# NAME OF THE GAME:
# DO STATISTICS CONFIRM THE LABELS OF PROFESSIONAL BASEBALL ERAS?

by

Mitchell T. Woltring

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Leisure and Sport Management

Middle Tennessee State University
May 2013

Thesis Committee:

Dr. Colby Jubenville

Dr. Steven Estes

ACKNOWLEDGEMENTS

I would not be where I am if not for support I have received from many important people. First and foremost, I would like thank my wife, Sarah Woltring, for believing in me and supporting me in an incalculable manner. I would like to thank my parents, Tom and Julie Woltring, for always supporting and encouraging me to make myself a better person. I would be remiss to not personally thank Dr. Colby Jubenville and the entire Department at Middle Tennessee State University. Without Dr. Jubenville convincing me that MTSU was the place where I needed to come in order to thrive, I would not be in the position I am now. Furthermore, thank you to Dr. Elroy Sullivan for helping me run and understand the statistical analyses. Without your help I would not have been able to undertake the study at hand. Last, but certainly not least, thank you to all my family and friends, which are far too many to name. You have all helped shape me into the person I am and have played an integral role in my life.

# ABSTRACT

A game defined and measured by hitting and pitching performances, baseball exists as the most statistical of all sports (Albert, 2003, p. ix). Probably more than any other sport, the game's present is couched in references to its history. Professional baseball has endured many changes (both overt and subtle) in rules, equipment, stadium structures, and competitive strategy over the course of its history. Because of such shifts, the modern era of Major League Baseball (MLB) has been segmented into six distinct eras (Lombardi, 2006): Dead Ball (1901-1919), Live Ball (1920-1941), Integration (1942-1960), Expansion (1961-1976), Free Agency (1977-1993), and Long Ball/Steroid (1994-2005). This study runs through the 2011 season and adds a seventh era, labeled "Post-Steroid" (2006-present).

The purpose of this research was to determine how the names and/or characteristics/perceptions associated with the actual offensive outputs of each era of MLB corresponded with the statistical realities related to each era's On-Base Plus Slugging Percentage (OPS), beginning with the 1901 season and MLB's Modern Era. The study's sole focus was the effect of team OPS to determine how hitting and pitching contributed to team winning percentage in each era.

Results were segmented by each defined era to determine any significant differences between the eras. Multiple regression and ANOVA were used to determine if perceptions and realities for each era's offensive output aligned descriptively. Results showed that perceptions for five of the seven eras matched statistical realities, while

perceptions of two eras did not. Results also showed significant statistical differences between the defined periods and illustrated how offensive output defined each era.

TABLE OF CONTENTS

Page

LIST OF TABLES

# CHAPTER I

## INTRODUCTION

As American historian Jacques Barzun (1954) wrote, “Whoever wants to know the heart and mind of America had better learn baseball, the rules and realities of the game” (p.159). This quote still rings true today. However, if someone wants to understand America and the game baseball in the 21st Century, they had better learn to follow the money. Zimbalist (1992) stated, “Money, was already a significant part of the game when the first professional team, the Cincinnati Red Stockings, was formed in 1869” (p.2). He (1992) adds that, “Although professional baseball’s institutions began to take clearer shape in 1876 and more solid form after 1903, the business of baseball has really has been in steady flux since its inception” (p.2).

Simply put, the game of baseball, at the professional level, is no longer a past time or tax write-off for owners; it is a primary source of business. Over the course of baseball, the economics of the game have dictated that management practices change by doing everything better, explaining the current trend in the game towards using empirical data with sophisticated statistical analysis to aid in decision making.

One could spend his or her entire lives dedicated to learning the game of baseball and merely scratch the surface of knowledge. At its core, the game of baseball has two very separate, yet continuously linked facets: hitting and pitching. Former pitcher Bob Veale (1966) concisely summed up over a century of baseball wisdom when he famously said, “Good pitching will beat good hitting any time, and vice versa.” On any given day,

it is impossible to know whether pitching or hitting will win a game of baseball. But thegame of baseball has provided innumerable statistics that can help to clarify whether hitting or pitching are more important.

Baseball is the most statistical of all sports, and the game is defined by hitting and pitching (Albert, 2003, p. ix). More than any other sport, the history of the game is referenced when talking about anything that is happening in the present. As author and baseball historian Stanley Cohen (1988) declared:

> Baseball, almost alone among our sports, traffics unashamedly and gloriously in nostalgia, for only baseball understands time and treats it with respect. The history of other sports seems to begin anew with each generation, but baseball, that wondrous myth of twentieth century America, gets passed on like an inheritance. (p.70)

Therefore, to understand baseball, one must understand both the history of the game and the economic drivers that lead to changes in the game on the field.

As Durant (1973) noted, the inception of the American League in 1901 as a rival to the already entrenched National League, “marked the beginning of the modern administration of baseball” (p. 47). The work of Adelman (1986) and Guttman (1978) show that the modernization of sport occurred long before what Durant notes as the modern administration of baseball, specifically between 1820 and 1870. Although the sport itself may have been modern prior to 1901, baseball was under rapid and constant change because the game was still working to get itself properly organized and rules were

changing from year to year. This meant that the game could be drastically different from year to year (Smith, 2010). Both the availability and the reliability of statistics prior to the 1901 season make it difficult to include them as part of the present research. The relative stability of rules and record keeping since 1901 makes it much more appealing for an empirical study. For the purposes of this study, the researcher focused on the modern era of baseball, beginning with the 1901 season.

In addition to properly organizing the game and setting a uniform set of the rules, the modern era brought forth the creation of the World Series which would pit the winners of each league against each other for a yearly championship. Structure and consistency put baseball at the forefront of American sports and baseball prospered as the modern era began. As Rader (1992) said, "During the first two decades of the twentieth century, professional baseball achieved a new level of maturity and stability as an American institution" (p.98).

Baseball has endured much change over the course of its history, and because of constant change, the modern era of baseball has been segmented into six distinct eras. Steve Lombardi (2006), a baseball analyst currently with Baseball-Reference, described the eras as the Dead Ball Era (1901-1919), the Live Ball Era (1920-1941), the Integration Era (1942-1960), the Expansion Era (1961-1976), the Free Agency Era (1977-1993) and the Long Ball/Steroid Era (1994-2005). This study runs through the 2011 season and a seventh era will be added and labeled the Post Steroid Era (2006-present). The rationale for the Post-Steroid Era will be explained in Chapter II.

## Delimitations

### *Scope of the study*

This examination of OPS in Major League Baseball began with the 1901 season, which marks the beginning of the Modern Era in baseball. It's focus is on the effect of team On-Base Plus Slugging Percentage on team winning percentage. The results were segmented into the defined eras for comparison in order to determine significant differences between each era.

### *Selection of the case*

The case for the study was any team that has been a member of Major League Baseball since the beginning of the 1901 season. Any franchise that has operated since then has played under the same set of rules as every other team in Major League Baseball in this timeframe. Availability of the relevant statistics, a prevalence of baseball-themed research, and the researcher's personal interest in baseball were vital to the selection of this case.

### *Definition of terms*

*Dead Ball Era* – The first segmented era of the Modern Era, spanning 1901-1919. It was characterized by low-scoring games, very few power hitters, and baseballs that would be used as long as possible, rendering them soft, mushy, and "dead." The 1908 season featured the lowest run average in Major League Baseball history, with teams combining to score only 3.4 runs per game. It was much more of a strategy-driven game because of the scarcity of runs scored. Bunting, stealing

bases, and hit-and-runs were very common, and to this day are referred to as playing "small ball." Teams tended to play in spacious parks that made hitting home runs with the dead ball even tougher, necessitating the reliance on small ball. (Rader, 1992; Thorn, 1974).

*Expansion Era* – The fourth segmented era of the Modern Era, spanning 1961-1976. It was so named for the unprecedented expansion in number of teams and geographical location of teams. The movement began with the Dodgers and Giants moving to California in the late 1950's and saw expansion across the United States and even Canada. Improvements in transportation, specifically air travel, aided baseball in its quest for expansion across the country. Baseball truly became America's pastime as people all over could now see Major League. Baseball (James, 2001; Neft et al., 1982; Rader, 1992).

*Free Agency Era* – The fifth segmented era of the Modern Era, spanning 1977-1993. It was so named because of the inception of free agency prior to the 1977 season. Free agency brought about fundamental changes to the structure of Major League Baseball that were unequaled since the inception of the Modern Era in 1901. Teams could no longer treat players as cheap commodities as player salaries immediately skyrocketed because of teams having to bid for their services. (Koppett, 2004; Neft et al., 1982; Rader, 1992).

*Integration Era* – The third segmented era of the Modern Era, spanning 1942-1960. It was named for the integration of black players into Major League Baseball starting with Jackie Robinson in 1947. The beginning of the era saw many

baseball players leave for World War II, which in many cases diminished the talent level in Major League Baseball for the early part of the era. The final team to integrate and play a black player in a Major League Baseball game would be the Boston Red Sox in 1959. (Bedingfield, 2009; Rader, 1992).

*Live Ball Era* – The second segmented era of the Modern Era, spanning 1920-1941. It was so named because of rule changes that made sure clean balls were being used throughout the game, and later. In the mid 1920's, the standard baseball was also changed in order to make it more "lively", meaning it flew further off the bat. Hitting took center stage with the emergence of players such as Babe Ruth. In this era, home run totals increased to numbers never previously witnessed. The era changed the life of the pitcher forever. From 1910-1920, eight pitchers recorded 30 win seasons; since 1920 only three pitchers have accomplished that feat, and the last was Denny McLain in 1968. (Honig, 1976; Rabinowitz, 1989; Rader, 1992).

*Major League Baseball* - Major League Baseball is the most well-known professional baseball league in the world. Major League Baseball uses 1869 as its founding year as that was the first year that a current organization in MLB, the Cincinnati Red Stockings (now Reds), began play. It serves as the governing body for both the American League and National League and is responsible for all the statistics that have been kept since 1901. (Neft, Cohen & Deutsch, 1982; Rader, 1992; Thorn, 1974).

*Modern Era* – Represents Major League Baseball from 1901 to the present. It includes each of the seven defined eras that are examined in this study. The modern era features an accurate record of statistics, a stable environment with both the National and American Leagues working congruently, a consistent numbers of teams (meaning no teams folding during a season), and a consistent set of rules across both leagues. (Smith, 2010).

