California's Unlevel Playing Field

Kylie Scharf[1], Nicole Yee[1], Curtis LeQuang[1], Emma Durler[1]

[1]Department of Statistics, Cal Poly - SLO

June 5, 2024

Abstract

In our report, we investigated whether California's collegiate expenses from 2015 to 2019 are truly greater for baseball than for softball. Under Title IV federal student financial assistance programs, every year colleges across the United States must open their books and reveal the inner workings of their athletic departments. This requirement aims to ensure transparency in athletic funding, participation, and staffing. Since 2003, the Department of Education has collected data from 2,074 schools nationwide through the Equity in Athletics Survey. Focusing on the expenses of specific sports can help us determine if equality in funding is truly ensured. Our research contains a subsetted dataset of 659 observations from 2015 to 2019. Each observation consists of a Californian school in a particular year with a collegiate baseball or softball team. From this data, we created a new variable to represent the difference in funding between baseball and softball for these 659 observations. We created multiple visualizations to better view observed average differences in expenses throughout the years. We then performed a paired t-test, finding convincing evidence that baseball expenses were significantly higher than softball expenses from 2015 to 2019 for California colleges like the ones observed. Additional research is needed to explore the reasons behind these disparities and to identify effective strategies for achieving funding equality. Studies could also be expanded to other states and sports to provide a broader understanding of the issue.

1 Introduction

In the sunny state of California, where the crack of the bat and the cheers of fans echo through the air, a tale of two sports unfolds. One is blessed with seemingly unlimited funds and admiration, and the other is left to fend for itself. A stark contrast in funding has long persisted between baseball and softball programs. For decades, college baseball has enjoyed a lion's share of financial support, with state-of-the-art facilities, top-tier coaching staff, and generous scholarship allocations. Meanwhile, softball programs, though equally deserving of investment and recognition, are often overlooked and underfunded.

Dana Drew Shaw (2019) writes of specific examples of this, where many times softball teams are not granted access to locker rooms and are forced to change in the parking lot while baseball teams have locker rooms and offices right next to the fields they play on.

The disparity serves as a barrier to the growth and development of female athletes, perpetuating inequality and hindering their ability to compete at the same level as their male counterparts. Denied access to resources and opportunities, female athletes are at a disadvantage both on and off the field, facing obstacles in their pursuit of academic and athletic success.

In our report, we aimed to answer the question of whether or not the average expenses of collegiate baseball are truly greater than softball in California from 2015 to 2019. The following section introduces the data, subsetting methods used to confine the data to only contain California colleges, and statistical analyses used to carry out our investigation. We then report our findings and discuss the implications of such. Finally, we provide a conclusion and direction for future research.

2 Data and Methods

2.1 Data Collection

The Equity in Athletics Disclosure Act requires all US colleges with a Title IV federal student financial assistance program to report on athletic funding, participation, and staffing for each sport and academic year (Office Of Postsecondary Education 2021). Since 2003, The Department of Education has collected data on collegiate sports budgets at 2,074 schools nationwide in obeyance of the act. The department created the Equity in Athletics Survey to collect such data. Each school with such Title IV federal student financial assistance programs completes this survey annually for each of its athletics teams. We accessed this data through the Equity in Athletics Data Analysis (2003-2023). The obtained data set contains all survey responses from 2015 to 2019. Although the data set contains various types of funding information for all participating schools, we are specifically interested in the expenses, in dollars, for each school's softball and baseball program.

2.2 Statistical Analysis

Prior to analysis, we subsetted the data to include only California schools's softball and baseball programs. Our final data set included 659 records of information on the expenses, in dollars, of schools in California from 2015 to 2019 for softball and baseball alike. Specifically, our final data set includes 133 universities in California in 2015, 130 universities in 2016, 132 universities in 2017, 132 universities in 2018, and 132 universities in 2019. We subsetted the data using the `tidyverse` (Wickham et al. 2019) and `readxl` (Wickham and Bryan 2023)

packages. We also used the `writexl` package (Ooms 2024) to create permanent Excel data sets. We then pivoted the data so that each school in each year would represent one observation and contain information about both the softball and baseball expenses at that school. We also created a new variable to represent the difference in expenses for baseball and softball (baseball - softball) at each school in each year. We displayed these average differences in a table using the `kableExtra` package (Zhu 2024). We intended to average these differences to determine whether or not these expenses are truly greater for baseball than for softball. To do such, we performed a paired t-test at a 5% significance level to analyze the average differences, averaged over all 5 years, and test whether or not there is a discernible difference in average expenses between baseball and softball.

