

Free Agency in Baseball: MLB Teams Aren't Getting The Player They Bargained For

Abstract

Every winter, professional baseball teams bid for the services of the MLB's most talented hitters.

Baseball's top stars often sign contracts for upwards of half a decade, guaranteeing them tens, if not hundreds of millions of dollars. However, some of these contracts work out better than others. When teams shell out inordinate amounts of money, they are gambling on their investment panning out; when it doesn't, the consequences can be drastic. The conventional wisdom in baseball states that players who sign contracts with new teams tend to underperform expectations in their new city.

Mostly supported by anecdotal and circumstantial evidence, this theory can be scientifically put to the test. By comparing player performance from 2021 to 2022 for players that switched baseball clubs this past winter, we can either support or contradict this long-repeated conjecture. In this paper we will explore the differences from year to year in numerous key baseball statistics. For each significant statistical finding, practical explanations will be offered based on existing literature and practical causal arguments.

Keywords

Baseball Contracts, Baseball Statistics, Free Agency, Information Issues, Injuries, Player Decline, Player Durability, Player Performance, Value Added

Introduction

On May 5th, 2022, Boston Red Sox second baseman Trevor Story walked back to the dugout after a strikeout, his fourth such result of the night. Boos rained down on him from Red Sox fans who had, to this point, watched their beloved team muster zero runs through seven innings. Story had made his Red Sox debut just 27 days ago after signing a six year, \$140 million contract over the winter with Boston. How did the multimillion dollar experiment go sideways so quickly?

Fans of major league baseball have heard a version of this story many times before. A player destined for stardom seeks a change of scenery, and then finds themselves struggling to meet expectations in their new city. Particularly in the 2022 MLB season, fans of several teams found themselves wildly disappointed with shiny new acquisitions that they were positive would generate elite output. However, these negative sentiments gave me pause. Sports fans are among the most emotionally fraught and reactive individuals that I can think of. Red Sox fans had written off the 29-year old Story before he had even spent a full month playing with his new club. I figured that there had to be a statistical method to test whether these fans' gripes had some merit or whether they were merely the reactionary groans of disillusioned and bitter sports fans. Consequently, it is that question which I seek to answer with the methods presented in this paper.

Author John Cassidy briefly explored a version of this phenomenon in his book entitled *How Markets Fail*, an economics book with an added statistical and mathematical flair. In Chapter 12, he references a study done by Kenneth Lehn, a financial economist at the Washington University School of Business.

He found that output dropped among players who switched teams, with the main culprit being health. Players who moved spent, on average, nearly eight more days on the disabled list than players who re-upped their contracts with their existing team. His theory was that ball clubs seeking to poach players from other teams suffered from a key handicap- they had vastly less information about said players than did the existing teams of the players. I'll explore this angle later on, but Cassidy and Lehn provided a strong introductory look into the matter.

Stephen M. Garcia et al. explored the many intricacies of baseball free agency in a 2020 economic journal. They found an increasing unwillingness for high-ranking teams to pursue high-profile free agent players. Teams seem to have soured on the idea of the splashy, big-money winter acquisitions. Haunted by memories of major flops from years past, more teams are looking for production from their young and developing homegrown talents rather than outside, expensive, established talent. Another trend they found is that when other teams bid for the services of their players, teams are becoming less likely to provide a counter-offer that would retain that player. Teams becoming more willing to part with their players aligns with the hypothesis that a player's existing team knows more about them than any other team does. A player's true state of fitness is technically hidden information- the consequences of nagging injuries often don't present themselves until a player ages, long after a team has already taken a gamble on them.

Baseball teams shell out *billions* of dollars (\$3.2 billion total in 2022) in guaranteed money every winter in their quests to hoist the World Series trophy. It's a high-stakes gambling game that produces more

losers than winners, and can either save or cost people their jobs. The contribution of this paper is a more advanced look into what Cassidy and Lehn contributed. The rest of the paper is organized as follows. I will first lay out the source of my data and how it was cleaned and organized. I will then describe and justify the methods I used to compare player performance across the selected years. There will then be a reporting of all output tables, a walkthrough of all statistical computations, and finally in-depth analysis of my findings and what they may mean for the sport at large.

