THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF FINANCE

HITTING THE MONEYBALL INTO THE LUXURY BOX

CHRISTOPHER LOGGIA
SPRING 2016

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Finance
with honors in Finance

Reviewed and approved* by the following:

Robert Novack
Associate Professor of Supply Chain Management
Thesis Supervisor

Brian Davis
Professor of Finance
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

This paper aims to determine if utilizing Moneyball roster management tactics improves the revenue generation and profitability of a franchise from a franchise owner's perspective. Ultimately, if a baseball team makes a specific effort to utilize Moneyball tactics, would the overall value of the franchise improve. All data is sourced from historical Major League Baseball statistics over eighteen seasons from 1997 to 2014. The analysis includes the creation of a new player valuation model by analyzing the career average Z-score to salary efficiency ratio. The analysis will test the relationship between a number of variables, but largely between the Z-score to salary ratio and total annual fan attendance and the Z-score to salary ratio and total annual revenue. The study found that by using a standard linear regression, there is a positive relationship between the Z-score to salary efficiency ratio with each of annual attendance and annual revenue. However, there is no relationship between the Z-score to salary ratio and winning percentage. Additionally, by using a forced constant regression, there are no positive relationships between the Z-score to salary ratio to any of annual attendance, revenue or winning percentage.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Novack for not only being my Thesis Supervisor, but an excellent AKPsi chapter advisor and friend to discuss baseball. I would like to thank Dr. Woolridge for teaching me something new every day in the Nittany Lion Fund and introducing me to the world of Finance. Finally, a thank you to Professor Davis for encouraging me to pursue a topic of interest and guiding my thesis experience.

Chapter 1

Introduction of Topic

The concept of Moneyball has existed since the beginning of the new millennium. This paper will utilize both the original concepts of Moneyball and more modern interpretations including the arbitrage implications.

Concept of Moneyball

"Swung on, there's a high drive hit way back to right center field. That one is gone! And it's 20 consecutive victories for the Oakland Athletics on an unbelievable night when they lost an 11-0 lead. And now they win it. Hatterberg is mobbed at home plate. The crowd comes back to insane life. Crazy…just plain crazy. How do you explain it?" Bill King's voice rang throughout the Oakland Coliseum on September 4th, 2002, marking an historic twenty game winning streak for the Billy Beane led Athletics on their way to clinching an AL West pennant. The scene is immortalized in Aaron Sorkin's film Moneyball setting the stage for the A's to win the 2002 World Series. Although the film dramatizes some of the Oakland A's 2002 season, both the film and Michael Lewis' book Moneyball: The Art of Winning an Unfair Game popularized the theory of Moneyball across Major League Baseball front offices.

As a small market team with limited financial resources, the Oakland Athletics implemented Moneyball in an attempt to win championships, and succeeded to an extent that no one had expected. After four consecutive playoff appearances, the A's managed to consistently compete with big market teams that were able to spend extensively to attract the best overall talent. In short, the theory of Moneyball holds that a disequilibrium existed between player salaries and expected contribution due to traditional scout analysis emphasizing and rewarding flashy statistics such as home runs, runs batted in (RBIs) and batting average. However, Bill James, the father of sabermetrics in baseball unearthed the failure of traditional statistics to capture the comprehensive contribution of all major leaguers. More specifically, James recognized that teams undervalued players' ability to get on base, regardless of the method. In doing so, the Oakland A's gave further consideration to the on-base percentage (OBP) rather than the simple batting average (AVG). (Hakes and Sauer, 2006). Batting averages calculate the frequency of a batter getting a hit for every at-bat, whereas the OBP calculates the frequency a batter reaches base for every at bat, including walks. In theory, the value of hitting a single or getting a walk has the same end result: standing on first base. James and the A's successfully demonstrated this inefficiency in traditional baseball theory by offering small market contracts to players who successfully reached base regardless of the method as opposed to spending big money on home run hitters and strong average hitters. Eventually, after comprising a roster full of strong on-base percentage players, the A's consistently reached the playoffs while consistently ranking towards the bottom of the league in total payroll. (Regan, 2012).

Moneyball Spreads

Moneyball has spread extensively to the Major League Baseball community over the past decade in an attempt to maximize team winning percentages. Over time, teams continue to provide statistical analysis to discover market inefficiencies regarding better ways to analyze the comprehensive value of a player. Consider the recent popularity of the wins above replacement (WAR) statistic These tactics have expanded to include applications outside of the Major League Baseball community. For a further example, in basketball the real plus or minus (RPM) analysis has become an integral role in free agent signings. The RPM in basketball adjusts for variables including ability to facilitate team play and defensive strength that were previously considered intangible measures (Ilardi, 2014) This advanced statistical method isolates the individual player's performance both offensively and defensively negating the impact of play of other players on the court. All of these statistical analyses have been conducted in order to scout better player talent to contribute to total team wins. However, this thesis intends to analyze the impacts of Moneyball in a different manner.

While it is safe to assume that the majority of Major League Baseball has adopted Moneyball tactics, each franchise and its scouts analyze and interpret the data differently depending on what they find to be the most valuable for their fan bases. Over the past half-decade, a number of front offices across a number of sports teams have increased their hiring of Wall Street professionals to maximize the implementation and return of Moneyball. Hopefully this research will provide an introductory player valuation model that can be formally adopted by a large number of sports franchises. Yet, as general managers begin to adopt similar models, the advantage the models or analyses discovered dwindle. Similar to the world of finance, increased capital and labor dedicated to analyzing Moneyball arbitrage opportunities may permit for high speed algorithms to immediately capitalize on any arbitrage opportunities. But the beauty of Moneyball in baseball is that signing players is more than just a simple contract, it requires a relationship and incentives to attract players to particular markets.

Potential Corporate Applications of Moneyball

When Moneyball was initially introduced, it was intended to be utilized by GMs and managers in order to put together a lineup that would maximize on-base percentage while maintaining a small budget. According to Forbes, Major League Baseball is worth a combined $36 billion (Badenhausen, 2015), less than only the NFL which is worth $45 billion. However, in the big picture each baseball team can be considered its own individual corporate entity attempting to maximize its profitability and growth from year to year. In doing so, there are significantly more variables to consider than simply winning individual games and championships. Different professional sports teams markets will have different expectations for what is an acceptable winning season. For example, the New York Yankees fan base expects to make a legitimate run at the World Series every year whereas the Oakland Athletics fan base finds a pennant title similarly exceeding expectations. Factors such as these will have similar impacts on fan happiness which ultimately impacts the attendance at baseball games which drives profits at an organization level. Can managers and GMs utilize Moneyball tactics to assemble winning (above 0.500) rosters while the owner of the team can still maximize the profit of his/her baseball team? How does the style of Moneyball impact fan attendance, fan happiness, fan base growth, championship likelihood, and in-game revenues? Can an organization assume Moneyball tactics and develop a portfolio valuation model to present predictive power for individual player contributions to overall firm profit maximization as opposed to simply winning percentages?

Thesis Statement

As calculated through the created player valuation model, Moneyball tactics will result in positive correlations of the Z-score / salary ratio to each fan attendance, total annual revenue and annual winning percentages.

Chapter 2

Literature Review

Initial Impressions on Fan Attendance and Moneyball

Although there has not been an extensive amount of academic and professional literature published on the topic at hand, several key themes emerged from the literature review. From 1998-2008 the Athletics were the most payroll efficient team in Major League Baseball (MLB), holding the fifth highest winning percentage in the league, despite having the seventh lowest payroll. For example, over the time span, the Athletics averaged a payroll of $52 million compared to the Yankees’ league high $167 million. Statistically speaking, the Athletics were two and a half standard deviations above the linear regression measuring overall efficiency, with the next closest competitors being the Minnesota Twins at one and a half standard deviations (Regan, 2012). The journal surprisingly states that teams that implement the Moneyball practices often improve their winning percentages but have unexpected negative consequences related to overall fan attendance. The efforts to conserve outs with the Moneyball theory contributed to the

Athletics' increased winning percentage but also helped cause the team's attendance records to fall from the sixtieth percentile to the eightieth percentile. Regan claimed that even though the team was winning more often and reaching the playoffs, the lack of homeruns and the increased focus on walks brought about a much slower and less engaging style of play that decreased fan attendance (Regan, 2012). Further, Moneyball tends to produce a significant amount of roster turnover in the constant pursuit of undervalued players and the shedding of overvalued players. A different study demonstrated that each percentage point increase in roster turnover leads to a 0.7 percent decrease in annual fan attendance (Kahane and Shmanske, 1997). The study additionally claims and supposedly proves that variables including winning percentage, lagging winning percentage, roster tenure, number of all-star players, newness of stadium and home runs per game all demonstrate a positive and direct relationship with fan attendance. Roster turnover and ticket prices are negatively related to fan attendance. Yet, a different source argues that fan attendance is positively correlated with Moneyball tactics when those strategies lead to increased winning percentages and playoff appearances.

Peripheral Revenues

The clear absence on studies of peripheral franchise revenues is apparent in Major League Baseball across the literature. Hakes and Sauer insinuate that there is a positive correlation between winning and off-field profit maximization but requires greater definition (Hakes and Sauer, 2006). While the majority of this analysis will focus on Moneyball's impact on fan attendance and other revenue generating streams, models exist that predict fan attendance based off variables such as weather, opponent, team performance, the night of the week and the time of the year (Kleps, 2014). These models will be useful in predicting the change in impact on fan attendance, ticket prices etc. from Moneyball, but will certainly need to be adjusted. Other authors have focused on potential intangible impacts. For example, in the professional cricket industry, one author claims that it is possible for a player's "non-cricket abilities to be more important than the actual talent considering it is a media entertainment business (Rastogi and Deodhar, 2009). Rastogi further implies that contract premiums should be awarded to players who stir up drama, interact with film stars and are fan favorites as these factors directly flow to positive changes in ticket revenues, broadcasting revenues, sports merchandising and other advertising opportunities. The validity in the increased peripheral revenues theory is acceptable but this research intends to attempt to relate the study to an American sports industry which may or may not have similar fan interests as the cricket industry overseas. To further the analysis of player popularity, studies have been performed in the National Basketball Association that demonstrate winning championships and individual player accolades tend to improve long term player popularity among fans, measured in all-star appearances. Additionally, this popularity is elevated when playing alongside other established stars in the league (Yang and Mengze, 2011). Again, once this analysis was done in the NBA, this research will be able to provide similar research and insights into the MLB realm regarding player popularity based off the fan voting for all-star appearances over the last decade. These results likely support larger-market teams benefitting from non-Moneyball strategies, but is it possible to attract stars before they are stars with Moneyball analysis and still reap the same franchise revenue benefits?

Salary Efficiency and Arbitrage Opportunities

Literature regarding the efficiency in baseball wages and the impact of sabermetrics on that efficiency surfaced and verified that hitters' salaries did not accurately reflect the contribution of various batting skills to winning games. For example, GMs frequently undervalued the ability to avoid making an out (need to account for walks) is just as important as scoring runs. As previously stated, the statistics of on-base percentage and slugging percentage tend to serve as better performance indicators than batting average and home runs because the aforementioned metrics consider total bases amounted and the value of sacrifices and walks. Since walks are ways to avoid outs, they should be considered equivalent to singles. Additionally, the study proves the value of the combination of these metrics as they explain 88.5 percent of the variation in a regression of winning percentage when taken together, but only 82.5 percent of variation when team OBP is compared to opponent OBP and 78.7 percent when team slugging percentage is compared to opponent slugging percentage (Hakes and Sauer, 2006). These statistics are quite valuable considering that "homerun hitters" make on average three to four million dollars more per year than average hitters. Yet, a team filled with players that tend to walk tend to force pitchers to throw more pitches resulting in fatigue which eventually forces those pitchers to throw a greater number of strikes or to be replaced.

