

# Does a Good Spring Increase the Mean? An Analysis of the Carryover Effect of MLB Players Spring Training Batting Performance to the Regular Season

Scott Powers<sup>1</sup>, Austin R. Brown, Ph.D.<sup>2</sup>

<sup>1</sup> Oak Mountain Academy, Carrollton GA

<sup>2</sup> Kennesaw State University, Department of Statistics & Analytical Science, Kennesaw GA

**Abbreviated abstract:** In Major League Baseball (MLB), Spring Training is an approximately month long preseason where players have the opportunity to practice and develop skills in live games. To the casual fan, when a fielder has a particularly strong batting performance, it is typically expected that the “hot streak” that player enjoyed will carry over into the beginning of the regular season. To assess this claim, 11 years of Spring Training and Regular Season data were analyzed to see what, if any, carryover effect exists using Regular Season batting average as the outcome variable.

## **Related publications:**

– Goossens, D. R., & Spieksma, F. C. (2012). The carryover effect does not influence football results. *Journal of Sports Economics*, 13(3), 288-305.

# Problem, Data, Previous Works

In sports, we've seen instances of players getting "streaky," meaning that they have sustained periods of excellent or poor performance. In MLB, fans have seen instances of players having excellent Spring Trainings in terms of batting performance. As is the case in fandoms, we expect the Spring Training "hot streak" to carry over into the beginning of the regular season.

However, a search into the carryover effect for team sports yielded only a single article (Goossens & Spijksma, 2012) which determined that no carryover effect was observed in an analysis of Belgian soccer.

Additionally, because of the rise of Daily Fantasy Sports betting (6 million registered Fan Duels users and 8 million registered Draft Kings users, according to Investopedia.com), it would be advantageous for users to know if Spring Training performance ought to be a consideration in drafting strategy.

Thus, the research question driving this study is:

***Does Spring Training performance in terms of batting average have an effect on batting average at the beginning of the regular season, controlling for other relevant variables?***



KENNESAW STATE  
UNIVERSITY

[abrow708@kennesaw.edu](mailto:abrow708@kennesaw.edu) - 2

UCSAS  
2020

# Methods

## Data Collection

Regular season batting data was gathered by year from Fangraphs. Python scripts were used to calculate and append covariates including Regular Season Opponent ERA. Spring training batting data was not found in an accessible format. Python packages urllib, Selenium, and BeautifulSoup were used to download data from html tables on the MLB Stats website and convert the tables to CSV sheets.

Spring training data was restricted to players with at least 30 PA while regular season data was restricted to the first month of the season and to players with at least 50 ABs. Data was taken from the 2009 – 2019 seasons and the total analyzed sample size was  $n = 2557$ .



## Analysis Plan

While the primary research question aims to determine the carryover effect of Spring Training performance, in terms of Batting Average and Strikeout per Plate Appearance, on Regular Season Batting Average (restricted to the first month of the Regular Season), other covariates controlled for were: *Season, Regular Season Opponent ERA, Regular Season Hits, Regular Season PA, Previous Regular Season wOBA*.

Because the outcome variable is on the  $[0, 1]$  interval, a **beta regression** was fit using the `betareg` function in R.



# Results and Conclusions

As shown, a player's batting average in the Spring was not found to be significantly associated with batting average during the beginning of the regular season. This suggests that a particularly good or bad spring training won't carry over into the regular season.

All the other predictor variables were found to be significant, most of which being intuitive. Surprisingly, Strikeouts per PA in Spring Training was found to be significantly and positively associated with regular season batting average. One plausible explanation could be that this is a proxy for playing time, meaning that as a player plays more in the Spring, the better we can expect them to perform in terms of batting average in the regular season.

## *Results of Beta Regression*

Variable Name	Exponentiated Coefficient	P-Value
Season	0.9983	0.0084
Regular Season PA	0.9871	< 0.0001
Regular Season Hits	1.0599	< 0.0001
Previous Regular Season wOBA	1.5310	< 0.0001
Regular Season Opponent ERA	1.0284	0.0004
Spring Training Strikeouts per PA	1.1062	0.0005
Spring Batting Average	1.0009	0.9742