changes, no one more than Alex Bregman, whose gains on fruitful banged PAs are worth 2+ WAR.

One shocking result is the one on the far right: for a lot of players cheating made little to no strides, rather they would have preferred to turn back the clock to 2016. Beltran is a head-scratcher: pointed as one of the instigators of the system along with Cora, and for that both they lost their jobs, he didn't fare well at all, returning poor performances if we cut away good and/or bad cheated PAs.

So, there we have it again: all the subterfuges and the wrongdoings weren't life changing for any player! Take our beloved Marwin Gonzalez: he broke out that season and earned a good contract in Minnesota, but in the worst possible statline operation we can perform he only loses 1 WAR, not much for a super-utility, switch-hitting swiss knife. He would have been paid either way, and all the others the same.

And therefore, the question remains: why did they do it? Not for winning, as they would have at least got a playoff rondo 9 out of 10 times, and not for their own personal sake, as performances are not that impacted unless we really shape them to be.

We would like to ask the Correas and Bregmans of the world for an answer, fearing that it might be what we think it is: we did it because we could, because we thought we wouldn't get caught and it was all gain without losses.

But they got caught and they lost, not the rings and the title mind you, but the respect of the fans, the sport and their peers, and that might be worth more than 15 WAR.

c. A flaw in the matrix: bias, injuries, booms and busts

In our analysis we cared about pointing out possible hardships and flaws in our choices and our method as to give some sort of push to any other fan of the subject to try and improve our work. Paraphrasing what Bill James once said: "A hundred years from now we'll think we got it all, yet we still won't have figured out nothing about baseball".

So, in this closing note we decided to briefly list all kinds of warnings and skipping stones we found along the way and provide some advice for future travelers.

The first dilemma in our study is the choice of WAR as cornerstone of our simulations as per B-T model. As everybody knows there are all kinds of WAR running around, so picking bWAR or any other construct instead of fWAR could be a way to find different changes.

Then again, in 2020 WAR is starting to be slightly outdated: the introduction of Statcast, a real time collector of on-field data done through cameras and sensors, cleared the skies on several underlying traits and characteristics of each player and event, from exit velocity and launch angle of a batted ball to spin rate and efficiency of a pitch, from sprint speed on the bases to route proficiency on defense.

A model based on a sole statline of basic results is therefore ancient and in need of further development, and so will be our work when a refurbished version of WAR, maybe some sort of xWAR as expected performance given Statcast data, will come out.

There are also some pure mathematical notes regarding WAR calculation: in our analysis for each scenario we kept the number of league runs and PAs as fixed, also when we cut out those where the Astros banged the can. Although it's a 400 PA sample on 180000+ total PAs, so quite the drop in the ocean, it is worth noting that eventual differences on WAR are to be found in league and positional adjustments, and there could be minuscule deviations in season's constants when we erase some batted ball events.

Another point where a different path could have been explored is the choice of a fixed WAR vector for each team, apart the Astros for that it changes as scenario-dependent, such that we consider seasons were teams end up with the same WAR over and over again. A smart choice could be that of drawing the team WAR value for each team from a different distribution that has mean as the original WAR of that specific team and a set standard deviation equal for all teams.

This would help with one of the main problems of fixing too many things: there's not that much external validity outside the scope of the 2017 season. As we set in stone some big comprehensive results such as team WARs, we excluded a melting pot of events: trades, injuries, suspensions and any kind of unforeseen happenings that could have deviated our values on one or the other way.

The same could be said for other teams' performances for their players: fixing a total means not allowing a roster to over/underperform. Saying that a team had X WAR means that the sum of that team's individual WARs must be X, no matter which player gave more or less to the cause. Who are we to say that the 2017 Angels had to have that WAR when all players could have been the same and human divinity Mike Trout even better?