Data

The data that has been collected for this study came mainly from Baseball Reference and Spotrac, two websites that contain nearly every morsel of baseball data one could desire. Baseball Reference mostly provided the true statistics and game-by-game breakdowns, while Spotrac mainly provided contract and free agency information.

To begin, I had to decide who, or what, to sample. Because we are comparing 2021 statistics to 2022 statistics, I had to look at only players who switched teams in free agency between those two seasons. For the sake of consistency, we will refer to the period between the 2021 and 2022 seasons as the “2022 offseason”. There were 64 batters who were free agents during the 2022 offseason. Of those 64, precisely 50 signed major-league contracts with a team that was different from the team they were on at the end of the 2021 season. Spotrac lists six other players who technically switched teams, however these players did not switch teams until after the commencement of the 2022 season, making them ineligible for this study.

Once I had my sample of 50 players, I had to then choose which statistics I was going to use to measure player performance. The first group of statistics I analyzed we will refer to as “counting statistics”.

These statistics are simple: they merely count the number of events of interest that occur in a player’s season. Runs (R) refers to the number of times a player touched home plate and scored a run for their team. Hits (H) refers to the number of times a player hit the ball into fair territory and reached base without a defensive error. Doubles (2B) refers to the number of hits that resulted in the batter reaching second base. Triples (3B) refers to the number of hits that resulted in the batter reaching third base. Home runs (HR) refers to the number of times a player hit a ball over the home run fence in fair territory, awarding them all four bases. Runs batted in (RBI) refers to the number of runs that scored as a direct result of the player being up to bat. Stolen bases (SB) refers to the number of times a player was on base and advanced by taking a base when they weren’t entitled to it. Finally, total bases (TB) refers to the number of bases gained by a batter as a direct result of their hitting.

The next kind of statistic is referred to as a “value added statistic”. These statistics are technically counting statistics, however they measure a player’s added value above a hypothetical league-average player. Wins above replacement (WAR) measures a player’s value by deciphering how many more wins they’re worth than a replacement-level player at their same position. Runs above replacement (RAR) measures a player’s value by deciphering how many more runs they’re worth than a replacement-level player at their same position.

The next two statistics are still technically counting statistics, but I portioned them off because they contribute to a different narrative than do the rest of the counting statistics. They will be referred to as “health statistics”. Games (GM) refers to the number of games a player appeared in. Days on the injured list (IL) refers to the number of days a player spent on either the 10-day, 15-day or 60-day injured list. When a player is on the injured list, they are ineligible to play due to injury.

The final group of statistics can be referred to as proportion statistics. These statistics are the only ones used in this study that can rise and fall, and are not dependent on the amount of game participation. For example, a player that plays 150 games has a much better chance of having more hits than a player who plays 50 games. The following statistics do not follow the same rule. Batting average (AVG.) is the proportion of at bats that result in a hit for the player. Slugging percentage (SLG.) is calculated by dividing total bases by the number of at bats. Home run percentage (HR%) is the proportion of a player’s plate appearances that result in a home run. Walk percentage (BB%) is the proportion of a player’s plate appearances that result in a walk. Strikeout percentage (SO%) is the proportion of a player’s plate appearances that result in a strikeout. Finally, stolen base percentage (SB%) is the proportion of a player’s attempts to steal a base that are successful.

All of the aforementioned statistics are extremely common in baseball data discourse. If the goal of the study is to measure player performance, these are the best parameters to use. When a player is at bat, the television broadcast of the game will display several key statistics, all of which are included in the study: batting average, home runs, and runs batted in. The chosen data will help answer the research question

by providing a multifaceted look at all aspects of a player's offensive game. Contact, power, speed, and discipline are all measured in at least one of the given variables.

Methods

The basic statistical mechanism used in this study was analysis of variance. By utilizing a one-way ANOVA, I was able to compare means for all of the aforementioned variables from 2021 to 2022. The one-way ANOVA assumes that the data follows a normal distribution, which is an acceptable approximation given the large sample size of players. Each counting statistic, value added statistic, and health statistic was run through this one-way analysis of variance test. For each of these statistics, the null and research hypotheses followed the same general form. The null hypothesis was that the mean value among all 50 players for these statistics in 2021 was equal to the mean value among the same 50 players in 2022. This hypothesis is the one we seek to reject, thus we need a research hypothesis that has the burden of proof. We must be able to measure the strength of the research hypothesis via our P-value. Therefore, our research (alternate) hypothesis is as follows. The mean value among all 50 players for these statistics in 2021 was not equal to the mean value among the same 50 players in 2022. We are seeking to uncover whether there was significant change in player performance between the two years, so these hypotheses will suffice.