We recognize that baseball and softball are not exactly equivalent sports, which will impact some of the differences between the data. Baseball and softball, while similar, have distinct rules, equipment, field dimensions, and player dynamics. These differences will therefore impact the expenses of each sport. For example, baseball often involves more games per season, longer game durations, and potentially more travel. This leads to increased costs related to transportation, housing, and meals. We will keep these differences in mind when completing our data analysis and hope to address as much of the impact of the differences as possible.

3 Results

In this report, we wanted to determine if, for the years 2015-2019, California collegiate expenses for baseball were on average greater than softball. California colleges reported their annual individual expenses for baseball and softball teams each year. Our data includes the expenses for softball and baseball teams at 133 universities in California in 2015, 130 universities in 2016, 132 universities in 2017, 132 universities in 2018, and 132 universities in 2019. Table 1 breaks down each year's average difference in expenses between the baseball and softball teams of an individual school, the standard error of such difference, and the number of participating California colleges.

Table 1: The average differences between baseball and softball expenses, the variability in such, and the number of participating California colleges for each year between 2015 and 2019.

Year	Difference	Standard Error	Universities
2015	$126,459.80	$20,152.60	133
2016	$147,310.83	$22,632.39	130
2017	$150,482.42	$24,847.36	132
2018	$168,719.20	$24,457.78	132
2019	$156,763.04	$23,088.92	132

Note:

Source: Department of Education

After looking at the table it becomes fairly clear that the expenses for baseball are much higher than the expenses for softball every year. The difference between the sports is over $100,000 every year (Table 1). Both baseball and softball expenses also appear to be increasing over time from 2015 to 2019 and then decrease in 2019. The variability, in terms of standard error, of the differences appears to remain relatively constant but slightly increases from 2015 to 2018 and then similarly decreases in 2019. The sample sizes also remain relatively constant, representing 130 to 133 schools per year (Table 1).

After performing a paired t-test, we found evidence that baseball expenses were discernibly greater than softball ($t = 14.542$; $df = 658$, $p < 0.0001$). Baseball expenses averaged about $149,919 more than softball expenses (95% CI: 129676.6, 170162.3). More specifically, we are 95% confident that at the collegiate level, baseball expenses were, on average, between $129,676 and $170,162 greater than softball expenses.

More prominently displayed in Figure 1, we can see that the average differences in expenses for a particular year are well above $0. The highlighted section in the figure represents a roughly 95% confidence interval for the true difference in expenses, where the width is two standard errors of the average difference for each individual year. The specific standard error for that particular year is shown above in Table 1. The entirety of the interval is clearly above 0 in the visualization.

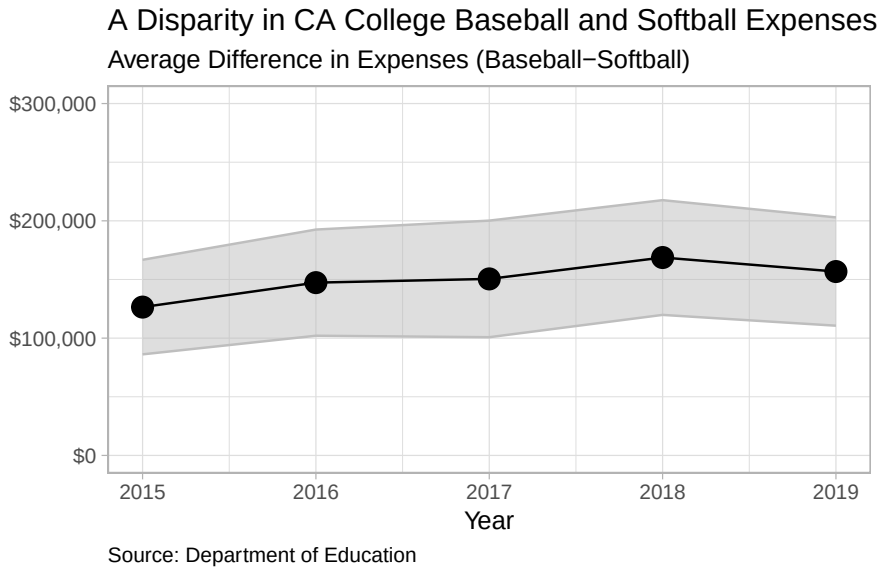

Figure 1: Above entails a visualization of the average differences in expenses from 2015 to 2019 for collegiate baseball and softball in California. The average difference in expenses for a particular year for all California teams is plotted as a black circle. The gray highlighted region represents adding or subtracting two standard errors of the average difference from the plotted average differences. This shaded region therefore represents a roughly 95% confidence interval for the true average difference for all collegiate baseball and softball teams in California between 2015 to 2019 that meet the conditions required for data collection for the Equity in Athletics Survey by the Department of Education. The average differences in expenses are consistently greater than $0 along with the interval for said average differences, meaning the average expenses for baseball are consistently significantly larger than for softball.