As we decided to focus on a specific season as a painting of a time gone by, we took a picture of the final moments and results of that season and used it to redraw it thousands of times, but that because we were interested in a single team and in particular changes on that team's roster run value. If someone would ever want to use this structure for different aims, we strongly advise them to be less strict and more flexible on the construction of WAR vectors, although not too much as we don't believe in the randomness of talent, but hey, that is us, and we are bound to be wrong somewhere!

A further matter is related to scenario 4: teleporting WAR from a year to another just in proportion on the number of PAs is like trying to realize a sculpture with a sledgehammer. Although year to year WAR has historically a decent correlation, it doesn't mean that players are who they were, not in this era of player development where someone can unlock hidden potential on a whim by working the numbers and polishing his craft in organizations such as Driveline.

Lastly the playoff conundrum: we decided to keep it simple and straight, eliminating some postseason matches and relating RS (Regular Season) and PS (Post Season) through the *win.league* function. We already warned about how to make playoffs more like those in real life, but we'd like to also remind how crazy, bizarre and streaky postseason baseball is.

Although we opted to simulate postseason results as the Astros won the WS in 2017 and not considering the playoffs could have been seen as a way to avoid an issue by a fan of the team, we don't advocate to perform playoff simulations based on such a model: too many times baseball has decided to go against all kinds of numbers and stats come October, with historically good teams losing early (those 2001 Mariners as one), medium teams going all the way to a WS win (2015 Royals and their bullpen) and classic franchises getting literally cursed and not winning for decades, from the Red Sox after trading Ruth to the Yankees to the Cubs and their 108 years drought ended in 2016.

And that's all she wrote for our analysis. There are all kinds of other ways to go about simulating seasons, maybe choosing a different model from the B-T one if existent or changing the random distribution from which results are drawn, or else making changes in the schedule or in the functions.

Whatever you want to do with this thesis, go ahead. Take this model, cut it and shape it as you want and as you think it should be. We tried to make it as easier as possible, from the season structure to the results analysis, as we wanted to reach a wider audience and not only those who speak stats and R as their main language.

Baseball is a game for the people, of every country, social status and level of instruction, and so our thesis tried to be the most inclusive and reader friendly we could come up with given the topic, and if we came up short well, may you be our dearest and most solid reliever in the pen.

VI. After the sunset: what's next for the 2017 Astros

After 5000 seasons, 5 different scenarios, a number of statlines and results there's only one thing we can say for sure: we don't know if cheating really changed anything.

As lowly and ugly our model looks like, the simulations don't lie: the Astros had their division in the bag also in case of disasters and mass injuries. Not even losing a third of their team value made their chances of winning the AL West go under 80%, a probability we really would love to have when playing poker or blackjack.

While it sounds like we want to diminish what the Astros did, we rather want to point out that it was even worse than what we thought: as disgusting and sinful human beings we would have understood if the trash-banging system gave that big of an edge to justify breaking rules and we would have shamefully felt somewhat close to the players' thinking in the matter, maybe not saying it out loud at least.

Instead what comes out is that there wasn't a need for such an awful trope, and that is baffling, both because it is counterintuitive to think that knowing the upcoming kind of pitch didn't change the whole batting thing as much as it should have, and because it sucks realizing that the players you admired and cheered for tarnished their reputations and careers for a pair of base hits and a longball.

To those who say that the Astros won because they cheated, we can't really answer, because as much as the numbers say no and I as a fan would love to think that the 2017 WS are somewhat a clean win, the randomness of a game such as is baseball and the fact that the cheating perdured also in the playoffs can't exclude the possibility that the season we saw in our lives was one of them where those two damn hits and a longball led to the trophy.

And so, we are left with a single burning question: what will the future look like for the protagonists of that tragically faithful 2017 Astros season?

It's been almost a year since the Rosenthal/Drellich report came out and subsequent investigations led to the scandal of the decade, and at the moment I'm writing there's still no baseball to be spoken of.

Summer camps are on the go and, given the COVID-19 pandemic, a 60 games regular season is due to start at the end of the month, but risks and intricacies are always on the menu, so nothing is written in stone for MLB in 2020.

