
Munich Personal RePEc Archive

An Econometric Analysis of the 2013 Major League Baseball Season

Fullerton, Steven L. and Fullerton, Thomas M., Jr. and Walke, Adam G.

University of Texas at El Paso

11 February 2014

Online at https://mpra.ub.uni-muenchen.de/59593/
MPRA Paper No. 59593, posted 31 Oct 2014 09:34 UTC

Research in Business & Economics Journal
Volume 9, Article 8, 2014

An Econometric Analysis of the 2013 Major League Baseball Season

Steven L. Fullerton
Management Department
New Mexico State University
Las Cruces, NM 88003-8001
fuller5@nmsu.edu

Thomas M. Fullerton, Jr.
Department of Economics & Finance
University of Texas at El Paso
El Paso, TX 79968-0543
tomf@utep.edu

Adam G. Walke
Department of Economics & Finance
University of Texas at El Paso
El Paso, TX 79968-0543
agwalke@utep.edu

Abstract

An econometric analysis of 2013 Major League Baseball season is conducted with respect to regular season victories. Results obtained confirm many, but not all, results reported in prior research. The importance of solid team pitching, defense, and offense is underscored. Outcomes for salaries and league affiliations differ substantially from prior seasons. History may serve as a guide to what occurs on the field, but it does not always repeat in the manner indicated in bygone years. Ultimately, 2013 represents a departure from the standard baseball norm.

JEL Categories
M21, Business Economics; L20, Firm Behavior

Key Words
Major League Baseball; Team Performance

Acknowledgements

Partial funding support for this research was provided by El Paso Water Utilities, Hunt Communities, the City of El Paso, the Water Research Foundation, the UTEP Center for the Study of Western Hemispheric Trade, and a UTEP College of Business Administration Faculty Research Grant. Helpful comments were provided by Jim Peach. Statistical research assistance was provided by Alejandro Ceballos and Alan Jiménez.

An Econometric Analysis of the 2013 Major League Baseball Season

Introduction

Major league baseball, the so-called national pastime of the United States, is closely watched by millions of fans every year. It has also been the subject of episodic empirical research that uncovers interesting insights with respect to team performance. Evidence has been reported with respect to influence of both standard and non-standard variables on win-loss records, sometimes with unexpected results. Generally speaking, no two seasons are alike and team performances are not always affected in equal measure by different variables on a year-by-year basis.

This study conducts a quantitative analysis of major league baseball team wins for the 2013 season. The analysis is carried out using least squares regression. Because of wide ranging differences in aggregate team payrolls, parameter estimation utilizes the White (1980) heteroscedasticity consistent standard errors and covariance matrix.

Subsequent sections are as follows. Section two provides a brief overview of related studies. Section three discusses data and methodology. Section four summarizes empirical outcomes and implications. The paper closes with a concluding section.

Related Studies

On average, more productive players receive higher salaries than less productive and/or less experienced players. Depken (2000) investigates whether greater wage disparity on individual team rosters affects team performances. Results obtained indicate that increased payroll disparity is negatively correlated with the number of wins for teams. Essentially, wider salary dispersion is found to damage roster cohesiveness and hamper club abilities to compete effectively.

Molina and Jewell (2004) extend this approach by incorporating variables that measure team abilities at playing fundamentally sound baseball. Among the different variables added to the analysis are team earned run averages as a measure of pitching effectiveness, team errors per game as a measure of defensive futility, and team on-base percentage as a measure of offensive prowess. Results obtained confirm the salary dispersion hypothesis that greater payroll disparity hampers win loss records.

Those results are not exclusive to professional baseball in the United States. San and Jane (2008) report similar evidence for Taiwanese baseball. Namely, total team salaries are positively correlated with team performances, but wider payroll dispersions hurt win-loss records. The cooperative nature of team success has also been documented as one reason behind the validity of the payroll dispersion hypothesis.

Breunig et al. (2014) also confirm the salary dispersion hypothesis. Representative of accumulated human capital, as team average salaries increase, the numbers of victories also grow. Results in this study are based upon data through 2010 and further indicate, however, that a fair amount of uncertainty regarding overall outcomes tends to be present every season. That implies that individual seasons may not always follow the same patterns as prior seasons.

Data and Methodology

This study empirically examines the outcomes of the 2013 major league baseball season in the United States. Regular season team victories for the 162 game schedules are used as the dependent variable. Even though the last game of the season between Tampa Bay and Texas is officially classified as a regular season game, it was an extra game to determine which of those two teams with identical records would play in the one-game wild-card playoff game. Essentially, the extra game was also a playoff game and is excluded from the sample.

