The Effect of Major League Baseball's New Rule Changes on the Amount of Stolen Bases

ECMT 463
Professor Georges Naufal
4 August 2023
Christopher Thomas
christopher.thomas@tamu.edu

Abstract

The purpose of this research paper was to determine if Major League Baseball's rules changes introduced in the 2023 season were successful in increasing the number of stolen bases per game. The game logs for every MLB team from every available game in the 2022 season and the 2023 season up until July 22, 2023, were compiled into a pooled cross-sectional data set from baseball-reference.com to conduct my research. I hypothesized that the new rules introduced would generate a positive correlation in the number of stolen bases for each game.

Statement of Contribution

I, Christopher Thomas, was responsible for compiling the data, generating the regression analysis, writing the report, and evaluating the regression results.

Section I: Introduction

Major League baseball has been researching and testing many methods to increase viewership of its product over the years. They found that the most effective way to do this was to "[improve] pace of play, action, and safety at the MLB level" (Castrovince 2023). The MLB wanted the average baseball game to include more stolen bases, more hits, more athletic defensive plays, safer gameplay, and they wanted games to end faster. The MLB decided to implement three new rules to try and achieve these goals. MLB introduced a pitch timer system, restrictions on defensive shifting, and an increase in base size for 1st, 2nd, and 3rd base. MLB introduced all three of these rule changes at the same time beginning in the 2023 season. In my research, I decided to measure the success of MLB's rule changes by determining their collective effect on the number of stolen bases per game by a team.

The pitch timer is a system in place in which the pitcher is allowed 15 seconds to get set on the mound and deliver a pitch to the batter. The pitcher gets 15 seconds if there are no runners on base and 20 seconds if there are runners on base. After a plate appearance ends, the pitcher has 30 seconds to get set on the mound and deliver a pitch. The batter is also required to be in the batter's box and facing the pitcher by the 8-second point. If either the pitcher or the batter is in violation, they receive a penalty. This rule also limits the amount of pickoff attempts a pitcher has on a runner to three attempts before the runner is awarded a stolen base.

The next rule put in place restrictions on defensive shifts. It is now required that defending teams have a minimum of four infielders within the boundaries of the infield and there must be two infielders on either side of second base. The infielders cannot change which side of the infield they are on from batter to batter. Penalties are administered for any violation of the defensive shifting rules.

Lastly, the most easily enforceable rule implemented in the 2023 season is the increase of base size. In the 2022 MLB season the size of 1st, 2nd, and 3rd base were 15 square inches. The bases were increased to a size of 18 square inches going into the 2023 MLB season. The increase in base size resulted in a "4 ½-inch reduction in the distance between first and second base and between second and third, which encourages more stolen-base attempts" (Castrovince 2023). It was hypothesized that this change in size could also improve player safety for both the defense and offense.

Determining the collective effect of these three rule changes is important for the MLB because it helps them determine if their product is becoming more entertaining for the audience. If these rule changes do create the desired increase in action and faster pace of play, then MLB can decide whether or not to make these rule changes permanent or whether to dispose of them going into any future seasons. These changes may also provide insight on how to change the rules of the game in the future to improve watchability of the product while also safeguarding the integrity of the sport.

I hypothesized that the introduction of these three new rules at the start of the 2023 MLB season would increase the number of stolen bases per game by a team. I compiled pooled cross-sectional data using the game logs found on baseball-reference.com. I used this data to prove that these rule changes did have a positive correlation on the number of stolen bases by a team in the average game. Section II of the paper is a literature review of related studies conducted on this topic. The next section will describe the model specification and discussion of variables. Finally, the paper will end with an empirical analysis of the data followed by a brief conclusion.

Section II: Literature Review

The Minor League Baseball organization affiliated with Major League Baseball has been a great way to test the effects of potential changes on the rules of the game. Zachary Houghtaling conducted a data analysis regarding Minor League Baseball's rule changes in 2021 and summarized the results of the study in a research report. He examined the effect of increasing the base size from 15 to 18 square inches, limiting pickoff attempts, and requiring the pitcher to remove his foot from the rubber on the mound to initiate a pickoff attempt. He found that these "rule changes have the potential to change how baseball is played at the professional level by increasing stolen base success rates and the amount of [stolen base] attempts in a game" (Houghtaling 2022). His research is very relevant to mine because in examining an entirely different data set, Minor League Baseball, his results led him to a similar conclusion that some of these rule changes influence the number of stolen bases per game.