The statistical method changes slightly when analyzing the proportion statistics. Means are not very useful when comparing these kinds of statistics. For example, if Player 1 gets 10 hits in 20 tries (a .500 batting average) and Player 2 gets 0 hits in 5 tries (a .000 batting average), the mean of their batting

averages will be .250. However, if you pool their attempts, which gives Player 1 additional weight due to the larger sample, they have a combined 10 hits in 25 tries, for a batting average of .400. This .400 figure is much more accurate than the .250 figure, because it creates a weighted mean that gives each player a weight that is directly proportional to their number of attempts. The same tactic was applied to our proportion statistics. For these figures, I created a pooled number that used totals for the necessary variables rather than calculating a proportion statistic for each individual player. With the unweighted method, the mean batting average for our players in 2021 was .237. This number is arrived at by adding together all 50 batting averages and dividing by 50. However, as we have established, this is a foolhardy strategy. Instead, I divided the total number of hits by our players in 2021 (4391) by the total number of times they were at bat in 2021 (17,591) to arrive at a more accurate weighted average of nearly .250. It would make sense that this number is higher than the original .237 because players that get hits more frequently are also more likely to have more at bats.

Once I calculated these pooled proportions, I utilized a 2-sample Z test for difference between proportions. The null hypothesis is only slightly changed from before. In this case, the null hypothesis is that there is no significant difference between the proportion from 2021 and the proportion from 2022. Our research hypothesis, again with the burden of proof, would then be that there is, in fact, a significant difference between the proportion from 2021 and the proportion from 2022. We can apply these hypotheses to all of the proportion statistics to seek out significant differences.

Simulation & Application

The following tables contain the results of the one-way analysis of variance for our counting statistics.

The first column describes what the statistic is that is being tested. The second column contains the average number of that statistic that the 50 players recorded during the 2021 season. The third column contains the average number of that statistic that the 50 players recorded during the 2022 season. The fourth column contains the raw increase or decrease in that statistic from 2021 to 2022. The fifth column contains the percentage increase or decrease in that statistic from 2021 to 2022. The sixth column contains the P-value for difference in the two means as calculated by SAS 9.4. The final column contains the significance marking for the statistic. A statistic with one asterisk indicates rejection of the null hypothesis at $\alpha = 0.10$. A statistic with two asterisks indicates rejection of the null hypothesis at $\alpha = 0.05$. A statistic with three asterisks indicates rejection of the null hypothesis at $\alpha = 0.01$.

| Stat | 2021 | 2022 | Diff. | %Chng | P-Value | Signif. |
|-------------|-------------|-------------|--------------|--------------|----------------|----------------|
| R | 51.34 | 36.94 | -14.4 | -28.05% | 0.0202 | ** |
| H | 87.82 | 71.54 | -16.28 | -18.54% | 0.1038 | |
| 2B | 16.64 | 13.76 | -2.88 | -17.31% | 0.1686 | |
| 3B | 1.38 | 0.92 | -0.46 | -33.33% | 0.1206 | |
| HR | 14.84 | 9.82 | -5.02 | -33.83% | 0.0172 | ** |
| RBI | 47.46 | 37.06 | -10.4 | -21.91% | 0.0635 | * |
| SB | 5.16 | 3.68 | -1.48 | -28.68% | 0.2769 | |
| TB | 151.74 | 116.6 | -35.14 | -23.16% | 0.0449 | ** |

These results on their own do a decent job telling a story. All eight of these statistical categories saw a precipitous drop from 2021 to 2022, with all of them falling over 15% year-over-year. Four of the eight categories contain a rejection of the null hypothesis at $\alpha = 0.075$. Two of the eight, runs and home runs, reject the null hypothesis at $\alpha = 0.025$. However, baseball pundits tend to scoff at analysis that only includes these basic counting statistics. There's vastly much more to the offensive game, they claim, than just hits and runs.