Variables collected for the analysis are listed in Table 1. The dependent variable, WINS, is the number of regular season wins by each team in 2013. ERA is team earned run average. Lower ERAs reflect better team pitching. Consequently, this variable should be negatively correlated with WINS. EPG is team errors per game. Lower EPGs represent better team defense. Given that, EPG is also hypothesized to be inversely correlated with WINS.

Table 1
Variable Names, Definitions, and Hypothesized Effect on Team Wins

Wins	Number of 2013 Total Team Wins	
ERA	Team Earned Run Average	Negative
EPG	Team Errors per Game	Negative
OPS	Team On-Base plus Slugging Percentage	Positive
TPR	Total Team Payroll in Millions of Dollars	Positive
PSD	Payroll Salary Standard Deviation in Millions of Dollars	Negative
DV	National League Dummy Variable, AL = 0, NL = 1	Negative

OPS measures team on-base percentage plus slugging percentage. Higher OPS numbers result from better team offense, so this variable should be positively correlated with WINS. TPR is total team payroll. Better players get bigger salaries, so this variable should be positively correlated with WINS. PSR is team player salary range. Bigger salary dispersions hurt team unity, so this variable should be negatively correlated with WINS.

The last variable is a dummy variable for the league affiliation of each team. American League teams use designated hitters and this gives them an offensive advantage in inter-league games (as well as in all-star games and world series games). DV is assigned a value of 0 for American League teams and a value of 1 for National League teams. Accordingly, this variable should be positively correlated with WINS.

As in previous studies (Breunig et al., 2014), the dependent variable is, for practical purposes, a continuous variable and least squares regression analysis is employed for parameter estimation. The sample is cross-sectional and exhibits substantial variability among the various teams that comprise major league baseball. Given that, heteroscedasticity testing will be utilized to determine whether ordinary least squares estimation will yield efficient parameter estimates and a consistent variance estimator (White, 1980).

Empirical Results

The computed White statistic for an ordinary least squares version of the equation is significant at the 1-percent level, rejecting the null hypothesis of homoscedasticity. Given that, heteroscedasticy corrected results are used for the t-statistics reported in Table 2. The outcomes support most of the hypotheses described above. As discussed below, however, there are three results for 2013 that run counter to what is hypothesized.

As expected, team ERA is critically important to win-loss records. The estimated parameter in Table 2 indicates that every one run increment in team ERA above the major league average, all else equal, reduces its WINS by 19 games under the conditions observed in 2013. Given this, the global search for major league quality pitching arms is completely understandable.

Team defense is similarly important. The coefficient for team EPG in Table 2 indicates that every additional error per game above the major league average reduces WINS by nearly 14 games under the circumstances observed during 2013. While glove work and accurate throws are not regarded with the same adulation as power hitting, it is no wonder that nightly news reels are filled with defensive gem videos.

Table 2
Heteroscedasticity Adjusted Output for 2013 Major League Wins
Dependent Variable: Wins
Method: Least Squares with White Heteroscedasticity-Consistent Standard Errors & Covariance
Sample: 1 30
Included Observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Probability
Constant	8.467	19.816	0.427	0.673
ERA	-19.253	1.583	-12.162	0.000
EPG	-13.783	6.160	-2.238	0.036
OPS	2.165	0.279	7.755	0.000
TPR	-0.069	0.032	-2.193	0.039
PSD	1.578	0.630	2.503	0.020
DV	-0.061	1.626	-0.038	0.970

R-squared	0.937	Dependent Variable Mean	81.000
Adjusted R-squared	0.920	Dep. Var. Std. Deviation	12.418
Std. Error of Regression	3.513	Sum of Squared Residuals	271.446
F-Statistic	54.653	Probability (F-Statistic)	0.000
Wald F-Statistic	83.105	Probability (Wald F-Statistic)	0.000
Log Likelihood	-73.578		

To win games, teams must score runs. To do so requires getting on base and advancing to home. Team OPS is measured in percentage points. For every one percentage point that team OPS above the major league average, WINS increase by approximately 2 games.

The uniqueness of the 2013 season becomes apparent once the effects of the off-the-field variables are quantified. Surprisingly, the TPR coefficient is both negative and statistically significant. Among other things, that outcome (from a high-end payroll perspective) undoubtedly reflects the collapse of the 2012 World Series Champion San Francisco Giants, plus the sub-par season of the Los Angeles Angels of Anaheim, and the third place finish of the New York Yankees in the American League East Division. From a low-end payroll perspective, the negative TPR coefficient also results, at least partially, from the 90-plus win seasons tallied by the Oakland Athletics, the Atlanta Braves, the Cleveland Indians, and, especially, the Tampa Bay Rays. Along the same lines, the 86 victories achieved by the Kansas City Royals eclipsed the 85

wins posted by the Yankees. While the sign of this coefficient is unexpected, it is similar to the finding reported by Keener (2013).