Alice Carlson of Capital News Service conducted an evaluation of the 2023 MLB season as of April 27th to determine the effect of the rule changes on the game. She found that most pitchers had to increase their pitch tempo in order to not violate the pitch clock rule. She also found that this shortened the length of the games as well as increased the number of stolen bases in the season. She then referenced data from Minor League Baseball's 2022 season and found that "stolen bases skyrocketed in 2022 after the base size was increased, with a 25% increase in stolen bases" (Carlson 2023). Her findings support the findings in my research that these particular rule changes can have a positive impact on the number of stolen bases in games. Her data also supports the idea that the rules result in a faster, more action-packed game.

Jay Cohen of AP News wrote an article preceding the 2023 MLB season in which he discussed the potential effects these rule changes could have on stolen base rates. He received

quotes from various players and managers that felt as though the change from 15 square inches to 18 square inches was a very noticeable difference that could influence the success rates of stolen base attempts. Cohen also found that in the past decade the number of stolen bases as well as the number of stolen base attempts have drastically decreased in Major League Baseball. This finding suggests it's reasonable that MLB should implement changes to increase action in games to bring its product up to the standard it used to be. The article also references the implementation of large bases in two minor leagues in half of their 2021 seasons where "one experienced a 2.2% increase in successful steals, and the other posted 0.7% increase" (Cohen 2023). These findings support the idea that an increase in base size would cause the desired effect of increasing the number of stolen bases.

Section III: Model Specification

The most efficient way to discover the relationship between the number of stolen bases by a team per game and the new rule changes in 2023 was to gather pooled cross-sectional data with the two time periods being 2022 and 2023. The data was acquired from baseball-reference.com, a website which contains the game logs for each game each team plays. Each game log has data on the batting statistics for selected team that game. Data from every game in the 2022 season and from every game in the 2023 season that had been played up until July 22, 2023, was used in this study. The data was compiled July 23, 2023, which is why less data is available from the 2023 time period. From all games observed in 2022 and 2023, two observations could be drawn from every single game because offensive statistics were available for the two teams playing in the game. In total 3,913 games were observed across the 2022 and 2023 seasons resulting in 7,826 total observations in my data set. There were 2,430 games observed in 2022 and 1,483 games observed in 2023. This means that there are 4,860 observations from 2022 and 2,966 observations from 2023. The data collected was observed at the Major League Baseball level.

I believe there is very minimal omitted variable bias on β_1. The variable associated with this coefficient is the presence or lack thereof of the new rules in the game. The determinant of the new rules in the game is whether the game was played in 2022 or 2023. The decision to implement the rules was a cause of the leaders of MLB evaluating previous season data sets, outside of my data set, and desiring a change in outcome of the games. The reason I chose to control for the variables I did was because I believed the variables selected could have a significant impact on Y.

An OLS regression model was used to estimate the relationship between our independent variables and the dependent variables. The t-test will be used to evaluate the results of the

regression. The t-test will be used to calculate the significance of β_1. The null hypothesis and proposed regression model are shown below.

$$H_0: \beta_1 = 0$$

$$H_A: \beta_1 \neq 0$$

$$SB = \beta_0 + \beta_1 Bsize + \beta_2 BB + \beta_3 H + \beta_4 GDP + \beta_5 SO + \beta_6 HBP + \beta_7 ROE + \beta_8 IBB + \beta_9 CS$$

Summary Table of Variables:

```
. summarize SB Bsize BB H GDP SO HBP ROE IBB CS
```

Variable	Obs	Mean	Std. dev.	Min	Max
SB	7,826	.5888065	.9066865	0	9
Bsize	7,826	.3789931	.4851674	0	1
BB	7,826	3.129313	2.004249	0	17
H	7,826	8.242525	3.379542	0	29
GDP	7,826	.7007411	.8111945	0	5
SO	7,826	8.474827	2.925607	0	20
HBP	7,826	.4233325	.6682675	0	5
ROE	7,826	.2260414	.4834821	0	4
IBB	7,826	.0940455	.3385333	0	5
CS	7,826	.1731408	.4082839	0	3

Explanation of Variables:

SB – Dependent variable. A discrete variable that represents the total amount of stolen bases for the team by the end of the game.

Bsize – X_1. Primary independent variable. This is a binary variable which is 1 if the game occurred in the 2023 season and 0 if the game occurred in the 2022 season. Given that the three rules were all introduced in 2023 at the beginning of the season, this variable helps distinguish which observations were taken with the new rules.

BB – X_2. Control variable. A discrete variable that represents the total amount of walks gotten for the team by the end of the game.

H – X_3. Control Variable. A discrete variable that represents the total amount of hits gotten for the team by the end of the game.

