Moneyball vs. Millionaires: Examining the Need for a Salary Cap in Major League Baseball

Thomas Sigall

Table of Contents

[Introduction 3](#_Toc152766545)

[The 2002 Oakland Athletics 3](#_Toc152766546)

[The Competitive Balance Tax 3](#_Toc152766547)

[Objectives 4](#_Toc152766548)

[Data 5](#_Toc152766549)

[Tables 5](#_Toc152766550)

[Exploratory Data Analysis 6](#_Toc152766551)

[Payroll 6](#_Toc152766552)

[Winning Percentage 7](#_Toc152766553)

[Population 8](#_Toc152766554)

[Methods 10](#_Toc152766555)

[Payroll 10](#_Toc152766556)

[Chi-Square Test 10](#_Toc152766557)

[Independent T-test 10](#_Toc152766558)

[Population 11](#_Toc152766559)

[Chi-Square Test 11](#_Toc152766560)

[Independent T-test 12](#_Toc152766561)

[Bootstrap Methods 12](#_Toc152766562)

[Conclusions 13](#_Toc152766563)

[Future Steps 14](#_Toc152766564)

[References 15](#_Toc152766565)

[Appendix 15](#_Toc152766566)

# Introduction

### The 2002 Oakland Athletics

The legendary 2011 film "Moneyball" tells the story of the 2002 Oakland Athletics, a team taken over by general manager Billy Beane who led them to a historic season, winning 103 games, winning their division, and making the playoffs. The reason that season was immortalized into a movie, however, was not simply because the team had a historically good season. The 2002 Oakland Athletics did something far more remarkable. They were not just able to win, but they did it as a small market team, with an incredibly low payroll, the third lowest in the league in 2002[[1]](#footnote-1).

The Athletics had seen periods of success in their franchise's history, but struggled to end the 1990's. The problems they faced during that time were not unique to the Athletics. Many teams around the league were struggling because they could not keep up with the "big guys". The Athletics were and are considered a "small market" team. This means they are located in an area with a lower population, meaning they bring in less money as an organization, meaning they have less money to spend on good players. They simply could not keep up with "large market" teams like the Yankees and Red Sox, teams with far deeper pockets. In 2002 alone, the New York Yankees spent $125 million on player salaries, the Boston Red Sox spend \$108 million, and the Athletics spend a measly $40 million[[2]](#footnote-2). In fact, many players would start with the Athletics and other small market teams, improve and become better ballplayers, then leave to the large market teams who could afford to pay them far more, essentially outbidding the smaller market teams.

The 2002 Oakland Athletics are a remarkable team because they were able to innovate and find value in players that did not cost nearly as much as the superstars New York and Boston were paying for. This was 20 years ago though, and now all teams use similar methods that were pioneered in that 2002 season. There is little value left to find out there on the smaller payrolls that these smaller market teams are given. This project is not about attempting to replicate what the 2002 Athletics were able to do. Instead, we will look at the system that forced them and all other small market teams to do more with less just to compete with large market teams. We will investigate the relationship between payroll and team success.

### The Competitive Balance Tax

To help decrease the gap between teams with deep pockets and teams without, most sports leagues in the United States have what is called a salary cap. This is as it sounds, an agreed upon limit on the amount teams can spend on player salaries. For example, the National Football league set its salary cap at $224,800,000 for the 2023 season[[3]](#footnote-3). Teams are not permitted to spend more than that amount on their players and spend a lot of time trying to maneuver contracts to stay under that figure. Player salaries are by far the largest expense any team has and this hard limit on that expense helps to balance the playing field and close the gap between large market and small market teams.