*On-Base Percentage (OBP)* – It represents the ability to get on base. Specifically, it is the measure of how often a batter, or a team, reaches base for any reason other than a fielder's choice, a fielding error, a dropped third strike, fielder's obstruction, or catcher's interference (Albert, 2010, p.2). A perfect OBP would be 1.000. The Major League average for OBP during the modern era is .340. Generally, an OBP of .350 or higher is considered good.

*On-Base Percentage Against (OBPa)* – It represents the ability to keep runners off of the bases. It is calculated exactly the same as OBP, but against a pitcher, or pitching staff (Albert, 2010, p.2). What would be considered good would be an inversion of OBP since the goal is to keep runners off base.

*On-Base plus Slugging Percentage (OPS)* – It represents the ability to both get on base and hit for power. It is a result of the simple addition of On-Base Percentage and Slugging Percentage (Albert, 2010, p.3). Baseball historian and Sabermetrician Bill James (2009) built a seven category Likert Scale for hitters based on their OPS. A "great" hitter will have an OPS of .9000 or higher, a "good" hitter will fall between .8333 and .8999, an "above average" hitter will fall between .7667

and .8332, an “average” hitter will fall between .7000 and .7666, a “below average” hitter will fall between .6334 and .7665, a “poor” hitter will fall between .5667 and .6333, and an “atrocious” hitter will fall below .5667 (p.24). The OPS for the entire Modern Era is .714, which makes sense in that all hitters combined fall into James’ “average” category. What is typically considered a good season for OBP (.350) and SLG (.500) can be added up to obtain a good measure for OPS. It adds up to an .850 OPS, which makes sense as it would fall into James’ “good” category.

*On-Base plus Slugging Percentage Against (OPSa)* - On-Base plus Slugging against uses the same calculation as OPS, except that it is it is a measure of a pitcher’s ability to both keep batters off base and prevent them from hitting for power (Albert, 2010, p.3). What is considered good is an inversion of what it would be for hitters, as the pitcher’s job is to limit base runners and power. The value of a hitter can be defined by their OPS, just as the value of a pitcher can be defined by their OPSa.

*Post-Steroid Era* – The seventh segmented era of the Modern Era, spanning 2006 to the present. It was so named for the fundamental changes to the banned substance policies of Major League Baseball. Beginning with the 2006 season, players who tested positive for a banned substance were subjected to incremental punishments of 50-game suspensions, 100-game suspensions, and lifetime bans from Major League Baseball for a third positive test. These changes were made in response to the outrage of baseball fans, writers and historians of the events that transpired

during the Steroid Era. Offensive numbers have returned to pre-Steroid Era levels and pitching numbers have improved dramatically. (Chen, 2010; Dittmeier, 2012; Ratto, 2012).

*Sabermetrics* - It is the specialized analysis of baseball through objective evidence, especially baseball statistics that measure in-game activity. The term is derived from the acronym SABR, which stands for the Society for American Baseball Research. It was coined by Bill James, who is one of its pioneers and is often considered its most prominent advocate and public face. Grabiner (1994) begins his *Sabermetric Manifesto* by saying, "Bill James defined sabermetrics as 'the search for objective knowledge about baseball.' Thus, sabermetrics attempts to answer objective questions about baseball…It cannot deal with the subjective judgments which are also important to the game" (p.1). Sabermetricians frequently question traditional measures of baseball skill and, therefore, question traditional views of baseball, including its history.

*Slugging Percentage (SLG)* – It represents the ability to hit for power. Specifically, it is the measure of the average number of bases a batter will get per at-bat (Albert, 2010, p.2). A perfect slugging percentage would be 4.00, and would mean that a player hit a homerun in every at-bat of their career. The single-season record for SLG is held by Barry Bonds at .863 in 2001. A slugging percentage over .500 is generally considered a good season for a power hitter.

*Slugging Percentage Against (SLGa)* – It represents a the ability to limit a batter's power numbers. It is calculated exactly the same as SLG, but against a pitcher or

pitching staff (Albert, 2010, p.2). What would be considered good would be the inversion of SLG since the goal is to limit power.

*Steroid Era* – The sixth segmented era of the Modern Era, spanning 1994-2005. It was so named because of the rampant use of steroids and other substances deemed by MLB as performance enhancing drugs (PEDs) during the era. Major League Baseball had extremely lax rules in place for both testing of banned substances and consequences of being caught. Offensive numbers saw huge jumps, especially in home runs as hitters were larger and stronger than ever before and swinging for the fences more than ever before. Following the 1994 strike that canceled the World Series, Major League Baseball did not truly see a recovery in its fan base until the (allegedly) steroid-spurred Home Run Race of 1998 between Mark McGwire and Sammy Sosa. The previous record for home runs in a season of 61, set by Roger Maris in 1961, would be surpassed six times between 1998 and 2001. (Grossman, Kimsey, Moreen & Owings, 2007; Hill & Schvaneveldt, 2011; Lenhardt, 2010).

**Statement of the Problem**

Each of the seven eras that comprise the modern era of Major League Baseball is defined by how it is perceived that games were won during the particular era. Hitting and pitching statistics were analyzed because, as Koop (2002) said, "data limitations make it difficult to model defensive performance" (p. 711). There have been great advancements in metrics to determine defensive performance, but they still face great criticism as to their reliability. Focusing on hitting and pitching is supported by Miceli & Huber (2009)

who found that pitching explains 2/3 of the variance in winning percentage and hitting explains the other 1/3, with no mention of defense (p.1). The problem, then, is to determine, definitively with statistics, whether hitting or pitching contributed more to winning percentage in a given era. Those definitive answers could then be compared to the perceptions of baseball historians for each of the given eras and expand the body of knowledge of baseball history.

The statistics can be run and analyzed because they represent an objective view of what happened. Determining perceptions is not a much more subjective process. An extensive overview of each era is necessary in order to understand the similarities and differences of the perceptions that make each era unique. By understanding each distinctive era, it is possible to compare and contrast the perceptions of each era to the reality that is presented by the statistics that were recorded.

**Case Study**

During the investigation, it became apparent that the best framework to build this study around was as a case study. A case study provides a systematic way of looking at events, collecting data, analyzing information and reporting the results. A case study fits well because of the nature of the information being analyzed, as it is an in-depth, longitudinal study examination of a single phenomenon: the effect of OPS on winning percentage. This framework allows a researcher to gain a sharper understanding of how the perceptions and statistical realities relate (Flyvbjerg, 2006, p.229).

Since this study tested a hypothesis, a case study is not always the first framework that would come to mind because of some misconceptions about the methodology. Some suggest that hypothesis testing is not the correct way to use a case study. But Flyvbjerg (2006) argued that case studies are often misunderstood because there is a perception that they are most useful in generating hypotheses in the first step of a research process, whereas hypothesis testing and theory building should be carried out by other processes (p.229). Eckstein's (1975) findings build on Flyvbjerg's rationale by saying that case studies, "are valuable at all stages of the theory-building process, but most valuable at that stage of theory-building where least value is generally attached to them: the stage at which candidate theories are tested" (p. 80).

It is important when conducting a case study to understand and clearly identify both the object and the subject of the study. Wieviorcka (1992) described the subject as the "practical, historical unity" through which the theoretical focus of the study is being viewed. It is the lens in which the researcher focuses through. The object is the "theoretical focus or analytical frame" (p. 160). In this study, the subject was represented by how perceptions and statistical reality that was recorded lined up by the eras. The object was the effect of OPS on winning percentage through the lens of the distinct eras.

The present study represented an evaluative approach to a case study. Thomas (2011) defined an evaluative case study as one where the purposes are first identified, then approaches are delineated, then processes are decided upon, and finally results are interpreted (p.512). The present study followed that approach as it began with the purpose, which was to determine if the perceptions of the eras match up with the

statistical realities. Next an approach was chosen, in this case testing the theory that each of the given eras do not match on subjective and objective levels. The processes that were decided upon will be outlined in the methodology section. And finally, the results will be interpreted, as they will be in the results section.

Using a case study framework also made it possible to incorporate both quantitative and qualitative data. In order to determine perceptions of each era, the researcher had to rely on qualitative accounts of the eras through baseball writers and historians. They shape what people believe and conceive about baseball history and are at the forefront of building the perceptions many baseball fans grow up with. But in order to objectively test each era the researcher must use quantitative data, in the form of statistics and their effect on winning percentage. Thus, a mixed method incorporating quantitative and qualitative data is necessary to complete the purpose of the study. A case study allows the researcher to do both.

**Importance of the Study**

Since the Cincinnati Red Stockings became the first professional baseball team, the game has been a business. Owners of teams have the primary goal of making a profit. Major league manager Terry Francona (2013) put it bluntly when describing the ownership group of the Boston Red Sox, “I don’t think they love baseball. I think they like baseball. It’s revenue, and I know that’s their right and their interest because they’re owners” (p.54). The core business of sport will always be entertainment on the field, because ultimately it is what the fan is paying to see. Putting together a winning team will draw more fans, both to the stadium and on television, which will increase revenue.

Therefore, strategies that can be implemented to increase winning percentage while lowering costs will always entice baseball management.

There are many studies that have looked at variables that correlate to winning percentage in Major League Baseball (Pujol & Nix, 1994). These variables can be divided into two groupings, the first of which contain variables that do not directly affect the product on the field. The second grouping consists of statistics that are a direct product of what is happening on the field of play. Each is important in explaining winning percentage, but they should be viewed differently in relation to their effect on winning percentage.

Variables that do not directly affect the product on the field have been analyzed to help determine their relationship to winning percentage in Major League Baseball. Specific studies have been done comparing the relationships of market size (Burger, J.D. & Walters, S.J.K, 2003; Butler, 1995; Schmidt & Berri, 2002), payroll (Hall, S., Syzmanski, S., & Zimbalist, A.S., 2002; Mizak & Stair, 2004; Wiseman & Chatterjee, 2003), and attendance (Davis, 2008; Davis, 2009; Schmidt & Berri, 2001) with winning percentage in Major League Baseball. While these variables have merit, they do not fall in line with the purpose of this study, which was to explain how the statistics derived from on-field play influence winning percentage.