GDP – X_4. Control Variable. A discrete variable that represents the total amount of times a team grounded into a double play by the end of the game.

SO – X_5. Control Variable. A discrete variable that represents the total amount of strikeouts received for the team by the end of the game.

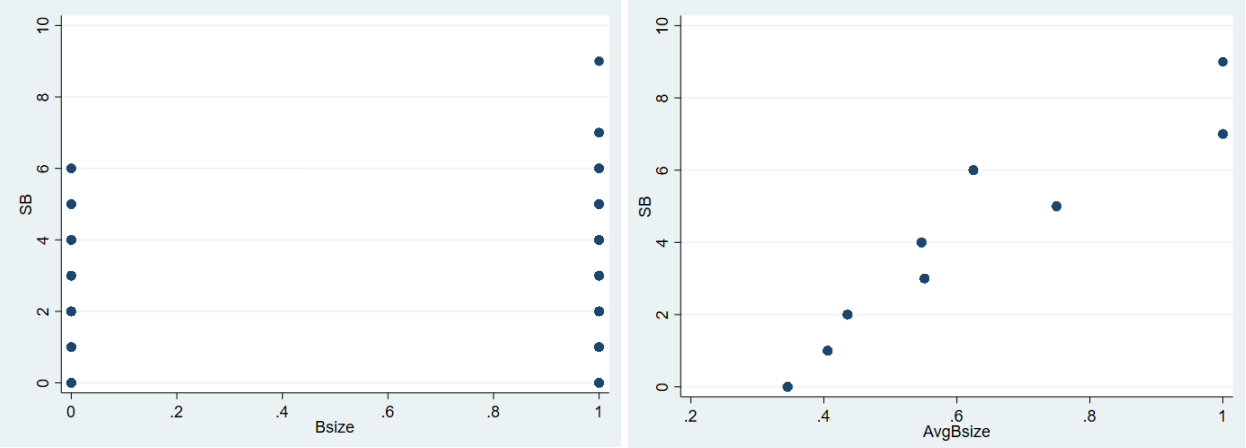
HBP – X_6. Control Variable. A discrete variable that represents the total amount of times a team got hit by a pitch by the end of the game.

ROE – X_7. Control Variable. A discrete variable that represents the total number of times a team reached base on an error by the end of the game.

IBB – X_8. Control Variable. A discrete variable that represents the total amount of intentional walks received for the team by the end of the game.

CS – X_9. Control Variable. A discrete variable that represents the total amount of times a team got caught stealing by the end of the game.

Section IV: Empirical Data



Scatter plots displaying the correlation of the data are shown above. The graph on the left plots the variable SB versus the variable Bsize. This graph does show that there is a correlation between the number of stolen bases by a team and the presence of the new rules in the game, but it is not the best way to depict the correlation. Using Stata, average value for Bsize was taken at each of the nine SB values in the data set using the command shown below.

```
. bysort SB: egen AvgBsize = mean (Bsize)
```

Given that the Bsize variable is binary and equal to 0 when the game was played in 2022 and 1 if played in 2023, the average Bsize for each unique value of SB is equal to the percentage of observations taken in 2023 divided by 100. These values were loaded into the AvgBsize variable. The second graph, on the right, shows the plot of SB versus this new AvgBsize variable. This graph effectively shows that there is a positive correlation between the number of stolen bases by a team and the introduction of the three new rules.

A potential source of bias in the plot on the right results from the number of observations in 2022 compared to the number of observations in 2023. Since 4,860 observations were taken from 2022 and 2,966 observations were taken from 2023, there could generally be an inflated percentage of observations from 2022 for each level of stolen bases.

SB	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Bsize	0.204*** (0.021)	0.187*** (0.021)	0.187*** (0.021)	0.188*** (0.021)	0.186*** (0.021)
BB		0.066*** (0.005)	0.066*** (0.005)	0.066*** (0.005)	0.068*** (0.005)
H		0.026*** (0.003)	0.026*** (0.003)	0.026*** (0.003)	0.025*** (0.003)
GDP		-0.079*** (0.013)	-0.079*** (0.013)	-0.083*** (0.012)	
SO		0.008** (0.004)	0.008** (0.004)		
HBP		0.030** (0.015)	0.030** (0.015)		
ROE		0.050** (0.021)	0.050** (0.021)		
IBB		0.117*** (0.030)	0.117*** (0.030)	0.123*** (0.030)	
CS		-0.002 (.024)			
constant	0.512*** (0.013)	0.054 (0.046)	0.054 (0.046)	0.147*** (.030)	0.102*** (0.030)

**Significance at the 5% level

*** Significance at the 1% level

The OLS regressions of the data were calculated using Stata and shown above. Regression 1 was chosen as the baseline regression to determine the effect of Bsize on SB without controlling for any other variables. Regression 2 was then taken to factor in all of the relevant control variables. Regression 3 was chosen to factor in all of the control variables that were found to be statistically significant at the 5% level based off of regression 2. Regression 4 was chosen to factor in the control variables that were statistically significant at the 1% level based off of regression 2. Regression 5 was chosen to determine the change on the β_1 after removing most of the control variables.