The MLB, however, does not have a salary cap. There is no limit on the amount of money teams can spend on player salaries. Instead, they have implemented a system called the "Competitive Balance Tax" (CBT)[[4]](#footnote-4). An amount is agreed upon by the league, the CBT threshold, though teams are allowed to exceed this amount in their payroll. If they do exceed this amount, they are taxed at a rate proportional to the number of years they have been over this amount. This system has been in place since 2002 and the values for the CBT threshold over the years are as follows (in millions of dollars):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Year** | CBT | Year | CBT | Year | CBT | Year | CBT |
| 2003 | $117.00 | **2008** | $155.00 | **2013** | $178.00 | **2018** | $197.00 |
| 2004 | $120.50 | **2009** | $162.00 | **2014** | $189.00 | **2019** | $206.00 |
| 2005 | $128.00 | **2010** | $170.00 | **2015** | $189.00 | **2020** | $208.00 |
| 2006 | $136.50 | **2011** | $178.00 | **2016** | $189.00 | **2021** | $210.00 |
| 2007 | $148.00 | **2012** | $178.00 | **2017** | $195.00 | **2022** | $230.00 |

Naturally, this value usually increases each year to keep up with various factors, such as inflation, increasing league revenues, and others. It is a high value, as most teams have never passed this threshold in the 20 years of its existence. Enough teams do pass it, though, that it gives a good baseline to determine which teams are large market teams that spend big and which teams are not.

The theory behind this threshold is that it should punish wealthier teams for spending large amounts of money that smaller teams do not have. This should level the playing field and allow teams to spend similar amounts of money and encourage on-field competition as the way to determine the best team, rather than the league being a competition decided on who has the deepest pockets.

## Objectives

Now that the system the MLB has put in place is understood, we can look at the objectives of this project. The overall question that will be looked at is as follows:

**Does the MLB need a traditional salary cap?**

**To answer this question, we will look at the relationship between spending and winning and attempt to determine if the CBT system is effective at encouraging a competitive league by discouraging teams from spending so much money that smaller teams cannot keep up. Various questions will be answered, such as:**

* **Do teams above the competitive balance tax threshold win more games than teams that are not?**
* **Is there a relationship between playoff teams and teams that are above the CBT threshold?**
* **Do teams in large population centers win more games than teams in small population centers?**

**There is some concern that a salary cap would do little to address the real issue, that teams in big cities will just always have more money than teams in small cities. There are more people in New York than Oakland, so it may be the case that Oakland is just in a smaller market and no salary cap will ever fix that. However, if spending is more indicative of team success than city population, we will have evidence that spending is what needs to be limited rather than population. This would be an encouraging sign as the MLB has far more control over how much teams spend than population growth in its cities.**

# ****Data****

## Tables

Data was gathered from many sources. First payroll data was obtained from Retrosheet[[5]](#footnote-5) and formatted into a single table to be used for exploratory data analysis. This table includes team, season, payroll (in millions of dollars), winning percentage, if the team made the playoffs, and if the team is above the CBT. Playoff data is from Baseball Reference[[6]](#footnote-6).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **team** | **year** | **payroll** | **wpct** | **playoffs** | **above\_cbt** |
| Los Angeles Dodgers | 2022 | $285.51 | 0.685 | 1 | 1 |
| New York Mets | 2022 | $255.92 | 0.623 | 1 | 1 |
| New York Yankees | 2022 | $240.29 | 0.611 | 1 | 1 |
| Philadelphia Phillies | 2022 | $221.74 | 0.537 | 1 | 0 |
| San Diego Padres | 2022 | $202.17 | 0.549 | 1 | 0 |
| Boston Red Sox | 2022 | $201.62 | 0.481 | 0 | 0 |

In addition, population data was needed to address the concerns about population being a better predictor than payroll as explained earlier. Population data was obtained from the US Census Bureau[[7]](#footnote-7) and from the UN for Toronto's population[[8]](#footnote-8).

|  |  |  |
| --- | --- | --- |
| **City** | **Year** | **Population** |
| Atlanta city, Georgia | 2003 | 403080 |
| Atlanta city, Georgia | 2004 | 397365 |
| Atlanta city, Georgia | 2005 | 396311 |
| Atlanta city, Georgia | 2006 | 403548 |
| Atlanta city, Georgia | 2007 | 410086 |
| Atlanta city, Georgia | 2008 | 414233 |

The main population center was determined for each team and can be seen in the appendix. Note that some population centers have multiple teams.

## Exploratory Data Analysis

### Payroll

The first thing looked at was how each team's payroll changes over time. There are 30 teams in total, so fitting them all into one graph is not very practical. Instead, I chose to group the graphs by division, meaning splitting the graphs into six graphs of five teams each. The dashed line on each graph represents the CBT threshold.