Specific studies looking at the relationship between offensive and pitching statistics and winning percentage in Major League Baseball have also been undertaken. These studies fall in line with the focus of this study as they help explain what is happening on the field that affects winning percentage. Previous studies have focused on

offensive statistics such as: stolen bases (Baumer, 2009; Baumer & Terlecky, 2010; Demmink, 2010; Turocy, 2005), batting average (Albert 2002; Bennett & Fluek, 1983; Houser, 2005), on-base percentage and slugging percentage as separate statistics (Deli, 2012; Farrar & Bruggink, 2011; Hakes & Sauer, 2006; Houser, 2005), and on-base plus slugging percentage (Hakes & Sauer, 2006; Lopez, Mundfrom, & Schaffer, 2011). Studies have also focused on pitching statistics such as earned-run average (Lackritz, 1990; Lopez, Mundfrom, & Schaffer, 2011; Olson, 2001), strikeouts (Chapman & Southwick, 1991; Houser, 2005; Sommers & Quenton, 1982) and walks plus hits divided by innings pitched, or WHIP (Beneventano, Berger & Weinberg, 2012; Houser, 2005;).

There are many other offensive and pitching statistics that have been covered in addition to the ones mentioned, but the point is that, as a whole, the analyses of offensive and pitching statistics have found that one statistic is better than any other at predicting a Major League Baseball team's winning percentage: On-Base Plus Slugging Percentage, or OPS. Besides being the best predictor of team success, OPS provides the ability to be used as both as offensive and pitching statistic, as it can be calculated for hitters or against pitchers. This allows the measurement of hitting and pitching on common ground, rather than trying to compare two statistics that are calculated by different means. Many of the aforementioned studies have looked at specific sets of times, or over the modern era of baseball as a whole. Despite all the relevant research, there seems to be a lack of studies that look at each of the eras of baseball independently.

Sabermetrics has provided a tool to empirically analyze baseball data and discern the empirical nature of what occurred. By studying what occurred and why, statistically

informed predictions about the future can be made. Major League Baseball is rich in tradition and history with clearly defined eras. With such clear distinctions, it makes sense to test each of these eras against perceptions. In order to delineate each era, it is thus worthwhile to determine the nature of each era.

**Statement of the Purpose**

The purpose of the present study was to use the statistics of OPS and OPS Against (OPSa) to determine if the perceptions of each era in Major League Baseball history align with the statistical reality. OPS represents a measure of hitting, while OPSa represents a measure of pitching. By using OPS and OPSa, it could be determined whether hitting or pitching contributed more heavily to winning percentage in each era. Additionally, the present study discovered whether there were significant differences in the importance of hitting and pitching between the eras as it related to each of the other eras.

**Hypothesis**

The success of Sabermetrics in disproving some traditional measures of baseball skill has cast doubts about how baseball has been viewed. Included in these doubts should be how each era is viewed; specifically whether hitting or pitching was more important to winning percentage in each given era. This study set out to demonstrate that OPS and OPSa will show that the subjective perceptions of each era do not match up with the objective, statistical reality

# CHAPTER II

## REVIEW OF THE LITERATURE

In order to properly assess the perceptions of each era, the subjective interpretations of each era must be relied upon. By compiling a multitude of opinions on each era, a wider, more general view of each era could be drawn. There are many scholarly articles to use in defense of the use of OPS, but, as noted earlier, there is an apparent lack of articles dedicated to the specific eras of Major League Baseball. For this reason, many accounts from baseball writers and historians were used to determine the perceptions of each era. As they were the people conveying opinions about baseball to the fans, it is imperative to use their writings as they truly are the people who steered the perceptions about each era.

### The Dead Ball Era (1901-1919)

As the modern era of baseball began, two specific rule changes certainly favored pitchers. The first rule change addressed the size and width of home plate. It was changed from a 12 inch square base to a five-sided figure that measured 17 inches across, making the strike zone much larger (Rader, 1992, p.87). The second rule change addressed foul balls. Prior to 1901 foul balls were not counted as strikes, so batters could foul off as many pitches as they wanted with no consequence. Beginning in 1901, foul balls began counting as strikes (Rader, 1992, p.87). As Rader (1992) noted, “with the larger plate and the new foul ball strike rule, strikeouts jumped more than 50 percent, while batting averages, home runs, slugging percentages, and runs per game sank to all-

time lows" (p.87). Because of factors including these rule changes, Neft et al. (1982) argued, "The game belonged to the pitcher. The parks were large and lacked the enclosures which could invite the long ball" (p. 11).

As if the hitters were not at enough of a disadvantage, Rader (1992) added that, "Hitters complained about the use of soft, discolored balls, leading baseball historians to label the age as the 'dead ball era' (pp.87-88). Thorn (1974) contended that, "the ball was so dead that hardly anyone could slug it into the stands" (p. 27). This argument is backed by the fact that the St. Louis Cardinals hit a major league leading 39 home runs, as a team, in 1901. By way of comparison, the recently concluded 2012 season saw six individual players hit more than 39 home runs. The works of baseball historians and writers embody a representation of a game dominated by pitchers in the Dead Ball Era.

**The Live Ball Era (1920-1941)**

The game of baseball witnessed dramatic changes during the Live Ball Era. Honig (1976) described the series of changes that baseball saw during the period by stating that, "Technological, economic and social upheavals…dictated a changing world, and nothing, including baseball, remained unaffected" (p.19). There were many internal and external factors that led to the array of changes, but the focus of the majority of changes was quite clear. Rader (1992) stated, "Beginning with the AL's 1920 season, the hitters went on a rampage that continued through the 1941 season" (p. 112). Pitcher Lefty Grove, a Hall of Famer and one of the greatest left-handed pitchers of all-time, described the Live Ball Era to Honig (1975), "The league was chock-full of hitters…those days if you didn't hit .300, they didn't think much of you" (p.81).

This change in the on-field production was the direct effect of more rule changes. "Trick" pitches, such as the spitball and the emery ball were abolished and an extreme effort was made to keep clean balls in play that the hitters could now clearly see (Neft et al., 1982, p.123). The result was an immediate increase in home runs, as Babe Ruth alone hit 54 in 1920. Ruth had led baseball with 29 home runs in 1919, and only 5 other players had even cracked double-digits that year. Neft et al. (1982) put it quite succinctly:

> The advent of power hitting did more than alter the face of the game. With its coming was the imbalance that is inevitable with such a drastic change. Those who paid the heaviest price were baseball's minority breed, the pitcher. Once a feared and revered figure, he became the object of many a cannon blast. (p.123)

Pitchers complained constantly about the rule changes, but the wheels had clearly been set in motion to not only keep, but add to the changes that steered the game in favor of the hitters (Rader, 1992, p.116).

Only five years after baseball first changed to a "livened" ball, it went a step further to increase the offensive output. Rabinowitz (1989) said that the change to an even livelier baseball "was based on the theory that fans prefer home runs to pitching duels" (p.54). Baseball historian Bill James (2001) echoed this sentiment by stating:

> When the owners discovered that fans *liked* to see home runs, and when the foundations of the game were simultaneously imperiled by disgrace, then there was no turning back. In 1925 a new 'cushioned cork center' ball was introduced,

> perhaps more lively than those before it, and offense was allowed to dominate (p.122).

Immediately following this change, Babe Ruth would hit 60 home runs in the 1927 season to establish a record that would stand until 1961. It was an important revelation to the owners that the fans enjoyed an offensive game, because external factors would soon affect baseball as much as any rule changes.

The Great Depression struck America with the crash of the stock market in 1929. Baseball was not unaffected as Rabinowitz (1989) stated, "Major league baseball, like virtually every industry, severely felt the effects of the Depression, though not immediately" (p.49). He further detailed that once the Depression hit baseball hard in 1931, the owners cut operating expenses by cutting player salaries, reducing the active roster, cutting coaching staffs and team personnel and adopting a truly standardized ball in 1934 (p.54). In order to keep fans coming to games during the most economically trying period in American history, Major League Baseball continued to tweak the game in order to maximize offensive output.

Offensive records, both on the individual and team levels, were continually shattered during the era. The internal and external factors associated with the game during the Live Ball Era all contributed to a burgeoning offensive game. A review of the era clearly shows that baseball historians agree that hitting dominated pitching in the Live Ball Era.

**The Integration Era (1942-1960)**

The Integration Era is aptly named because on April 11, 1947 the Brooklyn Dodgers signed Jackie Robinson, making him the first black player in Modern Major League Baseball (Effrat, 1947, April 11). On April 15, 1947 Jackie Robinson would play in his first game as he started at second base for the Brooklyn Dodgers. It would take 12 years, in 1959, until the Boston Red Sox would become the last team in baseball to integrate and play a black player (Rader, 1992, p.152). It may have taken 12 years for each team to integrate, but the game certainly saw tremendous change during the Integration Era.

It is important to note that the Integration Era begins in 1942, which is five years before baseball actually integrated black players. The reason for marking the beginning of the era in 1942 is because a major external change occurred in the 1942 season that facilitated in setting it apart from the Live Ball Era. As the United States entered World War II and young men were drafted into service, Major League Baseball was not unaffected. According to Gary Bedingfield (2009) at total of 1,363 players, managers, coaches and umpires from the major and minor leagues served in World War II, including 29 Hall of Famers. This meant that replacement players at all levels had to fill in the rosters and the quality of play was arguably diluted as a result. The war took its toll on baseball as players were drafted into active service as early as 1940 (Neft et al., 1982, p.208).

Offensive averages took a slight downturn as compared to the live ball era, but due to its popularity with fans and changes by the players, home runs continued to rise

(Neft et al., 1982, p.265). A slight drop in batting average was not enough to turn the tides back in favor of the pitchers. The perception of the era is abstracted well by Neft et al. (1982) with:

> The pitcher's plight, as it had begun in 1920, did not improve. Instead, it only got worse as light bats became the mainstay of the batter's arsenal. Armed with a light weapon that responded to a good pair of wrists that could help negate the blazing fastball, the batter was able to continue his dominance over the pitcher. The result was that choking up on the bat became a rarity and nearly everyone went for the pump, or home run (p.265).

The batting averages may have lowered, but home runs and integration continued to bring fans coming out to the games. Rader (1992) reinforced this claim, "Thanks largely to (Jackie) Robinson, five National League teams set new season attendance records in 1947" (p.151). .

Technological advances during this period elicited changes to the game as well. The first night game was played in 1935, but it was mostly viewed as a novelty prior to the Integration Era. Honig (1976) said that, "night baseball, infrequent before the war, came to dominate the playing schedule, its popularity unquestioned". He added that many ballplayers of the era felt the record books should have started anew with the advent of night baseball, as it favored the pitcher (p.19). This was clearly a factor in the slight downturn in offensive numbers in comparison to the Live Ball Era.