4 Discussion

Our analysis consisted of a paired t-test performed in R that compared the difference in expenses between collegiate baseball and softball in California from 2015 to 2019. We found statistically significant evidence to reject the null hypothesis and conclude that baseball expenses are significantly greater than softball expenses. We are 95% confident that at the collegiate level, baseball expenses were, on average, between $129,676 and $170,162 greater than softball expenses throughout the time period. A difference of between $129,676 and $170,162 is very large for a collegiate sports team so we therefore believe our results are practically significant as well as statistically significant. This substantial difference has real-world implications for resource allocation, program sustainability, equity, and strategic decision-making within collegiate athletic departments.

Although the differences in expenses between the sports were proven statistically significant, there are limitations to such a conclusion. One such limitation is that we cannot assume causation from these results as our data came from an observational study. We can therefore not make the claim that gender is the cause of the significant difference in average expenses. The data being collected from an observational study also implies that we can only generalize the results obtained to colleges/universities similar to those included in the study. That is, we can only conclude that the expenses for baseball are significantly higher than softball for schools in California from 2015 to 2019 with a Title IV federal student financial assistance program.

The result we obtained was expected as this is what past research and experiences support. One plausible reason for the statistically significant difference in expenses is the viewership difference between the two sports. Collegiate baseball traditionally attracts a larger audience than softball. This is reflected in higher attendance at games, more significant media coverage, and better TV ratings. The larger viewership for baseball translates into higher revenue from ticket sales, broadcasting rights, and sponsorships. This increased revenue may explain some of the differences in expenses between the sports. Athletic departments might then allocate more resources to baseball due to its potential for higher revenue generation, leading to better funding for baseball compared to softball, and continuing the cycle of funding inequality. We also mentioned in the data collection section that these two sports are not completely equivalent, which likely also explains some of the differences in expenses.

Amidst the challenges lie opportunities for change. Such numerical disparity serves as a rallying cry for equity and fairness within collegiate athletics. It is a reminder that every athlete, regardless of gender or sport, deserves access to resources and opportunities that will enable them to succeed on and off the field. Collegiate athletics must come together to address this issue head-on, committing to providing female athletes with the support they need to thrive, both academically and athletically. It's not just about leveling the playing field; it's about creating a culture of inclusivity and empowerment, where every athlete, regardless of gender or sport, has the chance to succeed.

5 Conclusion

Through our investigation of the difference in expenses between collegiate baseball and softball, we found that baseball consistently has higher average expenses than softball in California from 2015 to 2019. However, since the data was collected from an observational study, we cannot assume causation from these results. We therefore cannot make the claim that gender is the reason these two sports have such a large difference in expenses. The data being collected from an observational study also implies that we can only generalize the results obtained to colleges/universities similar to those included in the study. That is, we can only conclude that the expenses for baseball are significantly

higher than softball for schools in California from 2015 to 2019 with a Title IV federal student financial assistance program.

We created multiple visualizations, such as a line graph and a table, and performed a paired t-test to come to this conclusion. This paired t-test was conducted using the differences in expenses between the baseball and softball teams of 133 universities in 2015, 130 universities in 2016, 132 universities in 2017, 132 universities in 2018, and 132 universities in 2019. While we have extremely strong evidence of this statistically, there is further supporting evidence through the visual trends in our report. As the entirety of the 95% confidence interval is above $100,000 per year, we believe our results are not only statistically significant but practically significant as well. In a sports program, $100,000 goes a long way in terms of the funding for a single collegiate team in a specific year.

For further research, we wonder if this pattern of the difference in average expenses between collegiate baseball and softball exists through 2024, but since our data only includes data up to 2019, we can only assume. Additional research is needed to explore the reasons behind these disparities and to identify effective strategies for achieving funding equality. Studies could also be expanded to other states and sports to provide a broader understanding of the issue. More research could also be conducted on the decrease in funding observed in 2019 for both baseball and softball.

References

Office Of Postsecondary Education. 2021. "TREND: Office of Postsecondary Education. College Athletics: Total Participation | State: California | Educational Institution: Loyola Marymount University | Collegiate Sport: Rowing | Variable: Total # Participants Women, 2005 - 2018. Data Planet™ Statistical Datasets: A SAGE Publishing Resource Dataset-ID: 069-001-001." https://doi.org/10.6068/DP176F7A216C817.

Ooms, Jeroen. 2024. "Writexl: Export Data Frames to Excel 'Xlsx' Format." https://CRAN.R-project.org/package=writexl.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the {Tidyverse}" 4: 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, and Jennifer Bryan. 2023. "Readxl: Read Excel Files." https://CRAN.R-project.org/package=readxl.

Zhu, Hao. 2024. "kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax." https://CRAN.R-project.org/package=kableExtra.