Further evidence of the 2013 departure from the norm is provided by the PSD coefficient. Historically, greater salary disparities are associated with sub-optimal team performances. The estimated parameter for PSD is greater than zero, indicating that the opposite was observed in 2013.

The final regression coefficient appearing in Table 2 is estimated for the qualitative variable defined for league affiliation. As defined, 0 = American League, 1 = National League, the parameter is hypothesized to be less than zero. The coefficient for DV is negative, but it does not satisfy the 5-percent criterion and its magnitude is so small that the traditional American League offensive advantage seems to have been largely negated in 2013.

Because the results obtained for the off-the-field variables representing total human capital, payroll structure, and league rule differences, another specification is tested. In that specification, only the on-the-field performance variables are included. Once again reflecting the departure of 2013 from the past, parameter magnitudes in Table 3 are almost the same as those for the corresponding slope coefficients in Table 2.

Table 3
Reduced Specification Heteroscedasticity Adjusted Output for 2013 Major League Wins
Dependent Variable: Wins
Method: Least Squares with White Heteroscedasticity-Consistent Standard Errors & Covariance
Sample: 1 30
Included Observations: 30

Variable	Coefficient	Std. Error	t-Statistic	Probability
Constant	18.671	19.269	0.969	0.342
ERA	-18.756	1.194	-15.713	0.000
EPG	-17.084	5.067	-3.371	0.002
OPS	2.019	0.257	7.869	0.000

R-squared	0.924	Dependent Variable Mean	81.000
Adjusted R-squared	0.915	Dep. Var. Std. Deviation	12.418
Std. Error of Regression	3.623	Sum of Squared Residuals	328.223
F-Statistic	101.283	Probability (F-Statistic)	0.000
Wald F-Statistic	125.419	Probability (Wald F-Statistic)	0.000
Log Likelihood	-76.323		

It would be tempting to attribute the anomalies of 2013 to injuries. The Yankees, for example, lost all-star players Curtis Granderson and Mark Texeira for extended periods. Although not shown here, additional equations containing team disabled list days as an explanatory variable were also estimated. The results were inconclusive and the 2013 team injury coefficient was consistently insignificant. The magnitudes of the other parameters do not change when the team injury variable is included.

The results discussed above highlight the uncertainties associated with major league baseball and, probably, all professional sports. The fundamentals of good pitching, good defense, and good offense are critically important to major league baseball team performances.

While off-the-field factors such as salary disparities and rules differences usually tend to affect team standings, there is no guarantee that will always be the case for every season.

Conclusion

An empirical analysis of the 2013 major league baseball season illustrates how much the most recent season departed from the historical norm. Estimation results confirm the ongoing importance of pitching, fielding, and hitting. Surprisingly, however, total player capital, as approximated by team payrolls, was not found to improve team win-loss performances. Similarly, salary disparities were not found to hurt team records as has tended to occur in the past.

During the free-agent era, the teams with the highest payrolls have won the most games and championships. They have not, however, won all of the games and championships. Seasons like 2013 when low salary teams like Oakland and Tampa Bay out-perform their high payroll counterparts may be exceptional, but they cannot be ruled out.

Will 2014 resemble 2013 or will it represent a return to the historical norms of yesteryear? Prior empirical evidence points to the latter. If that occurs, the results discussed above will not be replicated. The safe bet is that 2013 was anomalous, but baseball is a professional sport in which serious departures can, and do, periodically occur.

References

Robert Breunig, Bronwyn Garrett-Rhumba, Mathieu Jardin, and Yvon Rocaboy, 2014, "Wage Dispersion and Team Performance: A Theoretical Model and Evidence from Baseball," *Applied Economics* 46, 271-281.

Craig A. Depken, 2000, "Wage Disparity and Team Productivity: Evidence from Major League Baseball," *Economics Letters* 67, 87-92.

R. Todd Jewell and David J. Molina, 2004, "Productive Efficiency and Salary Distribution: The Case of US Major League Baseball," *Scottish Journal of Political Economy* 51, 127-142.

Mary Hilston Keener, 2013, "The Econometrics of Baseball: A Statistical Investigation," *Research in Business and Economics Journal* 9, Article 2. Retrieved from http:// http://www.aabri.com/rbej.html

Gee San and Wen-Jhan Jane, 2008, "Wage Dispersion and Team Performance: Evidence from the Small Size Professional Baseball League in Taiwan," *Applied Economics Letters* 15, 883-886.

Halbert White, 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica* 48, 817–838.