Baseball had presumably found the recipe to keep fans coming in droves: let black players in the majors, keep belting home runs, and continue allowing hitting to dominate pitching

**The Expansion Era (1961-1976)**

The expansion era is appropriately named because the game expanded in numbers of teams, locations of teams and types of players. As Neft et al. (1982) said, "Until 1960, baseball had managed to preserve its limited franchise sanctuary of 16 teams since 1901" (p.345). Prior to 1957 when the Dodgers moved to Los Angeles and the Giants to San Francisco, they had also managed to keep all their teams in a geographic area that went no further south than Washington, D.C. and no further west than St. Louis. By 1976, Major League Baseball had increased to 24 teams and invaded previously untapped markets such as California, Texas, Atlanta, Minnesota, Seattle (briefly) and even Canada, with the expansion Montreal Expos.

Neft et al. (1982) added that in addition to the team expansion, the establishment of black and Latin players led to a different game as, "stars of all backgrounds stocked the clubs which excelled over the period" (p.345). As Witte & Weick (2006) explained, "By the late 1960s, Latino players were becoming commonplace on MLB rosters and were beginning to exert more and more influence on the game as their presence within the league continued to expand" ("The Latino Boom", para. 4).

In addition to the various expansions of the game, rule changes continued to set the precedent for changes on the field. For the 1963 season, the strike zone was officially

redefined (Rader, 1992, p.169). Previously, it had extended from the armpits to the top of the knees. It now extended from the shoulders to the bottom of the knee. This redefinition led to immediate changes. Rader (1992) said:

> For a brief interlude, the six seasons of 1963 through 1968, pitchers regained an ascendancy over the hitters that they had not enjoyed since the first two decades of the twentieth century. Although batting averages had been slipping downward prior to 1963, in the 1963 season major league run totals fell by 1,681, home runs by 297, batting averages by 12 points, and bases on balls by 1,345. Pitchers recorded 1,206 more strikeouts than in 1962 (p.169).

Bill James (2001) goes a step further as he claimed that in the 1963 season, "Baseball's second dead-ball era had begun" (p.249). Baseball was not happy with this trend as they had equated more runs scored and home runs hit with higher attendance.

In response, baseball enacted another rule change for the 1969 season that would swing the favor back to the hitters. James (2001) noted that, "until 1969 no one was regularly checking the height or the slope of the pitcher's mound" (p.250). So in 1969, the rule was changed lowering the mound from 15 inches above home plate to 10 inches above home plate. Additionally, the slope of the mound was defined and the height and slope rules would be strictly enforced.

As if the pitcher did not have enough to deal with as far as rule changes in favor of the hitters, the American League began using a Designated Hitter, or DH, in 1973. The designated hitter would assume a regular spot in the batting order in replacement of

the pitcher (McKelvey, 2004, p.2). He would go on to say that the creation of the DH rule as the most drastic in baseball history, "It was so drastic that one league accepted it and the other league did not" (p.2). He proposed the reason for this change:

> The American League owners…were hoping that designated hitters would provide a spark for the sagging offenses in their league. The American League's owners were also hoping that an explosion in hits, homers and runs would entice more people to come to their ballparks and enable them to catch and then pass the National League in the annual attendance race (p.2).

As is the case with most rule changes since the Live Ball Era, changes were made in favor of hitting because the owners believed offense drove attendance.

The expansion era is named because of the moves of the teams, and not necessarily the rule changes that affected the play on the field, possibly because of the relative equilibrium of the era. James (2001) claimed that, "Expansion favors neither the hitter nor the pitcher, on balance; it does as much to create a shortage of good hitters as it does to create a shortage of good pitchers" (p.307). James (2001) also asserted that there was a great balance between pitching and hitting in the 1970's (p.277).

Alternatively, and because of the six year period where pitching clearly dominated, Cohen (1988) claimed that "Pitchers had assumed control and command of the game as they had at no time since the pre-Ruthian era of the dead ball" (p. 13). It is clear that hitting did not dominate for the entirety of the expansion era, but there appears

to be no clear consensus among baseball historians as to whether the era as a whole was dominated by pitching or if it was simply balanced.

**The Free Agency Era (1977-1993)**

The foundations for free agency were established in 1966 with the Major League Baseball Players Association hiring Marvin Miller to be its executive director. Rader (1992) said:

> Miller's appointment was a decisive turning point in the history of baseball's player-management relationship. Before Miller, the player association had been moribund, an organization routinely used and manipulated by the owners for their own ends. After Miller, the MLBPA became a powerful counterweight to management, the reserve clause fell into shambles, and the players eventually obtained astronomical salary increases (p.186).

The inception of free agency in Major League Baseball in 1976 marked a complete upheaval of a structure that had remained relatively unchanged for almost a century.

The assault on the system began in 1969 when Curt Flood, an outfielder with the St. Louis Cardinals, refused to be traded to the Philadelphia Phillies, and took his case for free agency all the way to the Supreme Court (*Flood v. Kuhn*, 1972). The Supreme Court voted in favor of Major League Baseball in 1969, but the siege was underway.

Only eight years later, after many concessions by Major League Baseball that slowly gave players more power over their careers, full free agency went into effect following the 1976 season (Rader, 1992, p.194). Twenty-four players became free

agents, headlined by Reggie Jackson who signed a five-year $3 million deal with the Yankees (Neft et al., 1982, p.427). This was what Marvin Miller had set out to do. The players saw immediate results in the form of their salaries. As Koppett (2004) put it, "by 1979 the average salary was around $120,000, six times what it had been when the players hired Miller 12 years before" (p.369). The specific rules on free agency have been tweaked many times since then, but the 1977 season saw the first class of true free agents.

Along with all the turmoil off the field between the players and management, the Free Agency Era brought about results on the field that had not been seen in any single era to date. James (2001) said that, "Baseball brought into the 1980's a mixture of styles as rich as the game had had in more than half a century" (p.296). He goes on to point out that in 1980 alone, three players hit over .340, ten players stole 50 or more bases, three players hit over 40 home runs, two pitchers won 24 or more games and one struck out 286 batters. He summed it up by stating, "There have been few ten-year periods in history that could boast of players succeeding dramatically in so many different ways" (p.296).

There was clearly a change in how the game was structured and how players were helping their teams win. But as Rader (1992) said, "Although offensive and defensive statistics remained essentially unchanged, fans witnessed a new kind of game. It featured raw power, dazzling speed, and specialized pitching" (p.209). In comparison to the era preceding it, the statistics had not undergone drastic changes, but the new structure and skillsets of the players made it seem like it. Like the last half of the expansion era, the

rules on the field again favored the hitters as runs per game rose and offense was perceived to have dominated pitching.

**The Steroid Era (1994-2005)**

Many may find it ironic that the steroid era really did not begin until after steroids had been added to Major League Baseball's banned substance list. As Grossman et al (2007) noted, "Steroids finally made it to baseball's banned substance list in 1991, however testing for major league players did not begin until the 2003 season" (p.2). The important lesson is that testing did not begin until 2003, and even then, testing positive held no consequences for the players as a first positive test resulted in treatment for the player (Baseball Almanac, 2012).

As baseball came back from the player's strike in 1994, attendance declined rapidly. Lenhardt (2010) argued that, "steroids may have saved baseball after the 1994-1995 strike, which angered fans and resulted in attendance dropping by almost 10 million in both the National and American leagues" (p.1). Major League Baseball needed a way to bring fans back to the parks after alienating many of them.

It was not until the famed Home Run Race between Mark McGwire and Sammy Sosa in 1998 that attendance numbers recovered. For this reason, the argument has been made that Major League Baseball allowed steroid use to go unchecked, and therefore allowed an explosion of offensive numbers. Penn State professor Charles Yesalis said in an interview with USA Today that owners valued home runs because, "When they were down in the dumps in the early '90s, they saw what pulled them out . . . balls going over

the fence". He gives another reason for Major League Baseball to allow steroid use to go unchecked by saying, "When billions of dollars are involved, they don't want to lose that" (Mihoces, 2003). They may have chosen to ignore it, but the evidence of steroid use was there.

Grossman et al. (2007) stated that, "Evidence of steroid use was rampant. Offensive numbers were way up. In 1996, the Orioles, Mariners, and Athletics all broke their single season home run records" (p.2). Hill & Schvaneveldt (2011) strengthen this argument by saying, "offensive performance sharply increased during the beginning of the Steroids Era and remained at relatively high levels through the 2008 season" (p.2). From the gathered works of baseball historians and statisticians alike, it is agreed that hitting dominated pitching in the steroids era.

**Post-Steroid Era (2006-present)**

The Post-Steroid Era is being added at the discretion of the researcher. There are two main factors for distinction of the Post-Steroid Era. First, beginning with the 2006 season, the penalties for testing positive for steroids (or any banned substance) became harsher than ever before (Associated Press, 2005). In 2005, players testing positive for steroids received a 10-day ban for a first offense, a 30-day ban for a second, a 60-day ban for a third, a 1-year suspension for a fourth, and the penalty for a fifth positive was to be a "Commissioner's decision" (Bodley, 2005). The 2006 season saw the implementation of a system where a player received a 50-game ban for a first offense, a 100-game ban for a second offense, and a lifetime ban for a third offense (Bloom & Molony, 2005).

Since the new system was implemented in 2006, only 20 positive tests for performance enhancing drugs have been confirmed at the Major League level. Since 2006, only Manny Ramirez, Guillermo Mota and Elizier Alfonzo have tested positive twice, inciting 100-game suspensions (Dittmeier, 2012). It is not a Post-Steroid Era in the sense that steroids have been removed from the game. Longtime columnist Ray Ratto (2012) said completely eliminating steroids isn't realistic because, "there will always be players who look for the envelope to see the best way to push it. If there are no more positive tests, that means baseball has essentially stopped caring". The goal is that the harsher penalties help deter players from using because the consequences are greater than ever.

The second factor that led to labeling the period since 2006 as the Post-Steroid Era is the resurgence of pitching. The 2010 season in Major League Baseball was frequently referred to as "The Year of the Pitcher" (Chen, 2010). This is a distinction that was also given to the 1968 season, which led to the immediate and dramatic changes to the pitcher's mound which were chronicled earlier in the Expansion Era (Chen, 2010). For only the third time in Major League Baseball history, six no-hitters were thrown in 2010. During the 2012 season, seven no-hitters were thrown including three perfect games. The resurrection of pitching dominance and harsher penalties for steroids appear to be more than a coincidence.