The pundits aren't entirely wrong. It is highly likely that a better measure of player performance might be the value added statistics that we mentioned earlier. These statistics take a holistic approach to the game, attempting to summarize the true impact that a player might have on the games they participate in. The next table will contain the results of the analysis of variance for our two value added statistics. The same column headings used in the previous table are used here, with the same functionality. Bear in mind that these statistics are dependent on game participation- they are more likely to increase the more chances a player has- making them still technically counting statistics.

| Stat | 2021 | 2022 | Diff. | %Chng | P-Value | Signif. |
|-------------|-------------|-------------|--------------|--------------|----------------|----------------|
| RAR | 13.96 | 8.02 | -5.94 | -42.55% | 0.1043 | |
| WAR | 1.332 | 0.754 | -0.578 | -43.39% | 0.1191 | |

These analyses fall victim to the lofty standards of statistical significance. However, it can still be seen that the same large drops can be seen in these value added statistics. Both categories fell over 40% from

2021 to 2022, even if neither was significant at the level $\alpha = 0.10$. It should be noted, however, that both P-values came extremely close to being significant at that level. While our ANOVA may not have produced any definitive conclusions here, we can still see that our players, on average, added nearly six more runs to their teams in 2021 than in 2022.

You may be asking: what's going on with these huge year-over-year changes? After all, the fifty-player sample for 2022 is just the fifty-player sample for 2021 with a mere 365 days of added age. This is where it feels appropriate to mention the theories of John Cassidy and Kenneth Lehn that were mentioned in the introductory section. Are players that switch teams simply less fit to play? Our health statistics are able to gauge if there are significant differences in game participation from one year to the next. The next table will contain the results of the analysis of variance for our two health statistics. The same column headings used in the previous table are used here, with the same functionality.

| Stat | 2021 | 2022 | Diff. | %Chng | P-Value | Signif. |
|-------------|-------------|-------------|--------------|--------------|----------------|----------------|
| GM | 110.3 | 86.42 | -23.88 | -21.65% | 0.0069 | *** |
| IL | 27.24 | 30.32 | 3.08 | 11.31% | 0.6698 | |

After looking at this table, it seems we may have found our culprit. The difference in game participation between 2021 and 2022 for our players, as well as its accompanying P-value is staggering. The null hypothesis is rejected for games played at $\alpha = 0.01$. The drop from 110 games to 86 games is substantial, especially in a season that is only 162 games long. The high P-value for days spent on the

injured list is disappointing from a null hypothesis perspective, but we can still learn from this finding. This simply shows that we can't definitively say that major catastrophic injuries are the reason why player participation dropped so heavily from 2021 to 2022.

While we surely can pull information from these statistical findings, we wouldn't be doing our due diligence if we didn't run tests that control for game participation. We've established that players who switched teams played vastly less on their new teams, but how did they perform when they actually played? To answer this question, we have to use our proportion statistics. As stated before, these statistics do not necessarily increase with increasing opportunity. Rather, they are moving indicators of how any player is faring in their time on the field, however substantial that time may be.

It is important to note a few things before we change course. These proportion statistics are more prone than counting statistics to anomalies due to insufficient sample size. If you check the leaderboards for these categories just a few weeks into the season, you will see players you've never heard of sitting squarely in the top twenty. However, the standard error of these statistical categories is greatly decreased given a large sample size. The central limit theorem states that as the season goes on and sample size increases, the numbers will slowly approach what they are bound to be at by the end of the season. I believe that the hazard of inadequate sample size is prevented by employing the pooled-proportion strategy that was mentioned earlier. With the exception of stolen base percentage, all of these proportions have denominators of over 14,000, which should be more than enough to avoid the pitfalls of potential wayward results and statistical anomalies.