While the majority of literature supports the theory that inefficiencies occurred during the initial implementation of Moneyball and was corrected by 2006, there will never be conclusive evidence that the arbitrage opportunities do not exist until all possible statistics and combinations of statistics have been examined, which is nearly impossible (Baumer, 2014). Potential talent and contract arbitrage also exists in an organization's decision to utilize organic or inorganic growth. According to traditional baseball wisdom, franchises benefit significantly from having a well-developed scout team and farm system. For example, during the early 2000s, one author claimed that high school players entering the MLB earned a higher premium than college drafted players due to scouts' opinions (Caporale and Collier, 2013). The intangible benefit, or detriment, and group think provide further opportunities for arbitrage. Further, to relate more to the MLB, the Houston Astros process under Jeff Luhnow in the early portion of the 21st century was driven by data analytics in an effort for middle-market teams to invest in their future at a cheaper price than pursuing big-name free agents who would only provide a short-term boon to peripheral revenues at an inefficient cost in the form of the salary. The author concedes that the strategy likely would be impossible to follow for large market teams due to the instant demand to win, such as New York or Boston (Green, 2014). In a similar argument, one author claims that arbitrage exists because teams undervalue the present value of current major leaguers and tend to overvalue the future value of young prospects. For example, the Oakland A's were able to trade several top prospects for a greater number of currently talented major league players that could contribute to winning now with certainty as opposed to taking a risk on future performance, injuries or other off-field issues with minor leaguers that may have never even made it to the major leagues (Lemire, 2014). Further arbitrage and competitive advantage occurs on both ends of the trade between small market and large market teams when small market teams find and develop underappreciated players who contribute now that eventually receive premium contracts from large market teams. Small market teams benefit (arbitrage in a sense) in that their peers overpay consistently average players while large market teams benefit from reduced development expenses and get guaranteed contributions (Grier and Cowen, 2011). Both of these studies offer further evidence that there are a variety of alternative arbitrage opportunities that exist and that there are likely a number more to be discovered in the future with additional research. The ability to attract consistent players is vital, but this research will analyze how to predict this consistency at a reasonable price to reduce franchise costs.

Through several articles, it is apparent that current studies have minimally studied the impact and correlation of salaries and Moneyball, but have left many other variables such as team chemistry, fan popularity, leadership, likelihood of injury, peripheral revenues generated, average pitch counts and more to be studied when maximizing total financial returns for the franchise as opposed to just winning championships. The challenge to traditional baseball wisdom and analysis must be consistent and relentless to discover new potential sources of arbitrage to ensure profit maximization (The Economist, 2015).

A New Direction

After reviewing the published literature on the topic, it appears that there are a number of concepts that have been discussed briefly, but lack complete analysis and at times, analysis in the proper realm of this research. Nearly all of the literature points in the direction that arbitrage opportunities still exist and will always exist as new ideas and perspectives reach the market. But most importantly for my intended study, that arbitrage opportunities for profit maximization currently exist that are not being utilized. For example, it appears that teams have focused on winning championships as the primary goal of the franchise. But it seems that there is merit, as a franchise owner, to potentially focus on maximizing fan attendance, signing fan popular players and minimizing contract expenses in order to optimize franchise profits. In doing so, these owners indirectly benefit from increased ticket sales, advertising revenues, broadcasting revenues and general brand value as opposed to simply winning championships. In order to further analyze these figures, this research will combine aspects of the individual research already completed in different industries and for specific correlations to project future revenue streams and profit changes.

Additionally, this research intends to utilize past and future projections to establish a player projection model for a variety of statistics that will have different weightings in terms of value. These weighted averages will then be combined and compared to a variety of other players without stating the name to demonstrate that arbitrage opportunities exist when general managers do not get caught up in big name acquisitions. These projections will be created via a player model that incorporates traditional statistics such as hits, RBIs, stolen bases, homeruns and newer metrics such as OPS, WAR and slugging percentage. Most importantly, this research will also include variables such as fans attracted per player, player popularity, injury likelihood, leadership and jersey sales. After analyzing these values for current active players in the MLB, this research will then be able to optimize thirty portfolios of random players from a variety of different teams and contracts and see how they contribute to overall franchise profitability and revenue generation. The concept of predicting revenue and profitability based off a career average Z-score / salary ratio has not been established in any of the previous research examined, particularly relating it to the overall profitability as a franchise owner. While this thesis will not test the success or failure of Moneyball as this has already been established, it will largely explore the impact on peripheral revenues and expenses for a franchise owner.

Chapter 3

Data Methodology

To best analyze the impact of Moneyball tactics on franchise value, the analysis is comprised of two parts. The creation of a player valuation model provides an individual analysis of all starting batters in Major League Baseball for eighteen years. Based on these results, a series of correlations serve to test the proposed thesis that utilizing Moneyball tactics in managing a baseball lineup improve the overall value of a baseball franchise.

Data Collection

Annual Player Statistic Data

Player statistics from 1997-2014 were sourced entirely from BaseballGuru.com's database which is readily provided to the public. Although not all of the following data points were used in the player valuation model, the collection includes games played, at bats, runs scored, hits, doubles, triples, homeruns, RBIs, stolen bases, times caught stealing, walks, strikeouts, intentional walks, hit by pitch, sacrifice hits, sacrifice fly, ground into double play, batting average, on-base percentage, slugging percentage and select years include data on fielding statistics such as assists, putouts, errors, double plays and fielding percentage. This analysis calculates a summary statistic called total bases earned which includes hits, doubles, triples, home runs and walks. These statistics were collected and organized for all batting positions.

A career average annual salary was derived for all players. BaseballGuru.com provided annual salary data from 1997-2004. After 2004, the remainder of the annual player salaries were sourced from Baseballreference.com which collected a variety of sources to determine specific salaries. By taking an average of every player's career earnings and the number of years played, the analysis provides an expected annual salary for the player valuation model to hold.

Annual Revenue Data

Major League Baseball franchises are private entities and are therefore not forced to publish annual financial documents available to the public. With that as a given, both Forbes and Bloomberg have online authors that publish annual estimates of franchise values, revenues, operating profit, a limited list of peripheral revenues and media deals. Bloomberg provides greater detail into line items including gate receipts, concessions revenues, sponsorship revenues, parking and other revenues. A further breakdown of Bloomberg's media analysis includes MLB revenue sharing, regional sports network share, other team owned enterprises and MLB advanced media profit sharing. However, Bloomberg's annual publications were not readily available dating back to 1997.

This analysis utilizes Forbes' estimates for annual total revenues for every MLB team during the period 1997-2014. The annual figures will be calculated in millions of U.S. dollars but will not be adjusted for inflation (See Appendix A: Annual Team Revenues Summary 1997-2014). While Forbes and Bloomberg differ on both team enterprise value and revenue figures, they are relatively similar which lends support to accept Forbes' estimates due to the availability of the otherwise private data. Please note that these figures are all estimates from Forbes since the data is private. The real figures likely differ slightly but should have an immaterial impact on the analysis as the same level of error would be made for all teams since Forbes maintains similar assumptions throughout. Note that Forbes' 2002 publication was unavailable and the analysis assumes that 2002 revenues are equivalent to those earned in 2001.

Annual Attendance And Winning Percentage Data

Annual attendance and winning percentage data from 1997-2014 were derived from different sources due to availability. Annual ballpark attendance for each MLB team was collected from Ballparksofbaseball.com (See Appendix B: Annual Attendance 1997-2014). While the data in the index is shown as reported, the analysis converts the data to be in millions in order to be consistent with the units of the Z-score / salary ratio.

Annual winning percentage data for each team from 1997-2014 was sourced from Baseballreference.com (See Appendix C: Annual Winning Percentages 1997-2014). Note that both the Arizona Diamondbacks and the Tampa Bay Rays were not in existence in 1997 and therefore have a zero placeholder for revenue, attendance and winning percentage.

Player Valuation Model

The construction of the player valuation model began with the collection of historical statistics for every MLB batter between the years 1997 and 2014. The categories include: games played, at-bats, runs scored, hits, doubles, triples, homeruns, total bases collected, runs batted in, stolen bases, times caught stealing, walks, strikeouts, intentional walks, hit by pitch, sacrifice hits, sacrifice flies, grounded into double plays, batting average, slugging percentage, on-base percentage, putouts, assists, errors and fielding percentage. These figures were individually established for catcher, first base, second base, third base, and outfielders. The data sets were sorted to only include players that started eighty-one games or more to represent half of an MLB season. Summary statistics, including both mean and standard deviation, were calculated for these abbreviated data sets for each position. To ensure that the data was normalized as per the law of large numbers, histograms within plus or minus three standard deviations from the mean were created for each statistic for each position.

Once the data sets were assumed to be normalized with an anchor of zero, a Z-score was calculated for each statistic for each player based off the previously calculated means and standard deviations. The Z-scores were then utilized in the player valuation model to create a weighted average player efficiency Z- score for each player's performance per year. The weighted average Z-score takes into consideration games played, total bases procured, runs batted in, strikeouts, batting average, on-base percentage, slugging percentage, at-bats and runs scored. The weightings were 10%, 25%, 10%, 5%, 10%, 10%, 10%, 10% and 10%, respectively as seen in Table 1. The noted statistics were the selected tools for analysis because as per most current scouts, they appear to be commonly analyzed statistics in modern baseball. Considering that most batters have played in more than eighty-one games for more than one year in their career, this analysis method will utilize a career average player efficiency Z-score over the span of the entire career into one figure. In aggregate, this figure will demonstrate the player's pure offensive contributions towards a team's total ability and the individual player's value.

Table 1: Player Valuation Model Z-Score Weightings

Z-Score Weightings				
Games Played	10%		Batting Average	10%
Total Bases	25%		On-base Percentage	10%
RBIs	10%		Slugging Percentage	10%
Strikeouts	5%		At Bats	10%
Runs	10%			

However, a statistic that may be more telling is the weighted average player efficiency Z score to average career salary ratio (Z-score / salary ratio). To better incorporate the concept of franchises analyzing a player's contributions relative to the financial returns and costs, as is the aim of the paper, the Z-score / salary ratio will utilize the player's average career annual salary. The more positive and further away from the mean that the Z-score / salary ratio demonstrates, the more cost efficient the player is at delivering offensive outperformance. This measure will be the most telling statistic to indicate the Moneyball value that an individual player affords a team. While Moneyball infers that there are arbitrage opportunities based off a lack of analysis for a particular statistic, this paper analyzes whether the weighted Z-score / salary ratio may be a more telling measure for scouts and GMs to follow. Players who have a positive ratio will be considered for the Moneyball analysis performed and detailed later in this paper.

Correlation Analysis

Before entering the realm of future implied correlation analysis, it is vital to take a step back and analyze historical MLB franchise performances both on and off the field. To further the study, the analysis backwardly defines the annual starting roster for every MLB team from 1997 to 2014 based off the player on the team with the most games played for each position. For each player in the starting lineup, the Z-score / salary ratio will be assigned and the team average for that year will be calculated. These statistics will be summarized from 1997 to 2014 for each team.