A graph of different colored lines

Description automatically generatedA graph of different colored lines

Description automatically generated

We can see here that most teams do not even come close to the threshold, while some are above it most years. The New York Yankees, located in the AL East graph, have consistently above the threshold year after year. Compare that to the Tampa Bay Rays, also in the same graph. They are consistently one of the lowest teams in the league, consistently sitting at around 50% of the threshold, never even coming close to it. The remainder of the graphs can be seen below.

A graph of different colored lines

Description automatically generatedA graph of different colored lines

Description automatically generated

A graph of different colored lines

Description automatically generatedA graph of different colored lines

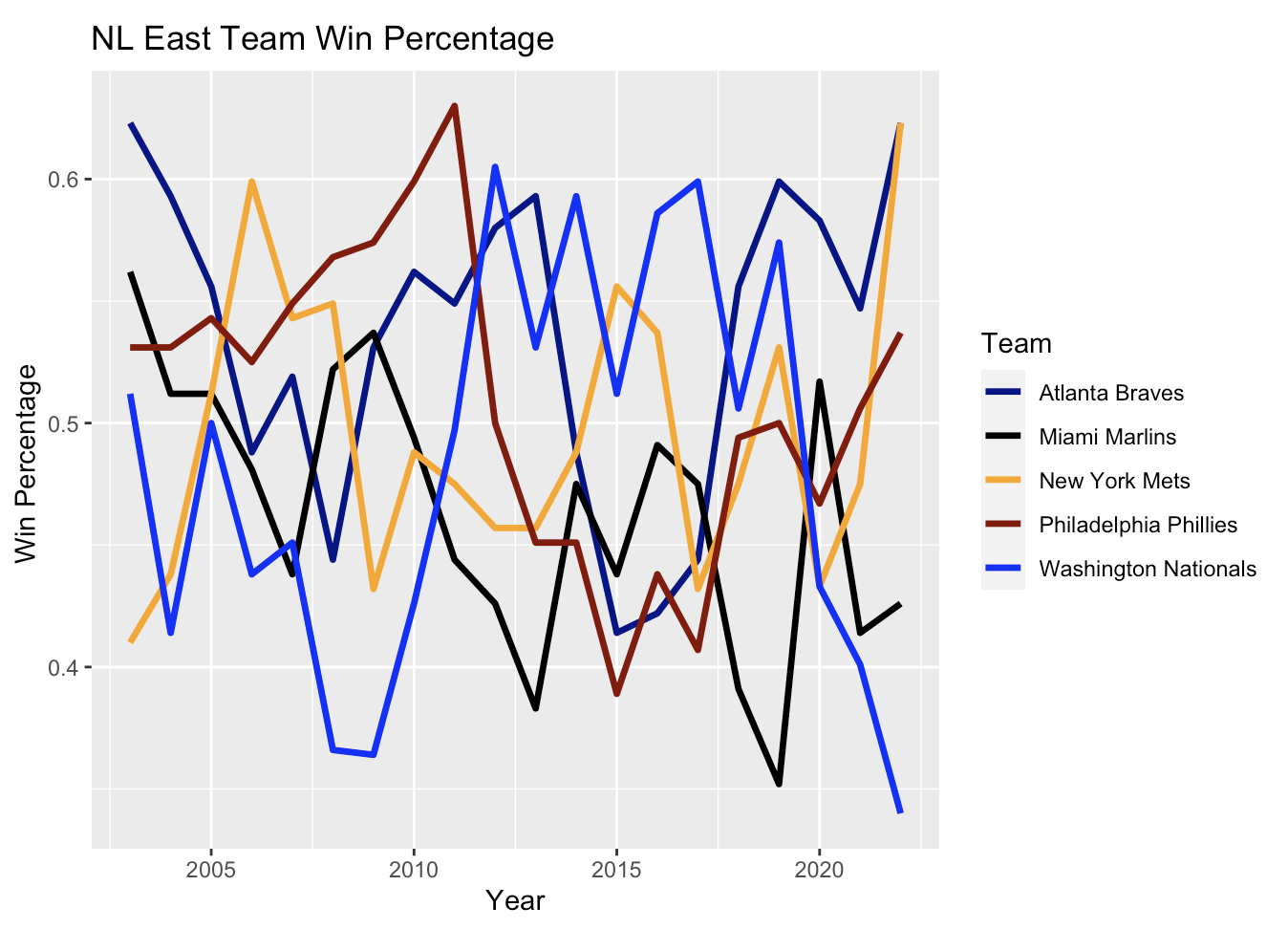
Description automatically generated

The same pattern can be seen here, though no teams are spending as much money year after year as the Yankees. We can see the Athletics here on the AL West graph, consistently one of the lowest spending teams in the league.