One of the main topics for the season, before all hell broke loose and schedules went flying away, was the kind of reception the Astros would have got from opponents and fans.

While it's true that both the manager and the GM have been fired, almost all the players that benefitted and participated in the scandal are still on the team, and if there's a sure thing in baseball, it's that there are two kinds of rules: written ones, that the Astros broke and were therefore punished by the book (for what has been called a slap on the wrist by many, and with good reasons), and unwritten ones, related to sportsmanship and respect, for that players are punished on the field.

It will be not as much the opposing team's fans (?) booing relentlessly the Houston side, rather what the players on the other teams will do to retaliate: get ready for a heavy dose of beanballs, hard slides on second base, maybe a brawl or two (although with a pandemic rules are even more strict).

We could discuss if these unwritten rules to be respected should be a thing or not (no, they shouldn't be in a civilized world), but some bruises on the ribs and spikes on the shins are not really the point, and neither is a yearly suspension for all managerial parties involved, Hinch and Luhnow and would-have-been managers Cora and Beltran.

The real question is: what is next for those 2017 Astros?

As crazy as it sounds, I fear that consequences are going to be much harsher for the players than the managers. While Cora and Beltran could be out of baseball permanently, a shame for the former, that won it all with the Red Sox in 2018 and got investigated for it again(Rosenthal and Drellich, 2020), and a grave for the hopes of the latter of being enshrined in Cooperstown, I suspect that both the former manager and GM of Houston are going to be back somewhere else in due time.

Hinch was one of the first to repent, on an interview with Tom Verducci (Footer, 2020) that aired nationally, taking partial blame for his inability to stop a system that was initiated by Cora and taken forward by Beltran among others, and time will be on his side, so that in close to five years he will be back at the helm of a team. The fact that he was a passive part rather than an active perpetrator could be his saving grace, and same for Luhnow.

As one of the most talented, analytically minded and forward thinking GMs in baseball, the figure of the former Astros GM out of this scandal turned out to be that of a scapegoat, if we were to listen to him: while it's true he was the first to know about the Codebreaker algorithm, and was amazed by its results, he never admitted to know what went on after that, the camera/monitor apparatus and the trash-banging.

His defence is both solid and trembling: on a more law-minded consideration he could really have been an unknowing accomplice, given that he wasn't in the dugout nor the clubhouse and so the whole thing could have gone on without him knowing anything, but on a common sense perspective how could he have not known what was going on?

Either way the fact that he was suspended for his inaction rather than an active participation in the scandal, combined with his talent and skills, lead me to believe that he'll be the first one back in another team's front office, not as a GM but probably on a smaller role.

What about the players? Why and how could they be hurt if they avoided any kind of suspension?

It's not about the possibility of physical damage from the opponents, and not even the risk of reconsideration of their performances for that season that could lead to a decrease in their future salaries. So, it's not harms or money, but what I consider most important for a player: the aftermath, as what is left after a career.

It all boils down to the kind of player we take in consideration: for all members of the 2017 Astros that were either at the end of the road or in the middle of a subpar career there’s no real danger.

The Evan Gattis, Jake Marisnick, all those middle relievers and bench bats are not going to be classified as certified cheaters, simply because they weren’t that great on the field to be remembered, when compared to other members of the team. A lot of them actually confessed their mistakes related to the scandal and are now enjoying a calm retirement or a quiet late part of their careers, while still getting to look at the WS ring.

That is not the same for the most talented players in that 2017 Astros team: pitchers could escape the danger as they weren’t really in on the act or benefitted of the scandal (although the contrary could be stated and studies could be made), so Verlander should be a safe bet for the HOF and McCullers could still carve out a great career, injuries aside, but the hitters? Here things get ugly.

There are a lot of Astros bats that are on their road to the Hall of Fame if their bodies and talents hold up into their late 30s: Altuve is on his way to some major milestones, 3000 hits first and foremost, and his berth on Cooperstown really only needs a compilator kind of career for another 8-10 years, Springer is a corner case but if his 2017-19 numbers stay the same for the rest of the way he could end up as a top 10 leadoff hitter in baseball lore.