Based on historical figures as collected by Baseballreference.com, Ballparksofbaseball.com and Forbes, annual winning percentages, annual total attendance and annual team revenue figures have been calculated (See Appendices A-C). These data are calculated for each franchise from 1997 to 2014 to match the collected data and statistics.

There will be three correlations performed for each franchise to test a number of relationships in order to further the analysis in the Moneyball field. The first, and potentially most telling relationship, will be the regression between the team average Z-score / salary ratio to the total attendance. The paper will also explore the regression of the team average Z-score / salary ratio to the total team revenue figures from 1997 to 2014. The Z-score / salary ratio to annual winning percentage will be examined to determine if there is a relationship. Finally, The analysis assumes that there is a positive relationship between attendance and total revenue.

Once these regressions have been computed, the established equations will be utilized for thirty new “portfolios” of Moneyball players to test whether the Moneyball tactics improve the profitability of existing MLB franchises. Each franchise will have a randomly selected Moneyball qualifying participant (Z-score / salary ratio above zero) for each position: catcher, first base, second base, third base, shortstop, three outfielders and a designated hitter for each year. The designated hitter will be assumed for both the American League and the National League and will simply be a duplicate of the first listed outfielder since there is insufficient data on designated hitters over the tested time period. Each new portfolio will then take the average ratio value for the specific year. The average ratio will then be substituted into the established formulas as stated above to derive the annual attendance, annual revenue and anticipated annual winning percentage for each major league team. The analysis will then compare how many seasons the derived Moneyball values were greater than the historical values.

Data Methodology Summary

The analysis in essence will test whether there is a correlation between the Z-score / salary ratio and attendance, the Z-score / salary ratio and total revenue and the Z-score / salary ratio and winning percentage. Please see Figure 1 below for a visual representation of the proposed analysis:

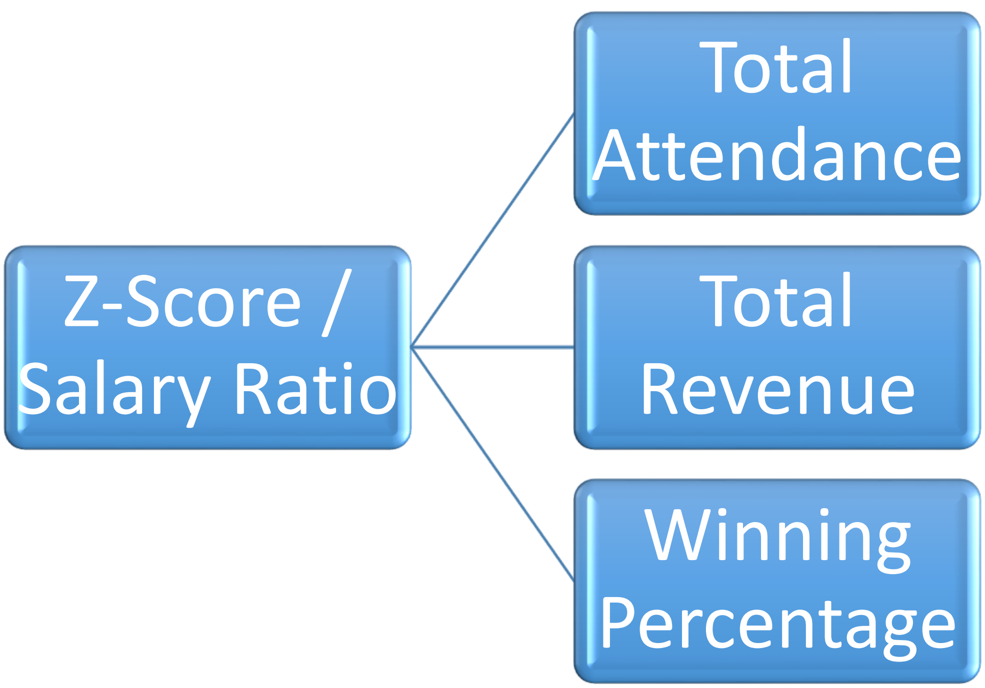

Figure 1: Z-Score / Salary Ratio Relationship Summary

H0: There is no positive correlation between Z score / salary ratio and total annual attendance.

H1: There is a positive correlation between Z score / salary ratio and total annual attendance.

H0: There is no positive correlation between Z score / salary ratio and total annual revenue.

H2: There is a positive correlation between Z score / salary ratio and total annual revenue.

H0: There is no positive correlation between Z score / salary ratio and winning percentage.

H3: There is a positive correlation between Z score / salary ratio and winning percentage.

Chapter 4

Data Analysis

The majority of the analysis performed in this thesis revolves around regression analysis. Throughout the analysis portion of this paper, there will be several mentions of simple linear regression and forced-constant linear regression. Demonstrations and results of both will be displayed and analyzed. However, to note the difference between the two, simple linear regression provides a bet-fit line without restricting the data set in any way. On the other hand, forced-constant linear regression ensures that the best fit line passes through the origin of the axis with an intercept of zero. Typically, forced-constant regression analyzes categorical data in which there is a binary choice in which a particular outcome is desirable, denoted by a one, and an unfavorable outcome is denoted by a zero. In this sense, forced-constant regression is most useful in a scenario in which the intercept is known to be zero. Yet, this research attempts to determine the power and predictive ability of the created player valuation model and will analyze both types of regressions regarding the F factor significance and its related R^2 measure.

In short, the closer to zero that the significance F result is the more powerful the model appears. Additionally, the closer to one that R^2 is infers that a greater portion of the variance is explained through the variables in the model. In tandem, regression results that indicate a low significance F and a high R^2 demonstrate that the model has derived a more meaningful relationship. Whether that relationship is positive or negative is independent of these two variables.

Analysis Wide Assumptions

Before reviewing the results of the regression, please note several assumptions. The Z-score / salary ratio assumes that all analyzed player statistics are normally distributed as proven by the histograms derived from league-wide figures. Only players that have played in 81 games or more per season qualify for that season to contribute to the normally distributed position statistics. Having a Z-score / salary ratio greater than zero infers that over a specific player's career they have performed higher than the standard normally distributed mean of zero, also known as the average league player at that position. Although a player may have been coined the starter at the beginning of the year, the historical portion of this analysis assumes that the starter is the player that appeared in the most games at a specific position. Both the American League and National League are assumed to utilize a designated hitter in starting lineups. Due to lack of designated hitter data available, this analysis assumes that the designated hitter is simply a repeat of the first listed outfielder for that specific year (See Appendix D).

Annual salaries are assumed to be an average of all career earnings for each individual player. Therefore, in this analysis teams are never paying the maximum annual salary of any player as their lower salaries reduce the annual price of the player. These annual expected salaries remain constant for each player regardless of which year they were analyzed in.

To qualify for the Moneyball rosters, players must have a Z-score / salary ratio of zero or greater to imply that the team receives a positive on-field contribution for every dollar spent. Every team was randomly generated a Moneyball starting lineup (See Appendix D). Random generation of rosters is assumed to simulate free agency to demonstrate that all teams would need to have an equivalent opportunity to acquire every player. Taken to extremes, this model would assume unlimited spending on salary, which is consistent with baseball's unlimited salary cap. For teams to maximize the benefit of the valuation model, all markets would need to be assumed to be large enough to have unlimited capital.

The analyzed relationships are determined to be somewhat linear in order for the linear regressions performed to hold any meaning. The input for all of these relationships is the Z-score / salary ratio which drives the dependent variables attendance, revenue and winning percentage, if any at all. To extend the analysis from individual teams to the entire Major League Baseball, the average attendance, revenue and winning percentages of all thirty teams are calculated as the annual results for the entire league. Similarly, rather than generating an MLB Moneyball roster, the analysis takes an average of all thirty teams' randomly generated Moneyball roster Z-score / salary ratios.

The regressions are conducted to establish a linear equation in the form of $Y = mX + B$ in which the X is the Moneyball Z-score / salary ratio. By using the Moneyball input, the analysis derives expected values for attendance, revenue and winning percentage which are then compared to the historical results. The analysis will detail what percentage of the years that the Moneyball generated team outperforms the historical results.

Z-Score / Salary Ratio And Total Annual Attendance

Regression Results

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.595762
R Square	0.354933
Adjusted R Sq	0.314616
Standard Error	0.090651
Observations	18

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.072345	0.072345	8.803615	0.009081613
Residual	16	0.131483	0.008218		
Total	17	0.203828			

	Coefficients	*Std Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	2.460931	0.043935	56.01321	8.64E-20	2.367793513
Z/Salary	0.579553	0.195327	2.967089	0.009082	0.165477953

Figure 2: Standard Regression MLB-wide

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.859757
R Square	0.739181
Adjusted R Sq	0.680358
Standard Error	1.234653
Observations	18

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	73.44319	73.44319	48.1794	3.3224E-06
Residual	17	25.91428	1.524369		
Total	18	99.35747			

	Coefficients	*Std Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	0	#N/A	#N/A	#N/A	#N/A
Z/Salary	-8.98035	1.293786	-6.94114	2.38E-06	-11.7099989

Figure 3: Forced Constant Regression MLB-wide

As previously mentioned, both a standard linear regression and a forced constant linear regression were conducted for each relationship. As can be seen in Figure 2 and Figure 3, both regressions returned a significance F factor lower than five percent, indicating statistical significance. These results both speak to the power of each model. The lower the significance F factor, the stronger the model. Clearly, the forced constant regression yielded stronger results in terms of predictability. It appears to be logical that the forced constant regression would yield more powerful results for all of the relationships because the Z-score / salary ratio benchmarks player performance to a league average value of zero, indicating that the player performs exactly average relative to the rest of the league. Utilizing an average of zero for player rankings, it appears to be logical to assume that the constant in a regression analysis and the established equations that the constant should be zero.

The forced constant regression has a stronger significance F factor and also explains greater variability that occurs in the statistics as is seen by its R^2 value of seventy-four percent relative to the standard regression's R^2 value of thirty-five percent.

While the forced constant regression better explains the model, the expected annual attendance of the Moneyball teams returns a negative value, which is impossible, after deriving the linear equations based off the regressions. This is a consistent result throughout all of the relationships. Due to this, although the forced constant regression model is more valuable for each of the three relationships, the results of the standard regression are much more intriguing, albeit much more volatile.

After running the regressions for the thirty MLB teams and the MLB on the whole, linear equations were established for each team to predict the expected annual attendance for each team and the MLB using the randomly generated Moneyball lineups. The Moneyball generated expected annual attendance for each MLB team was held constant for every year and compared to the historic value of each team's annual attendance in that year. If the Moneyball expected attendance was greater than the historic attendance in that year, it was considered to be a "beat." This infers that the Moneyball lineup based off the Z-score / salary ratio would have provided greater attendance in that year utilizing the player valuation model. The table demonstrates the summation of the results over eighteen years for each of the thirty MLB teams. The percentage beats column represents the number of times that the Moneyball attendance value exceeded the historic value for each MLB team out of eighteen seasons. As is seen in Table 2, when using the standard linear regression, twenty out of thirty teams had a beat percentage greater than fifty percent using Moneyball player valuation techniques. When averaging all thirty teams'

Moneyball annual attendance figures to represent the MLB on the whole, the MLB's Moneyball team outperformed its historic annual attendance figures ninety-four percent of the time.