**The Case for On-Base Plus Slugging Percentage (OPS)**

Baseball has no shortage of statistics to use in attempting to make predictions. In baseball, games are won by scoring more runs than the opponent. Bill James'

Pythagorean Percentage is widely considered the most accurate for predicting winning percentage. It is:

$$Win\% = \frac{\text{RunsScored}}{(\text{RunsScored} + \text{Runs Allowed})}$$

Grabiner (1994) supported James' model by saying:

> At the team level, a good measure of offense should have a strong correlation with runs scored. This means that it should be possible to predict runs scored reasonably well from the measure; the best teams by this measure should score a lot of runs, while the worst teams should score very few (p.1).

Therefore, finding the best statistics to predict a difference between the runs scored and the runs allowed, provides powerful information in order to build a winning team.

In his popular book *Moneyball*, Michael Lewis (2003) relayed what former Oakland Athletics general manager Sandy Alderson found out about predicting differences in runs scored:

> By analyzing baseball statistics you could see through a lot of baseball nonsense. For instance, when baseball managers talked about scoring runs, they tended to focus on team batting average, but if you ran the analysis you could see that the number of runs a team scored bore little relation to that team's batting average. It correlated much more exactly with a team's on-base and slugging percentages. A lot of the offensive tactics that made baseball managers famous – the bunt, the

> steal, the hit and run – could be proven to have been, in most situations, either pointless or self-defeating (p.57).

In order to cut through the "nonsense" that Alderson referred to, one must not simply determine what the greatest predictors of runs scored are, but must utilize them in order to increase winning percentage. Grabiner (1994) reinforced Alderson's claim by saying that batting average does not have a strong correlation with runs scored, and that in fact, it is common for the team with the best batting average to be below average in runs scored (p.1).

Hakes & Sauer (2007) strengthen the argument in their findings that, "two statistics explain the bulk of the variance in winning percentage across teams: the team's on-base percentage and its slugging percentage, relative to the same percentages it allows for opponents" (p.178). On-base percentage (OBP) measures a player's ability to reach base. OBP is calculated by:

$$OBP = \frac{\mathrm{H} + \mathrm{BB} + \mathrm{HBP}}{\mathrm{AB} + \mathrm{BB} + \mathrm{HBP} + \mathrm{SF}}$$

In the equation, H = hits, BB = walks, HBP = hit by pitches, AB = at-bats and SF = sacrifice flies. Slugging percentage (SLG) distinguishes between different hit values and calculates the average number of bases reached for each at-bat. SLG is calculated by:

$$SLG = \frac{1\mathrm{B} + (2 \text{ x } 2\mathrm{B}) + (3 \text{ x } 3\mathrm{B}) + (4\ x\ HR)}{AB}$$

In the equation, 1B = singles, 2B = doubles, 3B = triples and HR = home runs. On-base plus slugging percentage (OPS) is simply the addition of OBP and SLG.

$$OPS = OBP + SLG$$

On-base percentage and slugging percentage on their own are each highly correlated to scoring runs. Albert & Bennett (2001) clarified the rationale for using OPS over either statistic by themselves by saying, "Clearly, SLG and OBP taken together as OPS produce a far-superior model than using either individually" (p.166).

The findings that Alderson used have been confirmed many times. Moy (2006) found that, "the proportion of the amount of variance in runs scored that can be explained by OBP and SLG, is .908" (p.22). Albert (2010) found that, "89% of the total variation in runs scored can be explained by the differences in OPS" (p.3). And in their findings, Hakes & Sauer (2006) found that the difference in OBP can explain 82.5% of the variation in winning percentage, the difference in SLG can explain 78.7% of the variation in winning percentage, and the difference in OPS can explain 88.5% of the variation in winning percentage (p.175).

Compared to all other baseball statistics, OPS has the strongest correlation with runs scored, runs allowed and winning percentage. OPS is typically viewed as an offensive statistic as it measures offensive output. But by calculating OPS Against (OPSa) for a team's pitching staff, offensive output can be calculated against that team in order to easily compare OPS to OPSa. This allows direct comparison of the same statistic so that the impact of offense and pitching can be measured on the same scale.

# CHAPTER III

# METHODOLOGY

## Case Study

It is important to understand why a case study is the preferred methodology for the particular research. As earlier explained, misconceptions about the appropriateness of the case study have led some to undervalue its merits. Creswell's (1998) definition of a case study captures the essence of this particular study as he describes it as, "an exploration of a bounded system or case (or multiple cases) over time through detailed, in-depth data collection involving multiple sources of information rich in context" (p.61). The use of both quantitative and qualitative information and the longitudinal aspect of this study create the rationale for a case study.

The attention given to looking at a single case can render a more complete understanding of what is actually happening. Stake (1995) described such by stating that a case study is, "an intensive, holistic description and analysis of a single instance, phenomenon, or social unit" (p. 27). By looking at a single phenomenon it is imperative to take a more intensive approach because the focus is so narrowed that it is necessary to draw meaningful conclusions. In the end, this provides for a greater explanation of what is being studied.

The purpose of conclusions drawn from a case study is to be able to apply what is learned in the same context. Cassell & Symon (2004) described it as the attempt to, "understand everyday practices and their meanings to those involved, which would not be

revealed in brief contact" (p.325). The objective of research dealing with baseball statistics is ultimately to determine trends to help understand why teams win games. There are so many variables involved, that it is necessary to take an exhaustive approach to each variable in order to fully comprehend its importance.

Yin (2003) supplied three specific criteria in which a case study would be considered the ideal approach. First, the research question must address the study through why, what or how questions. Next, the degree of manipulation of antecedent conditions must have attention paid to it. A case study should have a low manipulation, meaning the researcher has little to no control over the actual events being studied. Finally, the focus of the study should concentrate on ongoing as opposed to historical events.

Based on these criteria, the study of the effect of OPS on winning percentage across the different eras of Major League Baseball aligns itself well as a case study. The research question emphasizes the understanding of winning percentage in different eras compared to the perceptions of each era by asking why, what and how questions. There is no control of antecedent conditions by the researcher necessary as it investigates events that have already occurred. Finally, although historical data is being researched, the focus is to better understand the history in order to better understand the present and future. Sabermetric-centered research relies on the bevy of statistics available throughout the past to better understand it and make more effective predictions about the future (Grabiner, 1994). This study certainly falls in that description.

**Sampling Procedure**

There are two types of sampling typically used, random sampling and purposive or non-random sampling. Each exists for their strengths based on the type of research being conducted. Merriam (1998) stated that purposive sampling, "is based on the assumption that the investigator wants to discover, understand, and gain insight and therefore must select a sample from which the most can be learned" (p.61). Because of this, the effect of OPS on winning percentage across the different eras of Major League Baseball was purposefully selected.

The relative strength of OPS as a predictor of winning percentage compared to other statistics, and its ability to be used as both a hitting and pitching statistic make it ideal for this study. In order to draw the meaningful conclusions that a case study can provide us, the strongest statistics must be used, otherwise the entire endeavor is compromised.

**Procedure of Data Analysis**

There is an abundance of sortable baseball statistics available online. For this study, MLB.com was used because it included all necessary statistics in one place. A Microsoft Excel (2010) spreadsheet was created that included team name, league (National League or American League), year, winning percentage, OPS, OPSa, and era. The era column was created for easy separation of each era for later analysis, starting with the Dead Ball Era represented as era 1 and the Post-Steroid Era represented as era 7.

IBM SPSS (20.0) was used for all analyses as a part of this study. Table 1 displays how the eras will be represented in later analyses.

Table 1

*Era Designations for Analyses*

| Numerical Value | Associated Era |
|---|---|
| 1 | Dead Ball Era |
| 2 | Live Ball Era |
| 3 | Integration Era |
| 4 | Expansion Era |
| 5 | Free Agency Era |
| 6 | Steroid Era |
| 7 | Post-Steroid Era |

All of the relevant information was entered into the Excel spreadsheet by hand. One challenge in the data entry began with the 1949 season. Prior to 1950, the statistic of OPSa was not available. After exhaustively searching various sites that deal in baseball statistics, OPSa could not be found prior to 1950. It was confirmed in an email by Neil (personal communication, March 16, 2012) from baseball-reference.com that the reason the statistic is not available is because prior to 1950 MLB did not officially track doubles and triples against pitchers, making it impossible to accurately calculate slugging percentage against (SLGa). Since SLGa is one half of OPSa, there is no reliable statistic of OPSa prior to 1950.

Not having OPSa prior to 1950 presents an obvious problem in trying to compare OPS to OPSa for each era of the modern era of baseball. As a possible alternative to OPSa, reliable statistics to calculate On-Base Percentage Against (OBPa) are available for every year of the modern era. There are several studies that have concluded that OBP is a better predictor of winning percentage than SLG. Two studies (Moy, 2006; Winston, 2009) both found that OBP is roughly twice as important as SLG. Hakes & Sauer (2006) found that the coefficients for OBP are more than twice as large as the coefficients for SLG (p.175). In a subsequent study, Hakes & Sauer (2007) found that the coefficients for OBP range anywhere from 2.4 to 3.1 times more important than SLG for various periods they examined (p.181). As a result of this, they concluded that, "the relative contribution of OBP to winning is about twice that of Slugging, for the sample as a whole, and in every sub-period as well" (p.181).

The existing literature unequivocally states that OBP is a better predictor of winning percentage than SLG, so in the absence of OPS, it is more effective to have OBP available than SLG. This still presented the problem of whether OBPa was a viable replacement of OPSa when OPSa is not available. Albert & Bennett (2001) suggested that it is a viable replacement by claiming that, "OBP appears to be at least on par with OPS in predicting runs scored for nineteenth century teams" (p.166). This led to the creation of a new column in the Excel spreadsheet, OBPa.

Having OBPa for each year of the modern era allowed a correlation to be run for OBPa and OPSa from 1950-2011, to determine if the claim of OBPa being a viable replacement was true. A Pearson Correlation was run that returned a correlation of

$R^2$=.797 between OBPa and OPSa. As was expected, a strong correlation existed, making OBPa a viable replacement for OPSa prior to 1950. This led to yet another column created in Excel labeled OPSa/OBPa. This simply represented the relevant statistic, either OPSa or OBPa, based on availability.

The next step was to run two separate multiple regression analyses to determine the significance of OPS and OPSa/OBPa on winning percentage for the modern era. The first analysis was run for the years 1950-2011 to represent when OPSa is available. The second analysis was run for the years 1901-1949 to represent when OPSa is not available and therefore OBPa has to be substituted.