### Winning Percentage

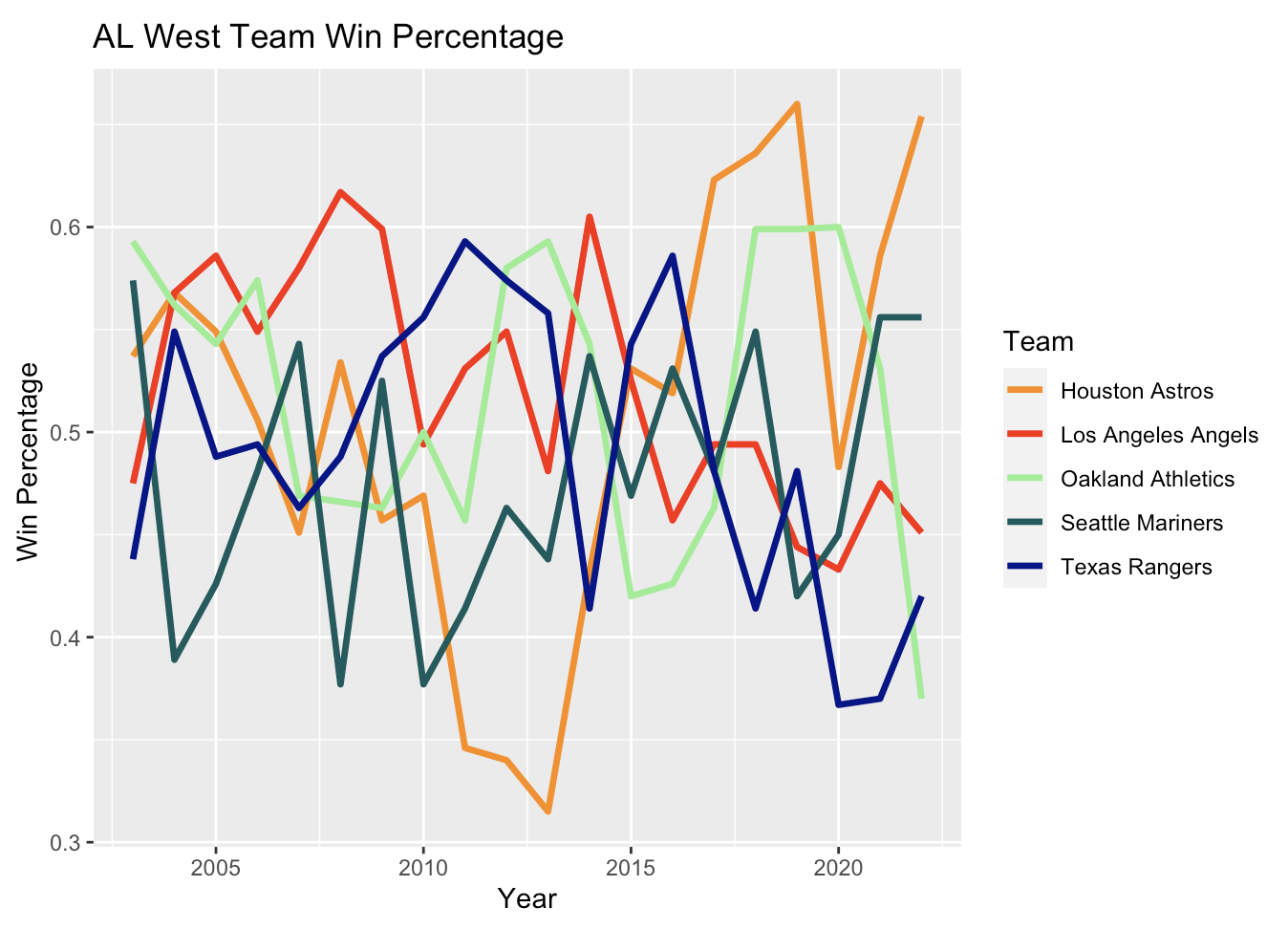
Moving on to graphs of winning percentage over the years, one of the main variables we will be using to evaluate success. These graphs will be in the same format, with teams split up by division in the same manner as seen previously.