The next table will contain the results of the analysis of variance for our proportion statistics. The column headings here are slightly different from before. The first column describes the statistic of interest. The second column contains the numerator, denominator, and final calculated proportion for that statistic in 2021. The third column contains the numerator, denominator, and final calculated proportion for that statistic in 2022. The fourth column contains the P-value for difference in the two proportions. The final column contains the significance marking for the statistic. A statistic with one asterisk indicates rejection of the null hypothesis at $\alpha = 0.10$. A statistic with two asterisks indicates rejection of the null hypothesis at $\alpha = 0.05$. A statistic with three asterisks indicates rejection of the null hypothesis at $\alpha = 0.01$.

| Stat | 2021 | 2022 | P-Value | Signif. |
|-------------|----------------------|----------------------|----------------|----------------|
| AVG. | 4391/17591 = .24962 | 3577/14557 = .24572 | 0.4237 | |
| SLG. | 7587/17591 = .43130 | 5830/14557 = .40049 | <0.0001 | *** |
| HR% | 742/19853 = 3.737% | 491/16198 = 3.031% | 0.0002 | *** |
| BB% | 1831/19853 = 9.223% | 1310/16198 = 8.087% | 0.0001 | *** |
| SO% | 4413/19853 = 22.228% | 3568/16198 = 22.027% | 0.6455 | |
| SB% | 258/332 = 77.711% | 184/244 = 75.410% | 0.5157 | |

We can see here that, even when adjusting for the fact that our players were playing less in 2022, their performance took a tumble after moving to their new team. There are remarkably low P-values for the differences in slugging percentage, home run percentage, and walk percentage. The difference in the

three other statistical categories was not significant at any alpha level below 0.10, but a drop-off in performance can still be seen. The batting average for our players dropped nearly four points year-over-year (the batting average for the entire league only saw a dropoff of less than one point) and our players found slightly less success when trying to steal bases as well.

Discussion

Our first table showed statistically significant declines in runs, home runs, runs batted in, and total bases. These categories share a sort of commonality in that they are all statistics that rely on a hitter being able to have lots of swing power. Home runs are almost exclusively considered a sign of power, and our players hit, on average, five less home runs in 2022 than they did in 2021. There could be several possible explanations for this finding. A baseball team plays half of its games at its home field; when a player switches teams, they are invariably switching the venue in which they play a large chunk of their games. Ballpark dimensions vary from stadium to stadium, with the fences differing in height, distance from home plate, and angle. A good example of this comes from Trevor Story, one of our 50 sampled players and the one from the anecdote at the beginning of my introduction. Coors Field, his old field in Denver, measures 390 feet from home plate to the center field wall. Fenway Park, his new field in Boston, measures 420 feet to center field- a whopping 30 foot difference. Hitting a baseball 420 feet is a tall task. Players who have played in Boston for years will tell you they never try to hit a ball to center field, lest they waste their power trying to get the ball over one of the farthest walls in the entire league. Conversely, Story, playing his first season in Boston, is likely unfamiliar with the quirks and oddities that come with being a batter in a particular ballpark with unique dimensions. Our sample of

50 players all had to adapt to hitting in a new ballpark, forcing them to restructure their entire approach at the plate. This seismic disruption in their play style would also explain why these players were slightly less valuable in 2022 than in 2021, as shown in our second table.

I would be remiss if I didn't also mention that in general, moving teams is as massive a change off the field as it is on the field. Anybody who has transplanted themselves into a new city will agree that a massive adjustment period is necessary in order to get one's feet underneath them in a new environment. Shortstop Andrelton Simmons, one of our sample subjects, made the transition from small-market Minneapolis to the Windy City of Chicago, the third-largest media market in the United States. The aforementioned Story went from Denver, a city over 5,000 feet above sea level, to Boston, which is only a hair over 100 feet above sea level. It is not difficult to fathom a potential link between player performance issues and the enormous life-altering decisions that directly precede them.