Then there are a pair of youngsters in Bregman and Correa who have just started but still are on a solid path to historic careers, a sure-fire HOF for the former and a good bet for the latter if they can avoid too many trips to the IL, a problem that the young shortstop has already had these past few years.

To close it out there were also a pair or players that ended their career after that WS or shortly after and with a shot at immortality given their resumes, namely Beltran and McCann.

For all these players the real question is the one that was asked at the beginning: was it worth it?

It depends on what baseball really means to the individual: is it just a job, a skit done day by day thanks to talent and disposition with the sole end of getting paid and maybe gaining some accolades on the way, or is there something more than just a profession?

For all people involved in the scandal that belong to the first category there’s no doubt that, given the final result, cheating has been a smart choice, although not a game-changing one as we saw, and yes, while having done so is cause of dismay and criticism, at the end of the day if it’s about sitting at home with comforts and a nice piece of jewellery, well it was a no-brainer to begin with.

But for all others, players, personnel, fans and whatnot, that consider baseball more than just a sport all this situation and its consequences are going to hurt terribly in the long run.

There are few sports in the world were a person, is it a player, manager or else, is remembered for a single fact, a perpetual image that usually stems from a bad episode in his history: going back to the beginning the ol’ Shoeless Joe Jackson was one of the best players in his time and yet today he stands only as the one who sold a WS.

Same can be said for all the greats involved in the PED scandals: Bonds has actually the homerun record and Clemens a baggage of CY Young Awards as Best Pitcher and yet all these milestones are buried by the heavy stain of steroids, lies and broken promises. To this day the fact that two of the best players to ever set foot on a diamond are not into the Hall of Fame is a testimony that the baseball community is not forgiving nor forgetting.

What does it mean for the 2017 Astros players? Well, that, if they have ever had some hope of being more than just great players, they should have not done it dirty.

While there’s no doubt that the Altuve, Springer, Bregman and Correa will be remembered as once big performers, their run to the good side of history, a shining plaque in Cooperstown and some fond memories among the baseball enthusiasts, is, at least for now, over.

Until the community will soften up, and in more than a decade it hasn’t been the case for the Steroid Era players, all those careers, the milestones and accolades are going to come with a distressing footnote about the year in which they won it all but lost respect in doing so.

There’s only one side left: what about those who really love baseball more than just a game? What about me, an unlucky Astros fan?

I think of the 2017 Astros as of the scars you have after getting hurt as a kid: you had fun, you messed up and there’s that little sign that reminds you of those days, with all the good and the bad.

That scar is my 2017 WS winner cap: I still dust it off and wear it sometimes, because being a fan of a team is close to a faith, but boy, looking at it stings.

After all the numbers I piled up, the seasons I’ve simulated and their results I could almost convince myself that really, cheating didn’t matter that much at the end.

But baseball is still more than those stats. And that season, those players, that cap will always be a bittersweet memory of the October we got into heaven and the November we were sent to hell.

Acknowledgements

First and foremost, my utmost respect and gratitude go to prof. Mazzocchi.

Without him I could have never had the chance of relating my thesis to the sport I love, and that’s not a small praise. Not only that, for all the help, advice and hours spent discussing on how to get to this point I can only tip my cap and bow as a thankful player would do to his fans.

As far as real life is concerned, no one has ever been more important than my family, and to them this thesis is dedicated. I consider this to be just the first step on the road leading to me repaying them of all the efforts and investments they made on my studies, and I will continue to pursuit the best I can do to show how thankful I am.

I have always been treated as the first pick of the draft by my family, so I’m going to work as hard as I can not to betray those expectations.

My final salute goes to all inhabitants of the diamond, from the brilliant minds of writers on specialized sites such as Fangraphs, the Athletic and a thousand more, to the restless souls of all fans that contributed to this paper one way or another, particularly Tony Adams for his work on the dataset that made all this madness possible.