A graph of different colored lines

Description automatically generated

A graph showing different colored lines

Description automatically generatedA graph showing different colored lines

Description automatically generatedA graph of different colored lines

Description automatically generated

These graphs exhibit far more stationarity that the graphs of payroll, showing that winning percentage is not as consistent year after year. This is evidence that parity does exist in this league, as the same teams are not always winning the most games. However, this is not enough evidence to confirm this fact, we should continue with our analysis to determine if there is a relationship between spending and team success.

### Population

A graph of different colored lines

Description automatically generated

This graph shows the population of centers that were determined to be “large markets” based on their size. This shows how little change there is in these values over time, especially compared to the payroll plots from earlier. This makes sense, as population values are often observed over a much longer period of time.

A diagram of a box plot

Description automatically generated

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Min.** | 1Q | Mean | 3Q | Max. |
| 286578 | 440034 | 1431513 | 1474210 | 8335897 |
|  |  |  |  |  |

Looking at the boxplot, taking every city above the 3Q value would be a good metric to determine what cities are “large market” cities by population. This ends up being the 7 largest cities in the league, which ends up including 10 teams, which can be seen in the following table.

|  |  |  |
| --- | --- | --- |
| **team** | **city** | **population** |
| New York Mets | New York city, New York | 8335897 |
| New York Yankees | New York city, New York | 8335897 |
| Toronto Blue Jays | Toronto | 6313000 |
| Los Angeles Angels | Los Angeles city, California | 3822238 |
| Los Angeles Dodgers | Los Angeles city, California | 3822238 |
| Chicago Cubs | Chicago city, Illinois | 2665039 |
| Chicago White Sox | Chicago city, Illinois | 2665039 |
| Houston Astros | Houston city, Texas | 2302878 |
| Arizona Diamondbacks | Phoenix city, Arizona | 1644409 |
| Philadelphia Phillies | Philadelphia city, Pennsylvania | 1567258 |

# Methods

## Payroll

### Chi-Square Test

Is there a relationship between playoff appearances and being above the CBT threshold? A chi-square test would be appropriate for this scenario. The null hypothesis is that making the playoffs and being above the CBT threshold are independent. The alternative hypothesis is that those two variables are not independent (there is some relationship between the two).

**H0 = Making the playoffs and being above the CBT threshold are independent**

**Ha = Making the playoffs and being above the CBT threshold are not independent**

A black text on a white background

Description automatically generated

The p-value for this test is essentially 0. We can reject the null hypothesis and conclude that playoff appearances and being above the CBT threshold are not independent, there is some relationship between the two. Conventional wisdom would say that this relationship is a positive one in the sense that being above the threshold makes a team more likely to make the playoffs. Now it would be useful to confirm that intuition and if it is true, then quantify just how much of an advantage being above the threshold gives to a team.

### Independent T-test

Is the average winning percentage for teams below the CBT threshold smaller than the average winning percentage for teams above the CBT threshold?

* Let μa = mean winning percentage for teams *above* the CBT threshold.
* Let μb = mean winning percentage for teams *below* the CBT threshold.

**H0 = μa – μb ≤ 0**

**Ha = μa – μb > 0**

**A white screen with black text

Description automatically generated**

The p-value from this t-test is essentially 0. We have enough evidence to reject the null hypothesis, teams with a payroll above the CBT threshold have a higher winning percentage than teams below the CBT threshold. We are 95% confident that the true difference of means between the winning percentage of the two populations is greater than 0.0575. This works out to a difference of 9.3 wins over an entire 162 game season, meaning we are 95% confident that teams who exceed the CBT threshold win at least 9 more games than teams that do not.

## Population

Is there a relationship between population and winning percentage? What about population and payroll? Is this a problem that is difficult to address because teams in bigger cities are better, not teams with wealthier owners?