However, please note that the stronger forced constant regression model had a consistent zero percent beat performance for all thirty MLB teams and the MLB on the whole.

Table 2: Annual Attendance Moneyball Performance (in millions)

	Std. Moneyball Factor	% Beats	Forced Factor	% Beats
MLB	2.544633	94%	-1.296977	0%
ARI	2.045728	0%	-2.576202	0%
ATL	2.482260	33%	-2.152940	0%
BAL	2.513789	56%	-0.451180	0%
BOS	2.940704	61%	-0.517004	0%
CHC	3.101884	67%	-0.673495	0%
CHW	2.313398	78%	-0.541007	0%
CIN	2.272993	67%	-0.371434	0%
CLE	2.388628	67%	-0.159557	0%
COL	2.735368	50%	-0.627887	0%
DET	2.730430	72%	-0.332677	0%
HOU	2.878685	72%	-0.528758	0%
KCR	1.679626	67%	-0.437396	0%
ANA	2.935382	39%	-0.384735	0%
LAD	3.468874	50%	-0.787522	0%
FLA	1.587435	67%	-0.645542	0%
MIL	2.934317	83%	-0.648456	0%
MIN	2.231707	50%	-0.088032	0%
NYM	2.747876	61%	-0.426660	0%
NYY	3.637748	56%	-0.316139	0%
OAK	2.098521	72%	-0.350443	0%
PHI	3.163274	67%	-0.394292	0%
PIT	2.019486	78%	-0.253469	0%
SDP	2.433457	67%	-0.435646	0%
SFG	3.034113	28%	-0.451542	0%
SEA	2.902454	67%	-1.485282	0%
STL	3.253952	44%	-2.324204	0%
TBD	1.418444	41%	-0.320267	0%
TEX	2.589949	50%	-1.350596	0%
TOR	2.005559	50%	-0.546730	0%
WSN	3.292952	100%	-1.198843	0%

Considering that both regressions return significance F factors below five percent, the analysis concurs that the standard regression implies that the newly created player valuation model demonstrates a positive relationship between the Z-score / salary ratio and the annual attendance. However, due to the small R^2 figure, the analysis can not conclude with certainty that the model has accurate predictive power due to the inconsistent correlation.

The forced constant regression demonstrates a better fit for the player valuation model, but fails to imply that there is any type of relationship between the Z-score / salary ratio.

Z-Score / Salary Ratio And Total Annual Revenue

Regression Results

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.584737
R Square	0.341917
Adjusted R Sq	0.300787
Standard Error	45.01531
Observations	18

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	16845.37	16845.37	8.313041	0.010810526
Residual	16	32422.05	2026.378		
Total	17	49267.42			

	Coefficients	*Std Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	212.7206	21.81699	9.750229	3.9E-08	166.4706717
Z/Salary	279.6587	96.99479	2.883234	0.010811	74.03895215

Figure 4: Standard Regression MLB-wide

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.739846
R Square	0.547373
Adjusted R Sq	0.488549
Standard Error	115.061
Observations	18

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	272174.8165	272174.8	20.5585	0.000338773
Residual	17	225063.6988	13239.04		
Total	18	497238.5153			

	Coefficients	*Std Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	0	#N/A	#N/A	#N/A	#N/A
Z/Salary	-546.69	120.571784	-4.53415	0.000294	-801.074557

Figure 5: Forced Constant Regression MLB-wide

Similarly to the annual fan attendance regression results, the forced constant regression of the annual revenues return a much more powerful significant F factor as is seen in Figure 5. The forced constant regression R^2 result implies that the model explains roughly fifty-five percent of the variance in the annual revenue results are accounted for. As for the standard regression, the significance F factor is still very impressive, returning a one percent significance F factor. Yet, the R^2 analysis explains only thirty-four percent of the variance. However, both regressions return a significance value lower than five percent, and thus indicate the fit of the model.

As established in the annual attendance regressions, the annual revenue forced constant regressions always return a negative expected Moneyball revenue figure based off the new player valuation model. While the forced constant regression has greater model fit and predictive power, the impossibility of having negative revenue figures will lead to insignificant results for the forced constant regression model.

When analyzing the standard linear regression, the player valuation model appears to have some significance. The revenue figures are listed in millions of U.S. dollars. The calculations and analysis were conducted in the same manner as the annual attendance figures in the section above but simply changes the figure from attendance to revenue. As is seen in Table 3, when using the standard linear regression, twenty-six out of thirty teams had a beat percentage greater than fifty percent using Moneyball player valuation techniques. When averaging all thirty teams' Moneyball annual revenue figures to represent the MLB on the whole, the MLB's Moneyball team outperformed its historic annual revenue figures ninety-four percent of the time.

As is noted, the forced constant regression returned a consistent zero percent beat rate when analyzing the expected Moneyball annual revenue figures.

Table 3: Annual Revenue Moneyball Performance ($ in millions)

	Std. Moneyball Factor	% Beats	Forced Factor	% Beats
MLB	253.110	94%	-78.955	0%
ARI	177.842	71%	-132.038	0%
ATL	190.821	67%	-139.910	0%
BAL	163.743	56%	-26.001	0%
BOS	249.938	56%	-37.584	0%
CHC	199.510	56%	-42.949	0%
CHW	187.743	56%	-38.772	0%
CIN	153.114	56%	-23.135	0%
CLE	163.139	56%	-12.981	0%
COL	172.838	61%	0.000	0%
DET	183.691	61%	-22.247	0%
HOU	143.380	39%	-43.296	0%
KCR	155.672	72%	-30.096	0%
ANA	198.779	56%	-5.774	0%
LAD	211.871	56%	-43.528	0%
FLA	138.730	61%	-47.199	0%
MIL	186.761	78%	-31.876	0%
MIN	141.698	56%	-5.149	0%
NYM	204.193	50%	-30.822	0%
NYY	345.493	61%	-8.053	0%
OAK	143.239	50%	-26.667	0%
PHI	215.148	61%	-22.473	0%
PIT	184.075	89%	-13.238	0%
SDP	149.718	39%	-24.621	0%
SFG	192.511	56%	-26.935	0%
SEA	211.146	89%	-97.823	0%
STL	206.084	72%	-115.178	0%
TBD	163.893	76%	-19.932	0%
TEX	202.402	72%	-69.588	0%
TOR	132.127	44%	-35.885	0%
WSN	337.171	100%	-91.308	0%

Considering that both regressions return significance F factors below five percent, the analysis once again concurs that the standard regression implies that the newly created player valuation model demonstrates a positive relationship between the Z-score / salary ratio and the annual revenue. However, due to the small R^2 figure, the analysis does not conclude with certainty that the model has accurate predictive power due to the inconsistent correlation.

The forced constant regression analysis demonstrates that there is no positive relationship between the Z-score / salary ratio and annual revenue. A consistent prediction of negative revenues shows that the regression and player valuation model fails to provide significant results.

Z-Score / Salary Ratio And Annual Winning Percentage

Regression Results

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.314445
R Square	0.098876
Adjusted R Sq	0.042555
Standard Error	0.004692
Observations	18

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.87E-05	3.87E-05	1.755597	0.203786811
Residual	16	0.000352	2.2E-05		
Total	17	0.000391			

	Coefficients	*Std Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	0.501411	0.002274	220.4729	2.7E-29	0.496589474
Z/Salary	0.013397	0.010111	1.324989	0.203787	-0.00803736

Figure 6: Standard Regression MLB-wide

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.872309
R Square	0.760924
Adjusted R Sq	0.7021
Standard Error	0.250961
Observations	18

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	3.407732	3.407732	54.10698	1.61931E-06
Residual	17	1.070683	0.062981		
Total	18	4.478416			

	Coefficients	*Std Error*	*t Stat*	*P-value*	*Lower 95%*
Intercept	0	#N/A	#N/A	#N/A	#N/A
Z/Salary	-1.93442	0.26298	-7.35574	1.12E-06	-2.4892576

Figure 7: Forced Constant Regression MLB-wide

The results for the relationship between the Z-score / salary ratio and the annual winning percentage are mixed relative to the prior two relationships. Once again, the forced constant regression maintains an impressive significance F factor in that is much below the five percent required figure as you can see in Figure 7. The R^2 results for the forced constant demonstrate that the model explains seventy-six percent of the variance in the winning percentage calculations. While this is a rather impressive figure, the regressions continue to produce negative expected Moneyball winning percentages which is impossible once again. Therefore, the forced constant regression holds more significance and power than the standard linear regression.

The standard linear regression for the annual winning percentage figures produces a significance F of twenty percent, as seen in Figure 6, well above the five percent implying statistical significance of the model. Additionally, the R^2 value is a mere ten percent which shows the lack of value in the standard linear regression for the annual winning percentage. While the results of the standard regression are essentially meaningless, it is interesting to note that twenty-four out of thirty teams registered a beat percentage of greater than fifty percent and the MLB on average earned a beat percentage of seventy-two percent using the Moneyball player valuation model.

Table 4: Annual Winning Percentage Moneyball Performance

	Std. Moneyball Factor	% Beats	Forced Factor	% Beats
MLB	50%	72%	-28%	0%
ARI	48%	41%	-46%	0%
ATL	58%	50%	-44%	0%
BAL	47%	61%	-8%	0%
BOS	57%	50%	-10%	0%
CHC	51%	61%	-10%	0%
CHW	54%	72%	-15%	0%
CIN	52%	72%	-9%	0%
CLE	52%	50%	-4%	0%
COL	48%	72%	-10%	0%
DET	52%	61%	-8%	0%
HOU	57%	83%	-11%	0%
KCR	45%	72%	-12%	0%
ANA	56%	61%	-7%	0%
LAD	55%	72%	-12%	0%
FLA	52%	83%	-19%	0%
MIL	54%	89%	-14%	0%
MIN	51%	50%	-2%	0%
NYM	50%	50%	-8%	0%
NYY	60%	56%	-6%	0%
OAK	59%	78%	-11%	0%
PHI	56%	78%	-9%	0%
PIT	45%	67%	-7%	0%
SDP	51%	72%	-9%	0%
SFG	56%	67%	-8%	0%
SEA	55%	72%	-29%	0%
STL	60%	89%	-38%	0%
TBD	51%	65%	-8%	0%
TEX	54%	67%	-24%	0%
TOR	49%	33%	-13%	0%
WSN	64%	100%	-37%	0%

Due to the high significance F factor of the standard linear regression, its results pertaining to the annual winning percentage are void and it does not imply any relationship between the Z-score / salary ratio and annual winning percentage, let alone a positive relationship.

Although the forced constant regression returns an impressive significance F factor, it fails to predict plausible annual winning percentages. For this reason, the regression results imply that there is no relationship between the Z-score / salary ratio and the annual winning percentage.

Chapter 5

Conclusion

Moneyball's Impact on Baseball Franchises

The success of the player valuation model utilized to generate Moneyball inspired lineups for each MLB team has mixed results. It is difficult to truly analyze the value of the player valuation model without sensitizing the variables used to establish the player Z-Scores. However, the results of the regression analyses speak largely for themselves.

After running all of the standard linear and forced constant regressions for all thirty MLB teams and the MLB as a league on the whole, it appears that in all aspects, the forced constant regression model exhibits greater strength and power of the model as dictated by its significance F factors all far below five percent and far below those of the standard linear regressions. In addition to the significance F factors, all of the forced constant regression models explain the variance of the annual attendance, revenue and winning percentage variables due to the Z-score / salary ratio as indicated by the R^2 results for each regression. Yet, with all of the impressive statistical results, none of the forced constant regressions produce a single projected Moneyball attendance, revenue or winning percentage that is greater than zero. Clearly, these results are impossible as the minimum for attendance, revenue and winning percentage is zero.