Baseball and sabermetrics are among the most inclusive and productive of all communities, overflowing with articles, analyses and genius in itself, and to them I owe almost all of the reasons I jumped the economics ship to the statistics one and went on a two year trip that turned out as a key event in my life.

And to all players of this fantastic game all over the world, keep up the good job!

“It’s a round ball and a round bat, and you’ve got to hit the ball square”

To my family, just the good ones

July 2020

References

In order of citation, divided by scientific and specialistic publications:

All player and team data are taken from FanGraphs, so that WAR in this case is fWAR on the last update regarding defensive and catcher metrics.

Scientific publications:

Ulloa, Juan Sebastian, Amandine Gasc, Phillipe Gaucher, Thierry Aubin, Maxime Réjou-Méchain and Jérôme Sueur. “Screening large audio datasets to determine the time and space distribution of Screaming Piha birds in a tropical forest” *Ecological Informatics.* Jan 2016.

Bongjun, Kim and Bryan Pardo. “Sound Event Detection Using Point-Labeled Data” *IEEE*. Dec 2019.

Mandel, Michael, Enis Berk Çoban, Dara Pir and Richard So. “Transfer Learning from Youtube Soundtracks to Tag Arctic Ecoacoustic Recordings” *IEEE.* May 14, 2020.

Bradley, Ralph A. and Terry, Milton E. "Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons". *Biometrika*. 1952.

Zermelo, Ernst. “Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung” M*ath.* 1929.

Hunter, David R. “MM algorithms for generalized Bradley-Terry models” *Annals of Statistics.* 2004.

Bradley, Ralph A. “Science, statistics, and paired comparisons” *Biometrics.* 1976.

McHale, Ian and Alex Morton. “A Bradley-Terry type model for forecasting tennis match results” *International Journal of Forecasting.* 2011.

Cattelan, Manuela, Cristiano Varin and David Firth. “Dynamic Bradley–Terry modelling of sports tournaments” *Journal of the Royal Statistical Society.* 2012.

Barry, D. and Hartigan, J. A. “Choice models for predicting divisional winners in major league baseball” *Journal of the American Statistical Association.* 1993.

Albert, Jim, Bennett, J. and Cochran, J. J. “Anthology of Statistics in Sports” *Society for Industrial and Applied Mathematics.* 2005.

Albert, Jim and Max Marchi. “Analyzing Baseball Data with R” *Chapman & Hall*. Dec. 6, 2013.

Albert, Jim and Max Marchi. “Analyzing Baseball Data with R” *Chapman & Hall*. 2nd ed. 2018.

Lyman, H. B. “Test scores and what they mean” *Allyn & Bacon.* 6th ed. 1998.

Specialistic publications:

Rosenthal, Ken and Evan Drellich. “The Astros stole signs electronically in 2017, part of a much broader issue for Major League Baseball” *The Athletic.* Web (on paywall). Nov 12, 2019.

Passan, Jeff. “Ex-Astros pitcher Mike Fiers: Team stole signs with camera” *ESPN.* Web. Nov. 12, 2019.

Diamond, Jared. “Dark Arts and ‘Codebreaker’: The Origins of the Houston Astros Cheating Scheme” *Wall Street Journal.* Web (on paywall). Feb. 7, 2020.

Diamond, Jared. “The Astros’ Front Office Created Codebreaker. The Players Took It from There” *Wall Street Journal.* Web (on paywall). Feb. 12, 2020.

Manfred, Rob. “In re Houston Astros. Decision” report link https://img.mlbstatic.com/mlb-images/image/upload/mlb/cglrhmlrwwbkacty27l7.pdf Web. Jan. 13, 2020.

Browne, Ian. “Cora named in MLB report, may face discipline” *MLB.com.* Web. Jan. 13, 2020.

Rosenthal, Ken and Evan Drellich. “Details emerge about Carlos Beltrán’s role in the 2017 Astros clubhouse and the team’s sign-stealing scheme” *The Athletic.* Web (on paywall). Feb. 11, 2020.