The final step to reach a point where hitting and pitching could be compared within each era was two-fold. First, a Univariate Analysis of Variance (ANOVA) for OPS, OPSa/OBPa and Era was run. Three separate ANOVA models were run; Model 1 with just OPS and Era, Model 2 with just OPSa/OBPa and Era, and Model 3 was a full model with OPS, OPSa/OBPa and Era. Interactions between OPS and Era as well as OPSa/OBPa and Era were run as part of the ANOVAs to determine if the differences between the Eras were significant. Then, as part of each model, correlations were run, sorting the cases by era, to return Pearson Correlations between OPS and winning percentage and OPSa/OBPa and winning percentage for each era. Using these correlations, hitting and pitching could be compared directly, within each era, to determine which was more influential to winning percentage.

In the process of running these analyses, it was realized that the Integration Era was split in half by the availability of OPSa. In order to measure pitching by the same

statistic for the entire era, OBPa was substituted for OPSa for the entire era even though OPSa was available for 1950-1960. This distinction was implemented for the ANOVAs and accompanying correlations that were run to compare eras.

# CHAPTER IV

# RESULTS

In order to the get the desired results where the eras could be compared, several statistical analyses had to be run. Included was multiple regression and ANOVA, with specific attention paid to the Pearson correlations accompanying the ANOVAs. The Pearson correlations provided the information to actually compare the eras, but the multiple regression and ANOVA analyses needed to be deemed significant for the Pearson correlations to have any significance to the study.

## Multiple Regression

First and foremost it had to be determined whether OPS was a significant predictor of winning percentage in the modern era, because if not, the remainder of the study would have been moot. This check of significance was determined through two separate multiple regression analyses.

Table 2
*Multiple Regression with Pearson Correlations*

| | |
|---|---|
| 1950-2011 | |
| OPS | .467** |
| OPSa | -.466** |
| 1901-1949 | |
| OPS | .435** |
| OBPa | -.091** |

*Note*. ** = correlation with winning percentage is significant at the .001 level.
For 1950-2011 $R^2$ = .788, for 1901-1949 $R^2$ =.247.
Pearson Correlation between OPSa OBPa = .893

Table 2 illustrates that all of the statistics being used for the analysis are significant at the .001 level. OPS has a positive correlation with winning percentage because it is a hitting statistic, so the higher it is, the more runs a team would score and increase winning percentage. Conversely, OPSa/OBPa is negatively correlated with winning percentage because it is a pitching statistic, so the lower it is, the less runs a team would allow and increase winning percentage.

OPS does not show a large change in its relationship to winning percentage between the two periods with correlations of .467 and .435. For the period of 1950-2011, OPS and OPSa showed nearly the exact same correlation as each other to winning percentage. This meant that despite potential differences in eras throughout this time period, as a whole, hitting and pitching has contributed the same to winning baseball games since 1950. Although the observed correlation for OBPa at -.091 is not as strong as the correlation for OPSa, at -.466, it is still a significant predictor. This strengthens the argument that OBPa is a viable replacement for OPSa. The difference in $R^2$ values from the 1950-2011 period (.788) and the 1901-1949 period (.247) possibly reflect the difference made by substituting OBPa for OPSa. The fact that they are different groups of eras may also have something to do with this.

**ANOVA**

Since the multiple regression returned significant results, it was possible to examine the differences in the eras themselves by our significant measures of OPS and OPSa/OBPa.

Table 3

*ANOVA results for OPS, OPSa/OBPa and Era*

| Model 1 | *p-value* |
|---|---|
| Era | 0.043 |
| OPS | <.001 |
| Era*OPS | 0.032 |
| Model 2 | |
| Era | <.001 |
| OPSa/OBPa | <.001 |
| Era*OPSa/OBPa | <.001 |
| Model 3 | |
| Era | 0.001 |
| OPS | <.001 |
| OPSa/OBPa | <.001 |
| Era*OPS | 0.002 |
| Era*OPSa/OBPa | <.001 |

*Note*. Model 1 $R^2$ = .332, Model 2 $R^2$ = .355, Model 3 $R^2$ = .790

Table 3 illustrates that each of the models provides results that are significant at least at the .05 level, with many significant at the .001 level. What was most interesting is that the interactions tell us that the differences between the seven eras are, in fact, significant. The way in which hitting and pitching contributed to winning percentage in each era is unique to how it contributed to winning percentage in the other eras.

The significant interaction in Model 1 suggests that the way in which OPS affects winning percentage depends on the era. It was clear from the literature and multiple regression analyses that OPS was a significant predictor. This was important because although OPS is a significant predictor of winning percentage in each era, OPS does not

affect winning percentage in the same way from era to era. On the basis of OPS, the eras are unique.

The significant interaction in Model 2 demonstrated the same thing as Model 1 did, but for OPSa/OBPa. Already knowing that OPSa/OBPa is a significant predictor of winning percentage, the interaction revealed that the way in which OPSa/OBPa affects winning percentage depends on the era. Just like for OPS, each of the eras is unique based on how OPSa/OBPa predicts winning percentage.

Model 3 builds off of what was learned from the first two models. Model 3 is a full model, including Era, OPS and OPSa/OBPa in the same analysis. When both statistics are included, it is observed that the interactions are still significant. The fact that the interactions stay significant when both statistics are included in the model signifies that each statistic is a different phenomenon. It is not simply the same thing looked at from different sides. This is important because it allows the confident comparison of the statistics within each era and between the eras.

**Pearson Correlations**

Pearson correlations for OPS and OPSa/OBPa for each era allowed direct comparison to determine whether hitting or pitching contributed more to winning percentage in each era. The multiple regression and ANOVA analyses had built on each other in order to allow the examination of the Pearson correlations. The results of the multiple regression and ANOVA analyses allow the confident comparison of the

observed Pearson correlations because each previous analysis had returned significant results.

Table 4

*Pearson Correlations for Each Era*

| Era | | Pearson's correlations to win% for each era |
|---|---|---|
| 1 | OPS | .540** |
| | OPSa/OBPa | -.619** |
| 2 | OPS | .603** |
| | OPSa/OBPa | -.615** |
| 3 | OPS | .643** |
| | OPSa/OBPa | -.616** |
| 4 | OPS | .553** |
| | OPSa/OBPa | -.548** |
| 5 | OPS | .554** |
| | OPSa/OBPa | -.507** |
| 6 | OPS | .567** |
| | OPSa/OBPa | -.644** |
| 7 | OPS | .522** |
| | OPSa/OBPa | -.564** |

*Note*. ** = significant at the .001 level

Table 4 illustrates the Pearson correlations for OPS and OPSa/OBPa for each era. As would be expected based on previous analyses, all correlations are significant at the .001 level.

It can be determined which statistic contributed more to winning percentage in each era by simply comparing the absolute values of the correlations from Table 4, with

the higher of the correlations contributing more. OPS is positively correlated with winning percentage because it is the measure of runs scored. OPSa/OBPa is negatively correlated with winning percentage because it is the measure of runs allowed.

**The Dead Ball Era**

It is clear from Table 4 that the correlation between OPSa/OBPa and winning percentage (-.619) in the Dead Ball Era is stronger than that of OPS and winning percentage (.540). This is a fairly large disparity between hitting and pitching, so based on this it can be concluded that pitching was more important than hitting in this era. Because of the large disparity, it can be argued that pitching not only was more important, but that it was dominant. This falls in line with the perception that the Dead Ball Era was dominated by pitching. The perception of the era is supported by the statistics.

**The Live Ball Era**

Table 4 shows that the correlation between OPSa/OBPa and winning percentage (-.615) in the Live Ball Era is stronger than that of OPS and winning percentage (.603). These correlations are much closer than the ones viewed from the Dead Ball Era, but they still show that pitching contributed more to winning percentage than hitting. The perception is that hitting drove the Live Ball Era, but the statistics show that pitching was more important to winning baseball games.

The correlations between the Dead Ball and Live Ball eras show pitching remaining almost constant (-.619 and -.615), but that offense clearly took a jump in the

Live Ball Era as the correlations with winning percentage drastically increased from .540 to .603. Hitting undoubtedly became more important than it had been in the Dead Ball Era, which might explain the perception of hitting reigning supreme, but it was not enough to surpass pitching in how it contributes to winning percentage. The perception of the Live Ball Era does not match up with the statistics.

**The Integration Era**

Table 4 shows that the Integration Era is the first era of the modern era of baseball where hitting was more important than pitching. The correlation of OPS and winning percentage (.643) is higher than that of OPSa/OBPa and winning percentage (-.616). The Integration Era saw the proliferation of hitting that had begun in the Live Ball Era finally surpass pitching in terms of its contribution to winning percentage. Pitching, again, remained quite constant as it had between the Dead Ball and Live Ball eras.

Of interest is the steady increase in the correlations between OPS and winning percentage through the first three eras. Of equal interest is the relative consistency of the correlations between OPSa/OBPa and winning percentage. It is understandable that the perceptions of the era were centered on the increased importance of hitting, because clearly, hitting was becoming more important. But the argument can also be made that as hitting was becoming better, the assumption was made that pitching must have been getting worse. The correlations show us that, on the contrary, the importance of pitching had remained almost completely steady for 60 years through the first three eras. In the end, the perception of the Integration Era holds true as the continued rise in home runs helped hitting overtake pitching in importance to winning percentage.

**The Expansion Era**

The Expansion Era exhibits the closest correlations to compare in Table 4. The correlation for OPS and winning percentage (.553) is slightly larger than the correlation for OPSa/OBPa and winning percentage (-.548). The small discrepancy between the two lends weight to the perception that hitting and pitching were quite equal throughout the era.

Of note is the fact that the correlations, both for OPS and OPSa/OBPa, dropped sharply from the correlations of the Integration Era. This begs the question of what would cause such a drop in both the importance of hitting and pitching in relation to winning percentage. Perhaps external factors that had not been present in previous eras took their toll during the Expansion Era. With expansion across the country, teams would have had to endure more travel than in the past. Increasing the number of teams meant more players, many of which were probably not of the same skill set than the present players. Many new teams also meant many new ballparks with new dimensions that players were not familiar with.

It is a reasonable assumption that the rule changes regarding the pitcher's mound in1969 had a part to play in the decline of pitching importance. Many of these variables cannot be factored into this study as it is structured, but it is worth noting for the possibility of future research. For the purposes of this study, the statistics show that hitting overall, contributed more to winning percentage, albeit by the slightest of margins.

**The Free Agency Era**

The Free Agency Era from Table 4 shows the correlations staying in the same range as those from the Expansion Era, as they are drastically lower than the preceding eras. The gap between the correlations for OPS and winning percentage (.554) and OPSa/OBPa and winning percentage (-.507) widened from what was viewed in the Expansion Era.