For regressions 2 through 4, β_1 remained fairly constant between a range of 0.186 and 0.188. β_1 in the first regression was 0.204 which varies slightly from the other regressions but is similar enough to conclude that there is very little omitted variable bias from the control variables. I decided the best regression for the data was regression 3 because all variables included were statistically significant within the 5% confidence level and the β_1 suffered least from omitted variable bias. The regression equation for model 3 is shown below.

$$SB = \beta_0 + \beta_1 Bsize + \beta_2 BB + \beta_3 H + \beta_4 GDP + \beta_5 SO + \beta_6 HBP + \beta_7 ROE + \beta_8 IBB$$

I determined that the null hypothesis, $\beta_1 = 0$, could be rejected based off of the regression data collected. In each regression β_1 was statistically significant at the 1% confidence level and had very high t-statistics. This means that the actual β_1 had to be significantly different from 0. This implies that the presence of the three new rules in MLB's 2023 season did have a significant positive impact on the number of stolen bases by a team per game. The presence of the new rules increased the average number of stolen bases by a team per game by 0.187. This means that the average number of stolen bases per game increased by 0.374 because two teams play in each game. This means that the average number of stolen bases per season should increase by 908.82 stolen bases because 2,430 total games are played each season.

Section V: Conclusion

In the 2023 MLB season, the MLB introduced three new rules to help improve pace of play, action in the game, and safety. I conducted the research in this paper to determine if the introduction of these new rules was effective in increasing action in the game. I chose to examine whether the presence of these three rules had an effect on the number of stolen bases by a team per game. Determining whether the rules do increase action in the game provided insight on whether these rules should remain in the game. In my research, I found that the number of stolen bases by a team per game did increase with the presence of the new rules.

For my regression analysis, I gathered pooled cross-sectional data on games occurring in the 2022 and 2023 MLB seasons. My target independent variable was the presence or lack of the new rules. I had 8 control variables for walks, hits, grounding into double plays, strikeouts, hit by pitches, reaching on error, intentional walks and caught stealing attempts. Using these control variables helped me confirm I had minor omitted variable bias and helped me determine that there was a positive correlation between playing with the new rules and number of bases stolen. I found that each season played using these new rules would generate an increase of around 909 stolen bases across the league. MLB was successful in increasing the amount of action in the game by introducing these three new rules into the 2023 season.

I, Christopher Thomas, was responsible for compiling the data using game logs from baseball-reference.com. I was also responsible for running an OLS regression analysis using Stata BE. I read other articles and wrote up a literature review on similar subject matter. I then compiled all this information into this research paper, writing the abstract, introduction, literature review, model specification, empirical data, and conclusion.

References

Carlson, Alice. “New MLB Rules Lead to Shorter Games and More Stolen Bases.” *CNS Maryland*, 2 May 2023, cnsmaryland.org/2023/04/27/new-rules-in-the-mlb-lead-to-shorter-games-and-more-stolen-bases/#:~:text=The%20number%20of%20stolen%20bases%20increased%20drastically%20after%20the%20implementation,a%20pitch%20clock%20in%202022.&text=At%20the%20minor%20league%20level%2C%20stolen%20bases%20skyrocketed%20in%202022,bases%2C%20according%20to%20Bleacher%20Nation.

Castrovince, Anthony. “Pitch Timer, Shift Restrictions among Announced Rule Changes for ’23.” *MLB.Com*, 13 Feb. 2023, www.mlb.com/news/mlb-2023-rule-changes-pitch-timer-larger-bases-shifts.

Cohen, Jay. “MLB’s Bigger Bases Could Lead to More Steals, Fewer Injuries.” *AP News*, 23 Feb. 2023, apnews.com/article/mlb-sports-baseball-manny-machado-san-diego-padres-9136a7f751d67698dfb48d4bce5db7a2.

Houghtaling, Zachary. *Analysis of Minor League Rule Changes Effect on Stolen Bases*, Apr. 2022, ideaexchange.uakron.edu/cgi/viewcontent.cgi?article=3011&context=honors_research_projects.