Clemens, Ben. “The Count Is King (Even After Accounting for Batter Skill)” *FanGraphs.* Web. Jan. 8, 2020.

Jomboy Media channel. “Astros using cameras to steal signs, a breakdown” *YouTube.* Web. Nov. 12, 2019. Video link https://www.youtube.com/watch?v=M2XNW1qHN9w.

Passan, Jeff. “Astros' Jeff Luhnow, AJ Hinch fired for sign stealing” *ESPN.com.* Web. Jan. 13, 2020.

Rosenthal, Ken. “Carlos Correa rips Bellinger, passionately defends Altuve and says the Astros deserve their 2017 title” *The Athletic.* Web (on paywall). Feb 15, 2020.

Waldstein, David. “Carlos Beltran Out as Mets Manager After Cheating Scandal” *New York Times.* Web. Jan. 18, 2020.

The Bangs dataset and a corpus of related analysis are thanks to Adams, Tony et al. at http://signstealingscandal.com/, 2019-2020.

Lindbergh, Ben. “There’s No Virtue in Signaling. But Is There Any Benefit?” *The Ringer*. Web. Nov. 22, 2019.

Sullivan, Jeff. “The Astros Have a Completely New Look” *FanGraphs.* Web. Jan. 31, 2017.

Tango, Tom. “RE288: Run Expectancy by the 24 base-out states x 12 plate-count states, recursively” *Tangotiger.* Web. Nov. 22, 2018.

Mailhot, Jake. “How Much Did the Astros Really Benefit from Sign-Stealing?” *FanGraphs.* Web. Nov. 20, 2019.

Mailhot, Jake. “Which Players Might Have Benefited from the Astros’ Sign-Stealing?” *FanGraphs.* Web. Nov. 27, 2019.

Arthur, Robert. “Moonshot: The Astros’ Offense Took A Huge Leap After They Started Stealing Signs” *Baseball Prospectus.* Web. Nov. 22, 2019.

Mailhot, Jake. “The Most Important Bangs of the Astros’ Scheme” *FanGraphs.* Web. Feb. 7, 2020.

Slowinski, Steve. “LI” *FanGraphs*. Web. Feb. 17, 2010.

Tango, Tom M., Mitchel G. Lichtman, and Andrew E. Dolphin. “The Book: Playing the percentages in baseball.” *Potomac Books, Inc.* 2007.

Smith, Sean. "Position Player WAR Calculations and Details." *Baseball-Reference*. 2010. Web. update Mar 2018.

All the formulas regarding WAR, from batting to WRAA, replacement and player runs can be found on Slowinski, Steve. “WAR for Position Players” *FanGraphs.* Web. Apr. 2, 2012 and correlated.

Weinberg, Neil. “Wins Above Replacement (WAR) Calculator” *NewEnglishD.* Web. Dec. 18, 2013.

Cameron, Dave. “Unifying Replacement Level” *FanGraphs.* Web. Mar. 28, 2013.

Slowinski, Steve. “Positional Adjustment” *FanGraphs*. Web. Feb. 26, 2010.

Slowinski, Steve. “What is WAR?” *FanGraphs*. Web. Feb. 15, 2010.

Albert, Jim. “Simulating a 60-Game Season” *BaseballWithR.* Web. Jun. 29, 2020.

Cistulli, Carson. “2017 ZiPS Projections – Houston Astros” *FanGraphs.* Web. Dec. 8, 2016.

Cistulli, Carson. “2016 ZiPS Projections – Houston Astros” *FanGraphs.* Web. Jan. 13, 2016.

Rosenthal, Ken and Evan Drellich. “MLB’s sign-stealing controversy broadens: Sources say the Red Sox used video replay room illegally in 2018” *The Athletic*. Web (on paywall). Jan. 7, 2020.

Footer, Alyson. “Hinch on sign-stealing: 'I tolerated too much’” *MLB.com.* Web. Feb. 7, 2020. Video link https://www.youtube.com/watch?v=uU3sK2kImKs.