For the annual attendance and revenue regressions, the standard linear analysis demonstrates statistical significance based on the significance F factors below five percent. Additionally, all three models consistently predict positive results for attendance, revenue and winning percentages for all thirty MLB teams and the MLB on the whole. The value of these results are somewhat muted by the fact that the standard linear regressions fail to explain more than forty percent of the variability as measured by the R^2 figures.

While the standard linear regression appears to be less consistent for all three measures, the attendance and revenue analyses still should be taken as statistically significant, but used with limited confidence. The player valuation model requires additional tuning and greater granularity in variable selection and weightings to further the predictive power and consistency of the projections of player value. The forced constant regressions should be understood to have statistical significance but require further modification to hold any value in enhancing franchise profitability.

The analysis of the standard linear regressions serves to reject the null hypothesis for the relationships between the Z-score / salary ratio and each of attendance and revenue and to accept that both of these relationships have a positive correlation. The results imply that the player valuation model's Moneyball defined players would more consistently provide MLB teams with enhanced revenue and attendance results. The standard linear regressions also serve to accept the null hypothesis that there is no positive relationship between the Z-score / salary ratio and the annual winning percentage.

The analysis of the forced constant regressions serve to accept the null hypothesis for all three relationships that the Z-score / salary ratio has no positive relationship between any of the annual attendance, annual revenue or annual winning percentage figures.

Compiling both regressions, the analysis concurs that there is no relationship between the Z-score / salary ratio and the annual winning percentage. Yet, with limited success, there is a positive relationship between the Z-score / salary ratio and the annual attendance and revenue figures.

The current player valuation model can be used to establish an enhanced valuation of every current MLB player in terms of cost efficiencies for the franchise. Assuming these players are available, teams can implement the created player valuation model to identify players with positive Z-score / salary ratios that, as a starting lineup, is greater than current rosters' Z-score / salary ratio to improve overall attendance and revenue.

Topics to Explore Further

While this analysis has provided an introductory method for valuing players differently than current literature suggests, there are a number of topics that would enhance the accuracy of the study.

What are the factors and their respective weightings in an accurate revenue projection model?

The analysis provided in this paper utilized historical revenue figures to predict the potential revenue generation of randomly generated portfolios. To better test the potential relationship between total revenue and total attendance or total revenue and Z-score / salary ratio, the revenue projections need to be more comprehensive. To begin with, instead of attempting to predict total revenue, there are a number of categories that should be independently projected based off a number of factors rather than just attendance. For example, total revenue should be comprised of gate receipts, concessions revenue, sponsorship, media rights, parking and other miscellaneous franchise revenues. With greater detail into individual team projections for future attendance, inflation in stadium pricing, promotions and other pricing surcharges could enhance the projection of actual revenue generation.

Greater detail relating particularly to fan demographics would also enhance the value of the analysis. Higher margin seats, such as luxury boxes draw significantly more revenues for franchises. Understanding the wealth distribution of the surrounding fan base and the likelihood of attracting higher wealth fans based on team and player popularity would provide greater clarity into the revenue projection model as well.

What is the proper way to value a Major League Baseball franchise?

The lack of public information on precedent transactions of sports franchises due to their private transaction nature has made it difficult to provide a fair valuation method for Major League Baseball teams. With better access to private documents and financial statements, analysts would enhance their analysis in several ways. First, a better grouping of comparable precedent transactions could be established to provide reference for how projections on the revenue model impact the true value of the team. Additionally, it would allow the analysts to provide discounted cash flow analyses on firms that have fairly predictable cash flows. With public data, it is nearly impossible to ascertain a breakdown of operating costs for these franchises which makes the valuation very difficult on any other metric than a revenue based multiple.

What is the inherent covariance between Z-score / salary ratio and total revenue?

In this analysis, the revenue projections are based off attendance figures that are derived from Z-score / salary ratios. Therefore, the relationships that were tested already have some hidden covariance that the analysis will not identify. A further analysis that isolates these factors would benefit the area of study.

What are the proper Z-score weightings and variables?

This analysis assumes a limited number of variables when calculating the average Z-score per player. It is possible that further research and direct contact with current scouts and general managers would show that there is a different set of variables that provide better insight into the offensive efficiency of each player. By minimizing the number of variables, the more direct the potential relationships may be. Additionally, the weightings in this analysis were essentially arbitrary. Assuming that the Z-score / salary ratio is an appropriate efficiency measure, the changes in the weightings of the variables utilized could provide a nearly endless number of further research opportunities. If possible, a method for determining optimal weightings that are customizable, or independent, for each franchise would be helpful to both the academic and baseball community. To further the analysis of this research, the current variables and a different set of variables should be sensitized in order to test the impact each variable has on the overall player valuation.

How is player popularity measured? How is player popularity related to total revenue?

At the start of this analysis, the number of all-star game selections was proposed as a method for player popularity. However, over the course of the study it was recognized that there was an historic change in the method of selecting all-stars in the MLB from player voting to fan voting. In this change, the selections do not necessarily reflect talent and contributions to a team performance, but rather a player's "likeability". Yet, it has been difficult to ascertain whether an increase in likeability drives increased fan attendance, memorabilia sales, concession sales and so forth. A better metric that incorporates particular changes in ancillary revenues based off fan popularity would be beneficial.

What is the value of a world championship when selling a MLB franchise?

The number of World Series victories are very unevenly distributed throughout existing franchises with the Yankees and the Red Sox leading significantly. Yet, these franchises are not multiples ahead of other franchises that have had less World Series success. How much is a single championship worth? A string of successful playoff series? A dynasty? In essence, these championships would likely be valued the same as goodwill in any acquisition scenario, but these intangible values are valued differently in every market depending on the size of the market.

What is the incremental change in franchise profitability from utilizing Moneyball generated rosters?

One of the shortcomings of this analysis is that it tests relationships pertaining to revenue, but fails to test the change in profit for the organizations. Although major sports franchises are typically valued on revenue based multiples, the annual net income should be considered for franchise owners that intend to rely on the franchise for a consistent source of personal cash flow. To do so, the average career earnings for each player in the Moneyball lineup should be calculated and compared to that of the historical lineups. In this way, the player valuation model could be demonstrated to not only have a relationship with revenue, but also potentially could have a relationship with profit as well.

Is linear regression the proper analysis to perform?

The analysis assumes that either the standard linear regression of the forced constant linear regression are the only two regression analyses to consider. However, there are a number of different regressions that could potentially demonstrate a better fit for the model and lend greater predictive power. The regressions performed in this analysis demonstrate mixed results and it is possible that other regressions would be more meaningful.

How to properly simulate player availability?

This research assumes that all MLB players are readily available to sign from a general free agent pool. The free agency period is simulated by randomizing which players are assigned to which Moneyball rosters. In doing so, teams are not able to simply select the players with the highest Z-score / salary ratios but are limited by market free agent forces on availability. The model is similar to a new franchise entering the league which forces every team to make several players available for the new team to draft.

While the current method is not perfect, it attempts to simulate free agency. It is possible that there is a better method for simulating free agency and the actual availability of the highly qualified Moneyball players. The operating assumption in this research is that all franchises discover the player valuation model at the same time and implement the methods at the same time. However, in reality this is highly unlikely as specific franchises tend to be more analytically progressive than others.

Appendix A

Annual Team Revenues Summary 1997-2014 ($ in millions)

Franchise	1997	1998	1999	2000	2001	2002
ARZ	0	116	112	109	127	127
ATL	120	143	155	146	160	160
BAL	135	141	139	124	133	133
BOS	92	107	123	126	152	152
CHC	82	93	105	112	131	131
CHW	82	74	74	93	101	101
CIN	50	54	49	78	87	87
CLE	134	150	152	143	150	150
COL	117	125	122	119	129	129
DET	51	54	72	121	114	114
HOU	68	83	93	122	125	125
KC	51	54	63	73	85	85
LAA	63	89	91	94	103	103
LAD	94	108	120	131	143	143
MIA	88	70	59	67	81	81
MIL	47	56	61	70	108	108
MIN	47	47	48	58	75	75
NYM	81	100	126	162	169	169
NYY	145	176	196	192	215	215
OAK	56	57	61	75	90	90
PHI	57	66	68	79	94	94
PIT	49	52	58	70	108	108
SD	58	79	79	84	92	92
SF	70	73	72	139	142	142
SEA	90	81	112	138	166	166
STL	83	98	105	111	123	123
TB	0	94	78	81	92	92
TEX	98	108	118	127	134	134
TOR	67	73	74	80	91	91
WAS	44	47	47	54	63	63

Franchise	2003	2004	2005	2006	2007	2008
ARZ	126	136	145	154	165	177
ATL	156	162	172	183	199	186
BAL	129	148	156	158	166	174
BOS	190	201	206	234	263	269
CHC	156	170	179	197	214	239
CHW	124	131	157	173	193	196
CIN	123	127	137	146	161	171
CLE	127	139	150	158	181	181
COL	124	132	145	151	169	178
DET	117	126	146	170	173	186
HOU	128	155	173	184	193	194
KC	98	104	117	123	131	143
LAA	127	147	167	187	200	212
LAD	154	166	189	211	224	241
MIA	101	103	119	122	128	139
MIL	102	112	131	144	158	173
MIN	99	102	114	131	149	158
NYM	158	180	195	217	235	261
NYY	238	264	277	302	327	375
OAK	110	116	134	146	154	160
PHI	115	167	176	183	192	216
PIT	109	109	125	137	139	144
SD	106	150	158	160	167	174
SF	153	159	171	184	197	196
SEA	169	173	179	182	194	189
STL	131	151	165	184	194	195
TB	101	110	116	134	138	160
TEX	127	142	153	155	172	176
TOR	99	107	136	157	160	172
WAS	81	80	145	144	153	184

Franchise	2009	2010	2011	2012	2013	2014	Average
ARZ	172	180	186	195	192	211	146
ATL	188	201	203	225	253	267	182
BAL	171	175	179	206	198	245	162
BOS	266	272	310	336	357	370	224
CHC	246	258	266	274	266	302	190
CHW	194	210	214	216	210	227	154
CIN	166	179	185	202	209	227	135
CLE	170	168	178	186	196	207	162
COL	183	188	193	199	197	214	156
DET	188	192	217	238	262	254	155
HOU	189	197	196	196	186	175	155
KC	155	160	161	169	178	231	121
LAA	217	222	226	239	253	304	169
LAD	247	246	230	245	293	403	199
MIA	144	143	148	195	159	188	119
MIL	171	179	195	201	197	226	135
MIN	162	213	213	214	221	223	130
NYM	268	233	225	232	238	263	195
NYY	441	427	439	471	461	508	315
OAK	155	161	160	173	187	202	127
PHI	233	239	249	279	265	265	169
PIT	145	160	168	178	204	229	127
SD	157	159	163	189	207	224	139
SF	201	230	230	262	316	387	185
SEA	191	204	210	215	210	250	173
STL	195	207	233	239	283	294	173
TB	156	166	161	167	181	188	123
TEX	180	206	233	239	257	266	168
TOR	163	168	188	203	218	227	137
WAS	184	194	200	225	244	287	135