The correlation observed between OPSa/OBPa and winning percentage is easily the lowest correlation in Table 4. This may lend more fuel to the belief that the rule changes to the pitching mound in 1969 were an over adjustment and favored the hitters far too much. Additionally, with the inception of the DH in the American League in 1973, pitchers had a tougher lineup to go through as they longer had another pitcher to throw to at the bottom of the batting order.

As it was in the two eras preceding it, hitting contributed more to winning percentage than pitching in the Free Agency Era. This falls in line with the perception that hitting led the way because of the rule changes that favored the hitters.

**The Steroid Era**

The difference in correlations for OPSa/OBPa and winning percentage (-.644) and OPS and winning percentage (.567) is not only the widest discrepancy of any era, but it does not align with the perception of the era. The -.644 correlation between OPSa/OBPa and winning percentage also represents the single largest correlation for either stat in any era.

Not only did the -.644 correlation represent the highest observed correlation in all of Table 4, it represented the single largest change in a correlation from one era to another. As discussed, the rules had swung heavily to favor the hitters, which is supported by the correlation of -.507 for OPSa/OBPa from the Free Agency Era. While the correlation between hitting and winning percentage saw a small uptick in importance (from .554 to .567), it pales in comparison to the change in the importance of pitching from the Free Agency Era to the Steroid Era.

The correlation between pitching and winning percentage saw a jump to levels that had not even been approached since the Integration Era. Offensive numbers may have skyrocketed in the steroid era, but this study shows that pitching dominated hitting as it relates to winning percentage in the Steroid Era.

**The Post-Steroid Era**

Table 4 illustrates that the Post-Steroid Era shows a return to the level of correlations that were observed in the Expansion and Free Agency Eras. The trends in the league that led to the distinction of the Post-Steroid Era are upheld in the statistics. The differences in correlations for OPSa/OBPa and winning percentage (-.564) and for OPS and winning percentage (.522) show that pitching contributed more to winning percentage than hitting. Lending more weight to the reasoning for creating the Post-Steroid Era is the fact that the .522 correlation between OPS and winning percentage is the lowest observed correlation for OPS and winning percentage in any era.

Perhaps the distinction between the Steroid Era and Post-Steroid Era holds weight simply for the rule changes, but the statistics presenting the relative importance of pitching shows strong similarities between the two. Not only has pitching been more important than hitting since 2006, but hitting has observed a lower correlation to winning percentage than at any time in the modern era.

# CHAPTER V

## DISCUSSION AND CONCLUSIONS

The results of the analyses provided several intriguing outcomes as they are compared to what is typically understood about each era. In all, the perceptions in five of the seven eras were confirmed, while the perceptions of the Live Ball and Steroid Eras were refuted. Is it merely coincidence that the two eras touted as the most offensively prolific were the two in which the statistics did not match up with the perceptions? Perhaps an explosion of offensive output somehow seems 'sexier' than a game dominated by pitching, so baseball historians tend to romanticize hitting.

The way in which a baseball game is termed may lend some insight into how the public views displays of both offense and pitching. The combination of home runs and many runs scored in general elicit the loudest cheers and are often referred to by terms such as "barn burners," or offensive explosions. The connotation associated with a highly offensive game is exciting, or exhilarating. Conversely, a game featuring a superb pitching matchup is referred to as a "pitcher's duel". The term implies that the game will showcase a struggle to score any runs and will be void of the excitement of an offensive-laden game. With little to no offensive output, many view such a game as boring as hitters are continually set down in order without hits and runs to break the monotony.

The perception that offense was the catalyst in the Steroid Era appears to be refuted by the statistics. The correlations of the Steroid Era in Table 4 provide the type of results this study set out in search of. With the era almost exclusively labeled as

an offensively driven era, the empirical data showing that pitching contributed more to winning percentage may represent the most poignant example of baseball historians misrepresenting an era.

It is interesting that while pitching in the Post-Steroid Era shows a larger contribution to winning percentage than hitting does, it is still a large drop from the correlation observed in the Steroid Era. This would seem to support a belief that the Steroid Era represented a gross miscalculation about the dichotomy of importance between pitching and hitting

The preference for offensive baseball that has driven so many changes throughout the eras may be no better characterized than in recent stadium construction. Beginning in the mid 1990's, baseball saw a wave of new stadium construction. Many of the facilities that were being replaced were built in the 1950's and 60's and were oftentimes built as multisport or multipurpose stadiums. Because of this, most followed a very similar design which ultimately made them aesthetically unappealing, provided poor sightlines and made them unattractive to sponsorship opportunities from corporate America (Egan, 2010).

With the increase in size and scope of the sport industry, the new stadiums built since the 1990's are by and large sport-specific (Egan, 2010). New baseball stadiums have the ability to be built strictly for baseball. Because of this change in design, many more visually attractive stadiums that can be tailored to baseball needs and wants have been constructed. Since the fans prefer an offensive game, many of these stadiums have been built to favor offense, and specifically home run hitters. Parks built since 1990 such

as Coors Field in Denver, Rangers Ballpark in Arlington, Chase Field in Phoenix, Great American Ballpark in Cincinnati, Citizens Bank Park in Philadelphia, US Cellular Field in Chicago, Oriole Park at Camden Yards in Baltimore, and Yankee Stadium in New York and are routinely referred to as "bandboxes" and find themselves in the top 10 for hitter friendly parks (Cockcroft, 2010). In baseball terms, a bandbox is a field with smaller dimensions that favors offense, and particularly home runs.

The fans' propensity to prefer an offensive minded game and the ballparks the game is played in are undoubtedly not the only explanations for the contradiction of statistics and perceptions in the Live Ball and Steroid Eras; but due to the results discovered from this study about the Steroid Era and Post-Steroid Era, the differences between the two eras may warrant a study of its own.

It is also possible that the hitting was so much better across the board in these eras that the teams who could simply field a pitching rotation of decent starters saw a significant advantage over teams who rolled out a rotation of fringe-average pitchers. It seems like more than a happenstance, though, that the eras in which the differences are perceived either in favor of pitching, slightly in favor of hitting, or equal between the two are the eras in which the statistics match up with the perceptions. Future studies should focus on how pitching truly differed between the eras to determine this. Also, a future study might include an appreciation of some aspects of the actual stadiums in which the games are being, or have been played.

It is easily observed by looking at the correlations that there is a noticeable variance in the correlations from era to era. For the purpose of this study, a two-way

interaction was done to determine merely if the eras were significant from each other. Future studies should look into whether or not the specific differences in the correlations from era to era are significant. This would require a post-hoc analysis doing pairwise comparisons of the correlations between specific eras. Additionally, a future study could include a three-way interaction to help determine if the ways in which OPS and OPSa/OBPa change from era to era are different depending on whether the focus is on OPS versus OPSa/OBPa. This would require a researcher with an in-depth understanding of the procedures.

While there are many studies on the subject of baseball statistics (Pujol & Nix, 1994), there seems to be an apparent lack of studies focusing on comparing the eras of baseball. With such sharp distinctions made between the eras for the purpose of looking at baseball history, it makes sense to study the differences between the eras to better understand them. Many of these distinctions were made well before advanced statistics found their home in baseball. There are too much good data and statistics available on baseball to ignore what they can help us comprehend.

It is important to note that no statistic can fully capture what is observed on the field. More clearly put, no statistic measures 100% of winning percentage. There are always factors beyond what is being studied. By no means is the way this study was conducted the only way to measure the importance of hitting and pitching in baseball. With the growth in the use of sabermetrics at the highest levels in the past decade as a tool for making baseball decisions (Woodrum, 2012), there is a plethora of statistics available that can be used to predict winning percentage. This study simply used a

statistic that was already shown to correlate highly with winning percentage and could be directly compared to examine the beliefs that many hold as truths. The examination of other relevant statistics is needed to gain a more complete understanding of the differences of the eras.

This study showed that there are differences between eras based on how winning percentage relates to both OPS and OPSa/OBPa. These differences do not always align with the perceptions about each era. As a whole, it was interesting that a majority of the objective information paralleled the subjective perceptions. Despite a majority of agreement, the differences between perceptions and reality in the Live Ball and Steroid Eras provided many great insights into the legitimacy of some of the perceptions and offer opportunities for further research. Further research should continue to uncover just how accurate the perceptions are.

# REFERENCES

Adelman, M. L. (1986). *A sporting time: New york city and the rise of modern athletics, 1820-1870.* Urbana and Chicago: University of Illinois Press.

Albert, J. (2002). A baseball statistics course. *Journal of Statistics Education*, *10*(2).

Albert, J. (2003). *Teaching Statistics Using Baseball.* Washington, D.C.: The Mathematical Association of America

Albert, J. (2010). Sabermetrics: The past, the present, and the future. In Gallian, J.A. (Ed), *Mathematics and Sports* (pp. 3-14). Washington, D.C.: The Mathematical Association of America.

Albert, J. & Bennett, J. (2001). *Curve ball: Baseball, statistics, and the role of chance in the game.* New York: Copernicus Books.

Associated Press. (2005, November 15). Steroid penalties much tougher with agreement. Retrieved from http://sports.espn.go.com/mlb/news/story?id=2224832.

Barzun, J. (1954). *God's country and mine: A declaration of love spiced with a few harsh words.* Boston: Little, Brown and Co.

Baseball Almanac. (2012). *Steroid suspensions.* Retrieved from http://www.baseball-almanac.com/legendary/steroids_baseball.shtml.

Baumer, B. (2009). Using simulation to estimate the impact of baserunning ability in baseball. *Journal of Quantitative Analysis in Sports, 5*(2), 1174.

Baumer, B., & Terlecky, P. (2010). Improved estimates for the impact of baserunning in baseball. *JSM Proceedings, Statistics in Sports, ASA.*

Bedingfield, G. (2009). Baseball in wartime. Retrieved from http://www.baseballin wartime.com/index.htm.

Beneventano, P., Berger, P. D., & Weinberg, B. D. (2012). Predicting run production and run prevention in baseball: The impact of sabermetrics. *International Journal of Business, Humanities and Technology, 2*(4), 67-75.

Bennett, J. M., & Flueck, J. A. (1983). An evaluation of major league baseball offensive performance models. *The American Statistician, 37*(1), 76-82.

Bloom, B.M. & Molony, J. (2005, November 16). MLB, union agree to stricter drug policy. *MLB.com.* Retrieved from http://mlb.mlb.com/news/article.jsp?ymd=2 0051115&content_id=1268529&vkey=news_mlb&fext=.jsp&c_id=mlb.