Of course, it must be discussed that a sizable portion of the drop-off can be explained by our players' lack of physical fitness in 2022. As could be seen in our third table, game participation fell by nearly 25 games per player from 2021 to 2022. This difference was deemed empirically significant, and there is a theoretical explanation for this finding as well. Chapter 12 of John Cassidy's book *How Markets Fail* is entitled "George Akerlof's Market for Lemons", wherein he explains that unequal access to information plagues almost every sector of our economy. This includes the market for baseball free agents, which Cassidy explains suffers from the problem of hidden injury information. He posits that "a player's original team knows more about his health than a potential acquirer, and the decision to

release him to free agency may be a sign that it has concerns about his health” (Cassidy 154). This hypothesis is consistent with the previously mentioned studies of Garcia et al. which found that teams are more willing than ever to part with their own players. While the drop in games played doesn’t necessarily reveal a regression in talent or skill, a player needs to be healthy in order to be useful. Our value statistics are tied to game participation because a player that is riding the bench with back problems isn’t valuable at all. Teams are, generally, required to pay players their massive salaries regardless of how healthy they are, making this trend of bad health in newly signed players troubling for big spenders. Teams see big free agency contracts as an investment, and a player can’t provide return on that investment unless they’re truly fit to play. It should also be mentioned the effect that switching organizations might have on a player’s health. Even if there aren’t any nagging hidden issues like the ones Cassidy hypothesized about, a new team means a new training staff and a new set of team doctors. Teams aren’t owed any extensive medical records when they sign a player, so medical staff have very little to work from when treating a newly acquired player.

Arguably the most troubling finding for our players, even more so than their dismal health, was their inability to meet expectations when they *were* on the field. Our fourth and final table contained statistical categories whose drop-offs could not be explained by not playing as many games in 2022. Every category decreased in some way except for strikeout percentage, with slugging percentage, home run percentage, and walk percentage decreasing significantly. It seems as though between power, contact, speed, and discipline, power is the performance tool that appears to have evaporated the most in our players. The drop in slugging percentage and home run percentage matches up with the drop in

other power-related categories from table one. Hitting power is a complicated concept. Gratuitous strength doesn't always correlate with good power, but it certainly does play a part. Perfect swing form and unimpeachable timing ability are inarguably more important when it comes to launching the baseball- strength means nothing if you cannot get the barrel of the bat on the ball. A baseball swing can be thought of as a symphony of moving body parts. If a batter develops even the slightest hitch in their swing, the entire orchestra is thrown out of balance and the player stands no chance of hitting the ball squarely. These hitches usually develop from, coincidentally, nagging injuries to the back, neck, and extremities. With so many of the players in our sample battling long-term ailments in 2022, it is not difficult to see how their power was essentially zapped when compared to the season before. These weeks-long bouts with injury can cause major performance slumps, draining a player's confidence and setting them up for failure. Baseball is partially a mental game- step up to the plate with a pessimistic mindset, and you're almost guaranteed to fail. Early-season struggles, especially with a brand-new team, can easily spiral into season-long, emotionally torturous campaigns that leave fans wondering what exactly the team signed up for when they shelled out millions of dollars.

All of this cynicism is not to say that our players can't turn their fortunes around. The numbers don't lie: the players in our sample saw a downturn in essentially every single statistical category we decided to investigate. However, we have seen players take their lumps with a new team and persevere. Kyle Schwarber, one of the sampled players, saw one of the largest year-over-year decreases in batting average from 2021 to 2022. But when his Philadelphia Phillies made their improbable run to the World Series this November, Schwarber reversed course and smashed 6 home runs in 17 games, giving Phillies fans

hope for a brighter campaign ahead. Baseball is, and will continue to be, an endearingly unpredictable sport. Players that switch teams will likely always have their growing pains, but the fans that stick it out might just be lucky enough to bear the fruit when those players bloom.

References:

Bibliography will be on a different document.

Acknowledgements:

Thank you to the University of Connecticut Department of Statistics faculty, and specifically to Dr. Jun Yan, for their contribution to and assistance with the writing of this paper.

About the Student Author:

Jason Boghosian is an expected graduate of the University of Connecticut with a B.A. in Statistics and a minor in both Public Policy and Urban Studies. He will begin his postgraduate studies in August of 2023 when he pursues a Master of Arts in Survey Research at the University of Connecticut.

Press Summary:

The purpose of this study was to examine the effects that switching teams had on professional baseball players in 2022. The hope was to investigate whether there were statistical grounds for the claim that players who change organizations generally perform worse on their new team. This study focused on the difference from 2021 to 2022 in 18 statistical categories for 50 sampled MLB players that switched

teams between the two seasons. After running statistical tests, the implication of the findings was that players' fitness level, health, and overall hitting performance appeared to drastically regress after switching teams.