Appendix B

Annual Attendance 1997-2014 (in millions)

	ARI	ATL	BAL	BOS	CHC	CHW	CIN	CLE
1997	0	3.464488	3.711132	2.226136	2.580325	1.864782	1.785788	3.40475
1998	3.61029	3.36086	3.68465	2.314704	2.955193	1.391146	1.793649	3.467299
1999	3.015948	3.284901	3.432099	2.446277	3.293659	1.349151	2.061324	3.468436
2000	2.942517	3.229082	3.296031	2.585895	2.789511	1.947799	2.599318	3.456278
2001	2.736451	2.82353	3.094841	2.625333	2.779465	1.766172	1.879757	3.175523
2002	3.198977	2.603484	2.682439	2.650862	2.693096	1.676911	1.855787	2.61694
2003	2.805542	2.401084	2.454523	2.724165	2.96263	1.939524	2.355259	1.730002
2004	2.51956	2.322565	2.744013	2.837304	3.170184	1.930537	2.28725	1.814401
2005	2.059331	2.521534	2.624804	2.813354	3.100262	2.342834	1.943157	1.973185
2006	2.091005	2.549524	2.15315	2.930768	3.123295	2.957414	2.135417	1.99807
2007	2.316507	2.745203	2.164822	2.970755	3.252462	2.684395	2.058593	2.275911
2008	2.509924	2.532834	1.950075	3.04825	3.29984	2.460749	2.058632	2.16876
2009	2.129183	2.373631	1.907163	3.062699	3.168859	2.284164	1.74792	1.766242
2010	2.056519	2.510119	1.733018	3.046445	3.062973	2.194378	2.060551	1.391644
2011	2.105432	2.37294	1.755461	3.054001	3.017966	2.001117	2.213498	1.840835
2012	2.177617	2.420171	2.10224	3.043003	2.882756	1.965955	2.347251	1.603599
2013	2.134795	2.548679	2.357561	2.833333	2.642682	1.768413	2.492059	1.572926
2014	2.07373	2.354305	2.464473	2.956089	2.652113	1.650821	2.476664	1.437393

	COL	DET	HOU	KCR	ANA	LAD	FLA	MIL
1997	3.888453	1.365157	2.046781	1.517638	1.76733	3.319504	2.364387	1.444027
1998	3.792683	1.409391	2.458451	1.494875	2.51928	3.089222	1.730384	1.811593
1999	3.235833	2.026491	2.706017	1.506068	2.25304	3.098042	1.36942	1.70179
2000	3.285711	2.533752	3.056139	1.677915	2.066977	3.010765	1.218326	1.573621
2001	3.166821	1.921305	2.904277	1.536371	2.000919	3.017143	1.261226	2.811041
2002	2.737838	1.503623	2.517357	1.323036	2.305547	3.131255	0.813118	1.969153
2003	2.334085	1.368245	2.454241	1.779895	3.061094	3.138626	1.303215	1.700354
2004	2.338069	1.917004	3.087872	1.661478	3.375677	3.488283	1.723105	2.062382
2005	1.915586	2.024505	0.272472	1.371181	3.404686	3.60368	1.823388	2.211023
2006	2.104558	2.595937	3.022763	1.372694	3.40679	3.758421	1.16512	2.335643
2007	2.327846	3.047124	3.020405	1.616687	3.365632	3.856753	1.370511	2.869144
2008	2.650218	3.202654	2.779487	1.578922	3.336747	3.730553	1.335086	3.068638
2009	2.66508	2.567185	2.521076	1.797887	3.240386	3.761669	1.464109	3.037451
2010	2.875245	2.461237	2.33149	1.615327	3.250814	3.56232	1.524894	2.776531
2011	2.909777	2.642045	2.067016	1.72445	3.166321	2.935139	1.520562	3.071373
2012	2.630458	3.028033	1.607733	1.739859	2.177617	3.324246	2.219444	2.831385
2013	2.793828	3.083397	1.651883	1.750754	3.019505	3.743527	1.586322	2.531105
2014	2.680329	2.917209	1.751829	1.956482	3.095935	3.782337	1.732283	2.797384

	MIN	NYM	NYY	OAK	PHI	PIT	SDP	SFG
1997	1.411064	1.766174	2.580325	1.264218	1.490638	1.657022	2.089333	1.690869
1998	1.165976	2.287948	2.955193	1.232343	1.715722	1.56095	2.555874	1.925364
1999	1.202829	2.726008	3.293659	1.434632	1.825337	1.638023	2.532538	2.078365
2000	1.059715	2.775661	3.227657	1.728888	1.612769	1.709119	2.423149	3.31533
2001	1.782926	2.65833	3.265907	2.154496	1.782054	2.428661	2.378128	3.311958
2002	1.924473	2.804838	3.465807	2.169811	1.618467	1.784988	2.220601	3.253203
2003	1.946011	2.140599	3.4656	2.216596	2.259948	1.636751	2.030084	3.264898
2004	1.879222	2.318321	3.775292	2.201516	3.206532	1.583031	3.040046	3.258864
2005	2.013453	2.782212	4.09044	2.109298	2.665301	1.794237	2.832039	3.140781
2006	2.285018	3.342449	4.200468	1.976625	2.701815	1.861549	2.659732	3.130024
2007	2.296383	3.853955	4.271967	1.921854	3.108331	1.749142	2.790074	3.223202
2008	2.302611	4.047404	4.298655	1.663262	3.422583	1.609076	2.427535	2.863837
2009	2.416237	3.154262	3.719358	1.392192	3.600693	1.577853	1.922603	2.861113
2010	3.22364	2.559738	3.765807	1.418391	3.647249	1.613399	2.131774	3.037443
2011	3.168107	2.378549	3.65368	1.476792	3.680718	1.940429	2.143018	3.387303
2012	2.776354	2.219444	3.542406	1.679013	3.565718	2.091918	2.123721	3.377371
2013	2.477644	2.135657	3.279589	1.809302	3.012403	2.256862	2.166691	3.369106
2014	2.250606	2.148808	3.401624	2.003628	2.423852	2.442564	2.195373	3.368697

	SEA	STL	TBD	TEX	TOR	WSN
1997	3.192237	2.364387	0	2.945228	2.589297	1.497609
1998	2.651511	3.195691	2.506293	2.927399	2.454303	0.914909
1999	1.457954	3.235833	1.749657	2.774501	2.163486	0.772737
2000	3.148317	3.336493	1.54944	2.800147	1.819885	0.926213
2001	3.507326	3.109578	1.298365	2.831021	1.915438	0.642745
2002	3.542938	3.011756	1.065742	2.352397	1.6379	0.812045
2003	3.268509	2.910386	1.058695	2.094394	1.799458	1.025639
2004	2.940731	3.048427	1.275011	2.513685	1.900041	0.74855
2005	2.689529	3.491837	1.124189	2.486925	1.977949	2.692123
2006	2.480717	3.407104	1.369031	2.388734	2.302182	2.15315
2007	2.672409	3.55215	1.389106	2.353862	2.360648	1.961579
2008	2.329702	3.430403	1.811986	1.945677	2.399786	2.3204
2009	2.195284	3.343252	1.874962	2.156016	1.876129	1.81728
2010	2.195284	3.301218	1.864999	2.505171	1.625555	1.828066
2011	1.896321	3.093954	1.529188	2.946949	1.818103	1.940478
2012	1.72192	3.262109	1.559681	3.46028	2.099663	2.370794
2013	1.761546	3.369769	1.5103	3.178273	2.536562	2.652422
2014	2.064334	3.540649	1.446464	2.718733	2.375525	2.579389

Appendix C

Annual Winning Percentages 1997-2014

	ARI	ATL	BAL	BOS	CHC	CHW	CIN	CLE
1997	0.00%	62.35%	60.49%	48.15%	41.98%	49.38%	46.91%	53.09%
1998	39.88%	65.03%	48.47%	56.44%	55.21%	49.08%	47.24%	54.60%
1999	61.35%	63.19%	47.85%	57.67%	41.10%	46.01%	58.90%	59.51%
2000	52.47%	58.64%	45.68%	52.47%	40.12%	58.64%	52.47%	55.56%
2001	56.79%	54.32%	38.89%	50.62%	54.32%	51.23%	40.74%	56.17%
2002	60.49%	62.35%	41.36%	57.41%	41.36%	50.00%	48.15%	45.68%
2003	51.85%	62.35%	43.83%	58.64%	54.32%	53.09%	42.59%	41.98%
2004	31.48%	59.26%	48.15%	60.49%	54.94%	51.23%	46.91%	49.38%
2005	47.53%	55.56%	45.68%	58.64%	48.77%	61.11%	45.06%	57.41%
2006	46.91%	48.77%	43.21%	53.09%	40.74%	55.56%	49.38%	48.15%
2007	55.21%	51.53%	42.33%	58.90%	52.15%	44.17%	44.17%	58.90%
2008	50.31%	44.17%	41.72%	58.28%	59.51%	54.60%	45.40%	49.69%
2009	42.94%	52.76%	39.26%	58.28%	50.92%	48.47%	47.85%	39.88%
2010	40.12%	56.17%	40.74%	54.94%	46.30%	54.32%	56.17%	42.59%
2011	58.02%	54.94%	42.59%	55.56%	43.83%	48.77%	48.77%	49.38%
2012	50.00%	58.02%	57.41%	42.59%	37.65%	52.47%	59.88%	41.98%
2013	49.69%	58.90%	52.15%	59.51%	40.49%	38.65%	55.21%	56.44%
2014	39.51%	48.77%	59.26%	43.83%	45.06%	45.06%	46.91%	52.47%

	COL	DET	HOU	KCR	ANA	LAD	FLA	MIL
1997	51.23%	48.77%	51.85%	41.36%	51.85%	54.32%	56.79%	48.15%
1998	47.24%	39.88%	62.58%	44.17%	52.15%	50.92%	33.13%	45.40%
1999	44.17%	42.33%	59.51%	39.26%	42.94%	47.24%	39.26%	45.40%
2000	50.62%	48.77%	44.44%	47.53%	50.62%	53.09%	48.77%	45.06%
2001	45.06%	40.74%	57.41%	40.12%	46.30%	53.09%	46.91%	41.98%
2002	45.06%	33.95%	51.85%	38.27%	61.11%	56.79%	48.77%	34.57%
2003	45.68%	26.54%	53.70%	51.23%	47.53%	52.47%	56.17%	41.98%
2004	41.98%	44.44%	56.79%	35.80%	56.79%	57.41%	51.23%	41.36%
2005	41.36%	43.83%	54.94%	34.57%	58.64%	43.83%	51.23%	50.00%
2006	46.91%	58.64%	50.62%	38.27%	54.94%	54.32%	48.15%	46.30%
2007	55.21%	53.99%	44.79%	42.33%	57.67%	50.31%	43.56%	50.92%
2008	45.40%	45.40%	52.76%	46.01%	61.35%	51.53%	51.53%	55.21%
2009	56.44%	52.76%	45.40%	39.88%	59.51%	58.28%	53.37%	49.08%
2010	51.23%	50.00%	46.91%	41.36%	49.38%	49.38%	49.38%	47.53%
2011	45.06%	58.64%	34.57%	43.83%	53.09%	50.62%	44.44%	59.26%
2012	39.51%	54.32%	33.95%	44.44%	54.94%	53.09%	42.59%	51.23%
2013	45.40%	57.06%	31.29%	52.76%	47.85%	56.44%	38.04%	45.40%
2014	40.74%	55.56%	43.21%	54.94%	60.49%	58.02%	47.53%	50.62%