Bodley, H. (2005, January 12). Baseball officials announce tougher steroids policy. *USA Today.* Retrieved from http://usatoday30.usatoday.com/sports/baseball/2005-01-12-steroid-policy_x.htm.

Burger, J.D. & Walters, S.J.K. (2003). Market size, pay and performance: A general model and application to major league baseball. *Journal of Sports Economics, 4*(2), 108-125.

Butler, M.R. (1995). Competitive balance in major league baseball. *The American Economist, 39*(2), 46-52.

Cassell, C. & Symon, G. (2004). *Essential guide to qualitative methods in organizational research.* London: Sage.

Chapman, K. S., & Southwick Jr, L. (1991). Testing the matching hypothesis: the case of major-league baseball. *The American Economic Review, 81*(5), 1352-1360.

Chen, A. (2010, July 5). Year of the pitcher. *Sports Illustrated.* Retrieved from http://sportsillustrated.cnn.com/vault/article/magazine/MAG1171503/index.htm.

Cockcroft, T.H. (2010, March 18). Ranking the ballparks. *ESPN.com.* Retrieved from http://sports.espn.go.com/fantasy/baseball/flb/story?page=mlbdk2k10ballparks.

Cohen, S. (1988). *A magic summer: The '69 mets.* San Diego, CA: Harcourt Brace Jovanovich.

Creswell, J.W. (1998). *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage.

Davis, M.C. (2008). The interaction between baseball attendance and winning percentage: A VAR analysis. *International Journal of Sport Finance, 3*(1), 58-73.

Davis, M.C. (2009). Analyzing the relationship between team success and MLB attendance with GARCH effects. *Journal of Sports Economics, 10*(1), 44-58.

Deli, D. (2012). Assessing the relative importance of inputs to a production function: getting on base versus hitting for power. *Journal of Sports Economics*. Advance online publication. doi: 10.1177/1527002511434056.

Dittmeier, B. (2012, May 8). Mota suspended 100 games for positive test. *MLB.com*. Retrieved from http://texas.rangers.mlb.com/news/article.jsp?ymd=20120507&content_id=30646892&vkey=news_mlb&c_id=mlb.

Demmink, H. (2010). Value of stealing bases in major league baseball. *Public Choice, 142*(3), 497-505.

Durant, J. (1973). *The story of baseball in words and pictures.* New York, NY: Hastings House.

Eckstein, H. (1975). Case study and theory in political science. In F. J. Greenstein & N. W. Polsby (Eds.), *Handbook of political science 7,* 79-137. Reading, MA: Addison-Wesley.

Effrat, L. (1947, April 11). Dodgers purchase Robinson, first negro in modern major league baseball. *The New York Times,* p.20. Retrieved from http://mitchell archives.com/wp-content/uploads/2008/10/jackie-robinson-article.jpg.

Egan, M. (2010, December 15). Multipurpose stadiums die out as sports world evolves. *Fox Business.* Retrieved from http://www.foxbusiness.com/markets/2010/12/14/multipurpose-stadiums-die-sports-world-evolves/.

Farrar, A., & Bruggink, T. H. (2011). A new test of the moneyball hypothesis. *The Sport Journal, 14*.

Flood v. Kuhn, 407 U.S. 258 (1972).

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry, 12*(2), 219-245.

Francona, T. (2013). *Francona: The red sox years.* New York, NY: Houghton Mifflin Publishing Company.

Grabiner, D. (1994). The sabermetric manifesto. Retrieved from http://www.sean lahman.com/baseball-archive/sabermetrics/sabermetric-manifesto/.

Grossman, M., Kimsey, T., Moreen, J. & Owings, M. (2007). Steroids and major league baseball. *Berkley University, 30.*

Guttmann, A. (1978). *From ritual to record: The nature of modern sports.* New York, NY: Columbia University Press.

Hall, S., Szymanski, S., & Zimbalist, A.S. (2002). Testing causality between team performance and payroll: The cases of major league baseball and English soccer. *Journal of Sports Economics, 3*(2), 149-168.

Hill, S.E. & Schvaneveldt, S.J. (2011). Using statistical process control charts to identify the steroids era in major league baseball: An educational exercise. *Journal of Statistics Education, 19*(1).

Hakes, J.K. & Sauer, R. D. (2006). An economic evaluation of the moneyball hypothesis. *Journal of Economic Perspectives, 3,* 173-185.

Hakes, J.K. & Sauer, R.D. (2007). The moneyball anomaly and payroll efficiency: A further investigation. *International Journal of Sport Finance, 2,* 177-189.

Honig, D. (1975). *Baseball when the grass was real.* New York: Coward, McCann & Geoghegan, Inc.

Honig, D. (1976). *Baseball between the lines.* New York: Coward, McCann & Geoghegan, Inc.

Houser, A. (2005). Which baseball statistic is the most important when determining team success? *The Park Place Economist, 13,* 29-36. Retrieved from http://www.iwu.edu/economics/PPE13/houser.pdf.

James, B. (2001). *The new Bill James historical baseball abstract.* New York, NY: The Free Press.

James, B. (2009). *The Bill James Goldmine 2009.* Skokie, IL: ACTA Sports.

Koop, G. (2002). Comparing the performance of baseball players. *Journal of the American Statistical Association*, *97*(459), 710-720.

Koppett, L. (2004). *Koppett's concise history of major league baseball.* New York, NY: Carroll & Graf.

Lackritz, J. R. (1990). Salary evaluation for professional baseball players. *The American Statistician*, *44*(1), 4-8.

Lenhardt, M. (2010). The business of steroids in baseball. *Illinois Business Law Journal.* Retrieved from http://www.law.illinois.edu/bljournal/post/2010/03/14/The-Business-of-Steroids-in-Baseball.aspx.

Lewis, M. (2003). *Moneyball.* New York, NY: W.W. Norton & Company.

Lombardi, S. (2006). Baseball eras defined. Retrieved from http://www.netshrine.com/era.html.

Lopez, J., Mundfrom, D. & Schaffer, J. R. (2011) What makes a winning baseball team and what makes a playoff team? *Multiple Linear Regression Viewpoints, 37*(2), 23-28. Retrieved from http://mlrv.ua.edu/2011/vol37_2/Lopezeta l_37_2_Final_3.pdf.

Major League Baseball Statistics (2012). Retrieved February 8, 2012 from http://mlb.mlb.com/stats/sortable.jsp?c_id=mlb&tcid=mm_mlb_stats#elem=[object+Object]&tab_level=child&click_text=Sortable+Player+hitting&sectionType=sp&statType=hitting&page=1&ts=1352406938377.

McKelvey, G.R. (2004). *All bat, no glove: A history of the designated hitter.* Jefferson, N.C.: McFarland & Company.

Miceli, N. S., & Huber, A. D. (2009). If the team doesn't win, nobody wins: A team-level analysis of pay and performance relationships in Major League Baseball. *Journal of Quantitative Analysis in Sports*, *5*(2), 1-18.

Mihoces, G. (2003, December 4). Baseball's steroid test program: Fair or foul? *USA Today,* p. 3c. Retrieved from http://usatoday30.usatoday.com/educate/ondcp/lessons/Activity16.pdf.

Mizak, D. & Stair, A. (2004). The relationship between payroll and performance disparity in major league baseball: An alternative measure. *Economics Bulletin, 12*(9), 1-14.

Moy, D. (2006). *Regression planes to improve the pythagorean percentage* (Master's thesis). Retrieved from http://www.stat.berkeley.edu/~aldous/157/Old_Projects / moy.pdf.

Neft, D.S., Cohen, R.M & Deutsch, J.A. (1982). *The sports encyclopedia: Baseball.* New York, NY: Grosset & Dunlap.

Olson, D. L. (2001). Comparison of three multicriteria methods to predict known outcomes. *European Journal of operational research, 130*(3), 576-587.

Pujol, T.J, & Nix, C.L. (1994). Revisiting team statistical factors contributing to winning percentage in major league baseball. *Research Quarterly for Exercise and Sport, 65*(1), A-59.

Rabinowitz, B. (1989). Baseball and the great depression. In Levine, P. (Eds.), *Baseball History* (pp.49-59). Westport, CT: Meckler Books.

Rader, B.G. (1992). *Baseball: A history of America's game.* Urbana and Chicago, IL: University of Illinois Press.

Ratto, R. (2012, August 23). MLB drug policy fatally flawed. *CSNBAYAREA.com.* Retrieved from http://www.csnbayarea.com/08/23/12/MLB-drug-policy-fatally-flawed/landing.html?blockID=761871.

Schmidt, M. B., & Berri, D. J. (2001). Competitive balance and attendance: The case of major league baseball. *Journal of Sports Economics, 2*(2), 145-167.

Schmidt, M.B. & Berri, D.J. (2002). Competitive balance and market size in major league baseball: A response to baseball's blue ribbon panel. *Review of Industrial Organization, 21*(2), 41-54.

Sommers, P. M., & Quinton, N. (1982). Pay and performance in major league baseball: The case of the first family of free agents. *The Journal of Human Resources*, *17*(3), 426-436.

Stake, R.E. (1995). *The art of case study research.* Thousand Oaks, CA: Sage.

Thomas, G. (2011). A typology for the case study in social science following a review of definition, discourse and structure. *Qualitative Inquiry*, *17*(6), 511-521.

Thorn, J. (1974). *A century of baseball love.* New York, NY: Hart.

Turocy, T. L. (2005). Offensive performance, omitted variables, and the value of speed in baseball. *Economics Letters*, *89*(3), 283-286.

Wieviorka, M. (1992). Case studies: History or sociology? In C.C. Ragin & H.S. Becker (eds). *What is a case? Exploring the foundations of social inquiry*. New York: Cambridge University Press.

Winston, W. (2009). *Mathletics.* Princeton, NJ: Princeton University Press.

Wiseman, F. & Chatterjee, S. (2003). Team payroll and team performance in major league baseball. *Economics Bulletin, 1*(2), 1-10.

Witte, C and Weick, A. (2006). *The history of Latinos in major league baseball and their growing influence on the game.* Retrieved from http://www.umich.edu/~ac213/student_projects06/witaw/index.html.

Woodrum, B. (2012, February 7). 2012 sabermetric teams: The market for saber players. *Fangraphs.* Retrieved from http://www.fangraphs.com/blogs/index.php/2012-sabermetric-teams-the-market-for-saber-players/

Yin, R.K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage.

Zimbalist, A. (1992). *Baseball and Billions: A probing look inside the big business of our national pastime.* New York, NY: BasicBooks.