### Chi-Square Test

Let’s start by doing the same tests as earlier but using a team’s market size as a variable instead of their payroll size. First, we repeat the chi-square test. The null hypothesis is that making the playoffs and being a large market team are independent. The alternative hypothesis is that those two variables are not independent (there is some relationship between the two).

**H0 = Making the playoffs and being a large market team are independent**

**Ha = Making the playoffs and being a large market team are not independent**

**A close-up of a number

Description automatically generated**

It is close, but a p-value of 0.058 means we cannot reject the null hypothesis at a 95% confidence level. We cannot conclude that there is a relationship between being a large market team and making the playoffs. Since we could reject the null hypothesis at a 90% confidence level, it is still worth continuing to the t-test as we did earlier to figure out how much of a difference being a large market team makes.

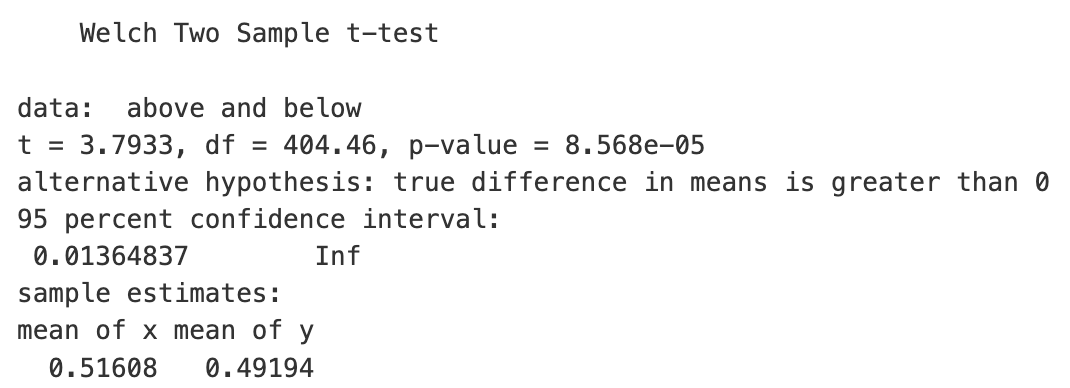
### Independent T-test

Is the average winning percentage for small market teams smaller than the average winning percentage for large market teams?

* Let μl = mean winning percentage for large market teams.
* Let μs = mean winning percentage for small market teams.

**H0 = μl – μs ≤ 0**

**Ha = μl – μs > 0**



The p-value for this test is essentially 0. We have enough evidence to reject the null hypothesis and conclude that large market teams have a better mean winning percentage than small market teams. We are 95% confident that the true difference in mean winning percentage between large and small market teams is greater than 0.0136. This works out to a difference of at least 2 wins over an entire 162 game season meaning we are 95% confident that large market teams win at least 2 more games than small market teams on average.

The fact that the difference between large and small market teams is 2 games per season, while the difference between teams above and below the CBT threshold is 9 games per season is an encouraging sign. Payroll is a much easier factor to control than population. If it were the case that teams in large markets won more games than the league would be helpless as it is not practical to limit population growth for the sake of competitive balance. Limiting payroll, however, is well within the capabilities of the league, and doing so would be much more effective at encouraging league parity than limiting market size.

## Bootstrap Methods

As we only have 600 points of data, bootstrapping may be useful here to get more accurate results. Let’s bootstrap the difference of winning percentage means between teams that exceed the CBT threshold and teams that do not.

A graph of a person with a number of numbers

Description automatically generated with medium confidence

Similar results can be seen here as the earlier t-test, as all differences are greater than 0 and most are around 0.07 and 0.08, showing that teams who surpass the CBT threshold have greater winning percentages than teams that do not.

Running a 95% confidence interval on these bootstrapped data, we get the following result:

|  |  |
| --- | --- |
| 2.5% | 97.5% |
| 11.37 | 13.25 |

We are 95% confident that the true difference in means between winning percentage for teams above and below the CBT threshold is between 11.3 and 13.3 wins each season. This is larger than the previously determined value of 9 wins, further supporting our hypothesis that teams that surpass the CBT threshold win more games.