	MIN	NYM	NYY	OAK	PHI	PIT	SDP	SFG
1997	41.98%	54.32%	59.26%	40.12%	41.98%	48.77%	46.91%	55.56%
1998	42.94%	53.99%	69.94%	45.40%	46.01%	42.33%	60.12%	54.60%
1999	38.65%	59.51%	60.12%	53.37%	47.24%	47.85%	45.40%	52.76%
2000	42.59%	58.02%	53.70%	56.17%	40.12%	42.59%	46.91%	59.88%
2001	52.47%	50.62%	58.64%	62.96%	53.09%	38.27%	48.77%	55.56%
2002	58.02%	46.30%	63.58%	63.58%	49.38%	44.44%	40.74%	58.64%
2003	55.56%	40.74%	62.35%	59.26%	53.09%	46.30%	39.51%	61.73%
2004	56.79%	43.83%	62.35%	56.17%	53.09%	44.44%	53.70%	56.17%
2005	51.23%	51.23%	58.64%	54.32%	54.32%	41.36%	50.62%	46.30%
2006	59.26%	59.88%	59.88%	57.41%	52.47%	41.36%	54.32%	46.91%
2007	48.47%	53.99%	57.67%	46.63%	54.60%	41.72%	54.60%	43.56%
2008	53.99%	54.60%	54.60%	46.01%	56.44%	41.10%	38.65%	44.17%
2009	53.37%	42.94%	63.19%	46.01%	57.06%	38.04%	46.01%	53.99%
2010	58.02%	48.77%	58.64%	50.00%	59.88%	35.19%	55.56%	56.79%
2011	38.89%	47.53%	59.88%	45.68%	62.96%	44.44%	43.83%	53.09%
2012	40.74%	45.68%	58.64%	58.02%	50.00%	48.77%	46.91%	58.02%
2013	40.49%	45.40%	52.15%	58.90%	44.79%	57.67%	46.63%	46.63%
2014	43.21%	48.77%	51.85%	54.32%	45.06%	54.32%	47.53%	54.32%

	SEA	STL	TBD	TEX	TOR	WSN
1997	55.56%	45.06%	0.00%	47.53%	46.91%	48.15%
1998	46.63%	50.92%	38.65%	53.99%	53.99%	39.88%
1999	48.47%	46.01%	42.33%	58.28%	51.53%	41.72%
2000	56.17%	58.64%	42.59%	43.83%	51.23%	41.36%
2001	71.60%	57.41%	38.27%	45.06%	49.38%	41.98%
2002	57.41%	59.88%	33.95%	44.44%	48.15%	51.23%
2003	57.41%	52.47%	38.89%	43.83%	53.09%	51.23%
2004	38.89%	64.81%	43.21%	54.94%	41.36%	41.36%
2005	42.59%	61.73%	41.36%	48.77%	49.38%	50.00%
2006	48.15%	51.23%	37.65%	49.38%	53.70%	43.83%
2007	53.99%	47.85%	40.49%	46.01%	50.92%	44.79%
2008	37.42%	52.76%	59.51%	48.47%	52.76%	36.20%
2009	52.15%	55.83%	51.53%	53.37%	46.01%	36.20%
2010	37.65%	53.09%	59.26%	55.56%	52.47%	42.59%
2011	41.36%	55.56%	56.17%	59.26%	50.00%	49.38%
2012	46.30%	54.32%	55.56%	57.41%	45.06%	60.49%
2013	43.56%	59.51%	56.44%	55.83%	45.40%	52.76%
2014	53.70%	55.56%	47.53%	41.36%	51.23%	59.26%

Appendix D

Moneyball Generated Rosters

LAA	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Meluskey	Hillenbrar	Alomar	Lugo	Peralta	Burks	Griffey	De Aza	Burks		
Z-score	0.583	0.147	0.547	0.031	0.173	0.277	0.463	0.071	0.277	0.286	0.322
Z-score/salary	2.285	0.061	0.290	0.006	0.038	0.057	0.052	0.053	0.057		

HOU	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Estrada	Garko	Spivey	Simmons	Gordon	Marte	Pierre	Fowler	Marte		
Z-score	0.014	0.053	0.370	0.079	0.155	0.148	0.215	0.030	0.148	0.135	0.141
Z-score/salary	0.010	0.117	0.339	0.106	0.043	0.297	0.049	0.012	0.297		

OAK	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Posada	Loney	Roberts	Cabrera	Blalock	Werth	Finley	Sheffield	Werth		
Z-score	0.832	0.084	0.498	0.254	0.797	0.196	0.199	0.699	0.196	0.417	0.080
Z-score/salary	0.106	0.032	0.054	0.073	0.291	0.032	0.041	0.062	0.032		

TOR	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Posey	Hatteberg	Kent	Ramirez	Randa	Mondesi	Alou	Lofton	Mondesi		
Z-score	1.342	0.040	0.852	0.937	0.074	0.233	0.767	0.051	0.233	0.503	0.083
Z-score/salary	0.248	0.035	0.123	0.119	0.033	0.032	0.118	0.010	0.032		

ATL	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Kendall	O'leary	Dozier	Renteria	Encarnacio	DeRosa	Drew	Hawpe	DeRosa		
Z-score	0.803	0.432	0.507	0.272	0.160	0.571	0.074	0.306	0.571	0.411	0.216
Z-score/salary	0.145	0.158	0.978	0.045	0.045	0.236	0.009	0.093	0.236		

MIL	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Lo Duca	Rizzo	Young	Aybar	Atkins	Baldelli	Gonzalez	Francoeur	Baldelli		
Z-score	0.798	0.052	0.301	0.164	0.693	0.233	0.266	0.016	0.233	0.306	0.133
Z-score/salary	0.237	0.052	0.040	0.046	0.252	0.252	0.063	0.005	0.252		

STL	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Santana	Gonzalez	Uggla	Lopez	Longoria	Pujols	Edmonds	Suzuki	Pujols		
Z-score	1.100	0.780	0.693	0.202	0.767	1.744	0.559	0.671	1.744	0.918	0.209
Z-score/salary	0.885	0.098	0.100	0.095	0.244	0.158	0.079	0.059	0.158		

CHC	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Rodriguez	Green	Johnson	Betancour	Nevin	Giles	Holliday	Brantley	Giles		
Z-score	0.841	0.468	0.277	0.271	1.208	0.688	0.916	0.387	0.688	0.638	0.135
Z-score/salary	0.110	0.052	0.104	0.138	0.307	0.110	0.082	0.203	0.110		

ARI	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Napoli	Carter	Valentin	Tulowitzk	Boone	Trout	Figgins	Willinghar	Trout		
Z-score	0.362	0.087	0.737	0.874	0.028	1.283	0.499	0.041	1.283	0.577	0.440
Z-score/salary	0.057	0.018	0.223	0.154	0.005	1.699	0.096	0.010	1.699		

LAD	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Martinez	Martinez	Hill	Rollins	Cantu	Green	Davis	Grandersc	Green		
Z-score	1.548	0.327	0.428	0.650	0.322	0.570	0.055	0.609	0.570	0.564	0.114
Z-score/salary	0.234	0.075	0.111	0.104	0.168	0.063	0.111	0.096	0.063		

SFG	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Lopez	Cantu	Baerga	Guillen	Iwamura	Harper	Jackson	Brown	Harper		
Z-score	0.683	0.395	0.005	0.530	0.176	0.049	0.245	0.007	0.049	0.238	0.075
Z-score/salary	0.113	0.205	0.003	0.100	0.057	0.038	0.113	0.006	0.038		

CLE	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Pierzynsk	Cabrera	Counsell	Peralta	Callaspo	Fukudome	Cruz	Wilson	Fukudome		
Z-score	0.502	1.155	0.084	0.528	0.034	0.092	0.192	0.090	0.092	0.308	0.055
Z-score/salary	0.120	0.100	0.052	0.118	0.017	0.009	0.047	0.021	0.009		

SEA	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Fick	Butler	Murphy	Young	Reynolds	Ellsbury	Salmon	Damon	Ellsbury		
Z-score	0.541	0.391	0.599	1.300	0.354	0.624	0.354	0.536	0.624	0.591	0.209
Z-score/salary	0.815	0.095	0.346	0.172	0.110	0.104	0.053	0.077	0.104		

MIA	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Lucroy	Delgado	Lopez	Escobar	Headley	Davis	Swisher	Baines	Davis		
Z-score	0.752	0.651	0.548	0.199	0.238	0.235	0.423	0.448	0.235	0.414	0.204
Z-score/salary	0.819	0.058	0.316	0.073	0.056	0.084	0.068	0.282	0.084		

NYM	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Molina	Coomer	Gennett	Bell	Gaetti	Klesko	Bradley	Ethier	Klesko		
Z-score	0.157	0.048	0.122	0.243	0.255	0.015	0.068	0.326	0.015	0.139	0.077
Z-score/salary	0.056	0.058	0.242	0.043	0.227	0.003	0.016	0.044	0.003		

WSH	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Hoiles	Martinez	Carpenter	Andrus	Bautista	Bonds	Stanton	Gordon	Bonds		
Z-score	0.014	0.099	1.490	0.328	0.309	1.413	0.305	0.691	1.413	0.673	0.334
Z-score/salary	0.004	0.019	2.254	0.132	0.048	0.101	0.154	0.192	0.101		

BAL	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Montero	Jackson	Loretta	Ramirez	Sandoval	Vaughn	Justice	Wilson	Vaughn		
Z-score	0.207	0.106	0.422	0.466	0.455	0.044	0.197	0.173	0.044	0.235	0.084
Z-score/salary	0.051	0.061	0.188	0.117	0.147	0.007	0.028	0.146	0.007		

SDP	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Wilson	Thomas	Cantú	Theriot	Kouzman	Grieve	Richard	Ordonez	Grieve		
Z-score	0.161	0.399	0.284	0.331	0.185	0.004	0.127	0.526	0.004	0.225	0.135
Z-score/salary	0.136	0.054	0.148	0.234	0.102	0.002	0.481	0.059	0.002		

PHI	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Rosario	Giambi	Callaspo	Escobar	Wigginton	Long	Berkman	Winn	Long		
Z-score	0.360	0.737	0.109	0.025	0.068	0.125	1.190	0.218	0.125	0.329	0.137
Z-score/salary	0.734	0.094	0.055	0.018	0.030	0.054	0.129	0.062	0.054		

PIT	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Johjima	Young	Hudson	Scutaro	Jones	Abreu	Rios	Quinn	Abreu		
Z-score	0.294	0.095	0.105	0.102	0.976	0.708	0.328	0.160	0.708	0.386	0.111
Z-score/salary	0.046	0.013	0.033	0.032	0.093	0.091	0.044	0.557	0.091		

TEX	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Molina	Pujols	Gordon	Garciapar	Prado	Cust	Gonzalez	Young	Cust		
Z-score	0.361	1.151	0.347	1.008	0.423	0.183	0.472	0.027	0.183	0.462	0.158
Z-score/salary	0.069	0.104	0.694	0.167	0.109	0.106	0.061	0.010	0.106		

TBD	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Hernande	Youkilis	Scutaro	Bellhorn	Freese	Bichette	Williams	Blake	Bichette		
Z-score	0.159	0.542	0.287	0.579	0.107	0.314	0.576	0.065	0.314	0.327	0.131
Z-score/salary	0.048	0.087	0.089	0.712	0.065	0.050	0.058	0.018	0.050		

BOS	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Wieters	Hafner	Aurilia	Rodriguez	Ensberg	Teahen	Anderson	Lee	Teahen		
Z-score	0.433	0.483	0.170	0.898	0.222	0.192	0.398	0.587	0.192	0.397	0.085
Z-score/salary	0.149	0.079	0.070	0.046	0.139	0.073	0.073	0.063	0.073		

CIN	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/Sa
	Davis	Bagwell	Castillo	Jeter	Ventura	Daulton	Prado	Matsui	Daulton		
Z-score	0.222	0.893	0.011	0.937	0.031	0.380	0.351	0.605	0.380	0.423	0.081
Z-score/salary	0.291	0.079	0.003	0.038	0.005	0.079	0.091	0.065	0.079		

COL	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Martin	Lee	Phillips	Castro	Machado	O'Neill	Dunn	Ramirez	O'Neill		
Z-score	0.462	0.505	0.419	0.457	0.040	0.377	0.776	0.793	0.377	0.467	0.084
Z-score/salary	0.102	0.077	0.063	0.166	0.079	0.061	0.090	0.058	0.061		

KC	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Piazza	Teixeira	Young	Adams	Chavez	Cameron	Puig	Choo	Cameron		
Z-score	1.239	0.551	0.049	0.154	0.538	0.093	0.463	0.342	0.093	0.391	0.091
Z-score/salary	0.118	0.040	0.018	0.307	0.102	0.018	0.125	0.076	0.018		

DET	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Gomes	Olerud	Kipnis	Eckstein	Hillenbran	Dye	Higginson	Mora	Dye		
Z-score	0.250	0.243	0.351	0.300	0.437	0.495	0.151	0.022	0.495	0.305	0.152
Z-score/salary	0.455	0.039	0.333	0.152	0.181	0.086	0.026	0.007	0.086		

MIN	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Barrett	Clark	Walker	Vizquel	Cirillo	Bruce	Wells	Guerrero	Bruce		
Z-score	0.229	0.515	0.236	0.163	0.098	0.197	0.447	0.580	0.197	0.296	0.074
Z-score/salary	0.114	0.092	0.177	0.049	0.030	0.045	0.048	0.069	0.045		

CHW	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Lieberthal	Howard	Menechin	Desmond	Spiers	Lawton	Calhoun	Alvarez	Lawton		
Z-score	0.258	0.577	0.235	0.380	0.157	0.103	0.162	0.289	0.103	0.252	0.183
Z-score/salary	0.053	0.040	0.670	0.189	0.109	0.029	0.324	0.202	0.029		

NYY	C	1B	2B	SS	3B	OF1	OF2	OF3	DH	Team Z S	Team Z/S
	Baker	Johnson	Walker	Aviles	Beltre	Scott	Byrd	Colvin	Scott		
Z-score	0.111	0.093	0.363	0.094	0.547	0.014	0.120	0.034	0.014	0.155	0.064
Z-score/salary	0.176	0.036	0.146	0.067	0.062	0.005	0.040	0.037	0.005		

BIBLIOGRAPHY

Badenhausen, Kurt, Michael Ozanian, and Christina Settimi. "The Business Of Baseball." Forbes. Forbes Magazine, 25 Mar. 2015. Web. 28 Oct. 2015.

"Baseball Reference." Baseball-Reference.com. N.p., n.d. Web. 30 Mar. 2016.

Baumer, Ben. "Quantifying Market Inefficiencies in the Baseball Players' Market." *Eastern Economic Journal* 40.4 (2014): 488-98. Web. 10 Apr. 2015.

Caporale, Tony, and Trevor C. Collier. "Scouts Versus Stats: The Impact Of Moneyball On The Major League Baseball Draft."Applied Economics 45.15 (2013): 1983-1990. Business Source Premier. Web. 2 Oct. 2015.

Forbes. Forbes Magazine, n.d. Web. 2 Feb. 2016.

Green, Jeff. "Extreme Moneyball: The Houston Astros Go All in on Data Analysis." *Business Week* Aug. 2014: n. pag. Web. 5 Apr. 2015.

Grier, Kevin, and Tyler Cowen. "The Economics of Moneyball." *Grantland*. N.p., 09 Dec. 2011. Web. 26 Apr. 2015.

Hakes, Jahn K., and Raymond D. Sauer. "An Economic Evaluation of the Moneyball Hypothesis." *Journal of Economic Perspectives* 20.3 (2006): 173-85. Web. 15 Apr. 2015.

Ilardi, Steve. "The Next Big Thing: Real Plus-Minus." ESPN. N.p., 16 Apr. 2014. Web. 26 Oct. 2015.

Kahane, Leo, and Shane Shmanske. "Team Roster Turnover and Attendance in Major League Baseball."*Applied Economics Letters* 29 (1997): 425-31. Web. 12 Apr. 2015.

Kleps, K. (2014). Indians are Forecasting, Studying Park Attendance. Crain's Cleveland Business, ,35(30), 6. Retrieved from http://search.proquest.com/docview/1550799465?accountid=13158.

Lee, Young Hoon. (2014). Stochastic Frontier Models in Sports Economics. International Journal of Sport Finance, 9(4), 360-74. Retrieved from http://search.proquest.com/docview/1625563191?accountid=13158

Lemire, Joe. "Billy Beane Finds New Moneyball Inefficiency: Oakland GM Has the A's Back in the Playoffs by Maximizing All 25 Roster Spots." *Wall Street Journal* 30 Sept. 2014: n. pag. Web.

Lewis, Michael. Moneyball: The Art of Winning an Unfair Game. New York: W.W. Norton, 2003. Print.

"MLB Team Valuations." Bloomberg.com. Bloomberg, n.d. Web. 2 Feb. 2016.

Rastogi, Siddhartha K., and Satish Y. Deodhar. "Player Pricing And Valuation Of Cricketing Attributes: Exploring The IPL Twenty20 Vision." Vikalpa: The Journal For Decision Makers 34.2 (2009): 15-23. Business Source Premier. Web. 2 Oct. 2015.

Regan, Colin S. "The Price of Efficiency: Examining the Effects of Payroll Efficiency on Major League Baseball Attendance." *Applied Economics Letters* 19 (2012): 1007-015. Web. 20 Apr. 2015.

Spring Forward. (2015, March 4). The Economist (Online), Retrieved from http://search.proquest.com/docview1660966434?accountid=13158

"The Baseball Guru - A World of Baseball!" The Baseball Guru - A World of Baseball! N.p., n.d. Web. 30 Mar. 2016.

Yang, Yupin, and Mengze Shi. "Rise And Fall Of Stars: Investigating The Evolution Of Star Status In Professional Team Sports." International Journal Of Research In Marketing 28.4 (2011): 352-366. Business Source Premier. Web. 2 Oct. 2015.

ACADEMIC VITA
CHRISTOPHER A. LOGGIA

EDUCATION

The Pennsylvania State University | The Schreyer Honors College — **University Park, PA**
Smeal College of Business | Bachelor of Science in Finance — *Class of May 2016*
College of the Liberal Arts |Minor in Economics

RELEVANT EXPERIENCE

Perella Weinberg Partners — **New York, NY**
Investment Banking Summer Analyst, Generalist — *Jun 2014 — Aug 2015*

- Advised on a fairness opinion and helped prepare internal presentations, various valuation analyses and conducted industry research and benchmarking for a potential $32 billion TMT merger
- Prepared a pitch book that included standalone and pro forma operating models to analyze organic and inorganic growth opportunities for a potential sale or series of acquisitions for Interxion, a $2 billion data storage firm
- Authored client presentations to explore potential strategic avenues for growth utilizing comparable and accretion dilution analyses for both an $180 billion consumer staples firm and a $9 billion industrials company

Nittany Lion Fund, LLC. — **University Park, PA**
President, Executive Board — *Dec 2014 — Dec 2015*

- Directed an investor relations program of over 75 investors to maintain formal relations through communicating the Fund’s performance, activities and internship success through annual, quarterly, monthly and weekly reports
- Enhanced the efficiency and performance of the $7.0 million Nittany Lion Fund by implementing faster voting mechanisms, thesis development and a streamlined portfolio analytics report with improved accuracy
- Ensured 100% internship placement for 44 Fund analysts and increased exposure for over 300 general members of PSIA by initiating weekly Monday morning market meetings, weekly educations and networking visits

Lead Analyst, Financials Sector — *Dec 2013 — Dec 2014*

- Managed the $1.1 million Financials Sector within the $7.0 million Nittany Lion Fund by using excess equity models, comparables, and ratios to analyze nine holdings
- Composed presentations using fundamental analysis and short-term catalysts to explain overall investment thesis for buy, sell, or hold recommendations to the Executive Board, Fund Managers, and a general body of 300

Wall Street Boot Camp — **University Park, PA**
Participant — *Jan 2013 — May 2013*

- Selected from 300 applicants to participate in weekly Wall Street educational and networking presentations

LEADERSHIP EXPERIENCE

Alpha Kappa Psi, Professional Co-ed Business Fraternity — **University Park, PA**
President, Executive Board — *Dec 2014 — Dec 2015*

- Augmented the professional and philanthropic development of 120 brothers to ensure 100% job placement amongst graduating seniors and eligible juniors by increasing professional exposure and workshops
- Devised and implemented a point system and corporate sponsorship program to accurately assess and reward active brothers by incentivizing greater attendance with scholarships sourced from affiliated corporate partners
- Facilitated the maturation of over 40 pledges to acclimate the classes to fraternity principles and essential professional skills through general guidance, providing 120 brother interviews and hosting several events

Vice President of Performance, Executive Board — *Dec 2013 — Dec 2014*

- Collaborated with 10 different chair positions to facilitate the evolution of various fraternal events, ranging from parent’s weekend to professional affairs, while focusing events on inter-business fraternity relations
- Engaged fellow Executive Board members in weekly meetings to discuss the progress of specific fraternal events and potential actions to further the professional, social and philanthropic development of 94 brothers

Leadership Jumpstart — **University Park, PA**
Leadership Jumpstart Teaching Assistant — *Jun 2013 — May 2015*

- Advanced the curriculum of Leadership Jumpstart by providing constant feedback and reflection on class events and designing class lectures that challenge and question students’ leadership abilities through thought and action
- Counseled a service project team of six students through attending bi-weekly meetings and offering insight into project proposals while analyzing and testing leadership horizons through personal and team journal reflections

Schreyer Honors College Orientation — **University Park, PA**
SHO-Time Mentor, Logistics Team and Hospitality Team — *May 2013 — Jun 2015*

- Guided a group of eight freshmen students through all four college years, specifically a three-day orientation, contributing advice on college life and an inviting environment to acclimate to a new home away from home

SKILLS, AWARDS & INTERESTS

- Working knowledge in Bloomberg, Factset, Capital IQ, Microsoft Office: Excel, PowerPoint and Word
- Walter Udovich Humanitarian Award, President’s Freshman Award, President’s Sparks Award, Evan Pugh Scholar Award, Marty D’Ambrosio Honors Scholarship and Sam Wherry Scholarship
- Interests include: playing baseball, football, collecting mini team bats, trying global cuisine, and traveling