# Conclusions

The most important conclusion was first seen in our t-test comparing average winning percentages for teams above and below the CBT threshold and then supported by our bootstrapped confidence interval. This interval suggested that teams that have a payroll surpassing the CBT threshold win between 11.3 and 13.3 more games than teams that do not. For simplicities sake, we will condense this conclusion by using the value of 12 wins, as that falls in the middle of this interval.

Teams that spend more on players than the value set by the CBT threshold win 12 more games than teams that do not. This should be a concerning conclusion from a league parity perspective because it means that the penalties for teams that go beyond the threshold are not sufficiently harsh. The tax does exist, but a competitive balance tax should make the league more competitively balanced by taxing – punishing – the teams that spend too much money, making their teams more comprable to those that do not surpass the threshold. The idea behind the CBT threshold is that if a team goes beyond it, they will be punished by having to spend a tax.

This project showed that the CBT threshold is effectively the opposite of a punishment. Teams that go beyond the threshold are not punished for their extravagant spending. Rather, they are essentially rewarded with 12 more wins than teams that do not. This shows that the current system is fundamentally flawed. The threshold system does not punish teams for spending too much money, making the league more competitive. Rather it encourages teams to spend as much as they can, effectively serving as a benchmark for teams to surpass, as if they do, they will have a significant advantage over their counterparts.

Just how much of an advantage is 12 wins each season? Below are the final standings for the 2023 MLB season, courtesy of Baseball Reference[[9]](#footnote-9). Only two out of six teams that finished first in their division did so by more than 12 games. In those four divisions decided by 12 games or fewer, had the division winner finished 12 games worse they would have dropped not just to second place, but all the way down to third place. 12 games is an incredibly large amount of wins to gain simply by surpassing the threshold set in place to discourage teams from doing exactly that.

A screenshot of a table

Description automatically generated

The results of this project are clear. A gap between rich and poor teams clearly exists. The systems in place to discourage rich teams from spending copious amounts of money and cruising towards success while the poor teams watch helplessly is clearly broken. All other major sports leagues in the United States have a salary cap that prevents this exact problem from occurring. Major League Baseball clearly needs to put a hard cap on spending. Their current CBT tax system does not work.

## Future Steps

I would like to continue this work by looking at how big spending influences team success in other professional sports leagues. The system that MLB has in place is clearly broken, do the other professional sports leagues do it better? What would past seasons look like had the MLB instituted a salary cap instead of moving to their current system in 2003? A further exploration of the relationship between salary and team success would be interesting. As the Oakland Athletics showed in 2002, however, it is possible to win games with a low payroll. Another extension of this project could be figuring out what those teams that experience success with a low payroll have in common. They are few and far between, but they do exist. Maybe we could learn something from their success.

# References

“Baseball Reference,” 2023. https://www.baseball-reference.com/teams/.

Bureau, US Census. “City and Town Population Totals: 2020-2022 — Census.Gov,” n.d. https://www.census.gov/data/tables/time-series/demo/popest/2020s-total-cities-and-towns.html.

“MLB’s Evolving Luxury Tax --- Blogs.Fangraphs.Com,” n.d.

OverTheCap. “Salary Cap Space,” 2023. https://overthecap.com/salary-cap-space.

Retrosheet. “2002 Payroll,” 2002. https://www.thebaseballcube.com/page.asp?PT=payroll\_year&ID=2002.

“World Population Prospects 2022,” 2022. https://population.un.org/wpp/

# Appendix

All code for this project can be found here: <https://github.com/tsigall/dsan5100-final>. Please consult the ReadMe for more information.

1. Retrosheet, “2002 Payroll.” [↑](#footnote-ref-1)
2. Retrosheet. [↑](#footnote-ref-2)
3. OverTheCap, “Salary Cap Space.” [↑](#footnote-ref-3)
4. “MLB’s Evolving Luxury Tax --- Blogs.Fangraphs.Com.” [↑](#footnote-ref-4)
5. Retrosheet, “2002 Payroll.” [↑](#footnote-ref-5)
6. “Baseball Reference.” [↑](#footnote-ref-6)
7. Bureau, “City and Town Population Totals: 2020-2022 — Census.Gov.” [↑](#footnote-ref-7)
8. “World Population Prospects 2022.” [↑](#footnote-ref-8)
9. “Baseball Reference.” [↑](#footnote-ref-9)