How Major League Baseball Scouts have Impacted Valuation Yields

By

Sandeep Satish

An honors thesis submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Science
Undergraduate College
Leonard N. Stern School of Business
New York University
May 2009

Marti G. Subrahmanyam
Charles E. Merrill Professor of Finance
Faculty Advisor

David L. Yermack
Professor of Finance
Thesis Advisor

Table of Contents

Abstract

The acquisition of talent in the Major League Baseball draft is a phenomenon few teams have mastered consistently since its inception in 1965. The draft provides a vital resource for teams searching for future stars of the game. Despite growing free agency spending and high team payrolls unrestricted by a salary cap, the draft allows small-market teams the opportunity to develop talent in-house. Or does it? Using a unique data set of the number of scouts employed in every other year from 1987 to 1999, I empirically test the relationship between player valuations and the scouting resources utilized by teams. My results indicate that the number of scouts a team employs within the domestic draft region (United States, Canada, and Puerto Rico) has no bearing on the successful valuation in a given draft year when controlled for the variables of league, payroll, first pick position, and prior year winning percentage. Based on this result, I analyze key metrics associated with a team's utilization of scouting resources. I find that while National League teams employ more scouts over the time, the American League has performed better. Financial metrics such as operating income and payroll have negative explanatory power on win-share valuation. As operating income and payroll increase, team draft performance falls. Furthermore, general managers who have experience in the player development system are much better at drafting than those who do not. My thesis concludes scouting adds no value based on the data, but the statistical analysis is ultimately inconclusive. Teams should source scouting resources to the Central Scouting Bureau and focus on international acquisition of talent.

Note of Thanks

Baseball has and always will be my passion. I wanted to take this opportunity to thank all those who have helped me not only with research, but in my personal development through understanding the thesis process. Professor David Yermack has been tremendously supportive since the summer of 2008 when I first met him. He is always positive and constantly provides constructive feedback. Without his due diligence and quickness in response, I would not be where I am today. Professor Subrahmanyam is a devout supporter of the thesis program and his kindness cannot be overlooked. Vince Gennaro, author of *Diamond Dollars*, has been proactive in his guidance and suggestions for this thesis paper. I have been in contact with him on a bi-weekly basis and his support and passion for the game is inherent in this paper. Additionally, the researchers at the Cooperstown Baseball Hall of Fame, Tim Wiles and Gabriel Schechter, offered a day's worth of time to provide the best resources for the project. Similarly, Victor Wang, a fellow college student, is light years ahead in his ability to analyze baseball data. His research, articles, and databases have been at the foundation of this thesis. I also have to thank my partner on the project, Vishal Gandhi, who dedicated his senior year to help with this project. Finally, the opportunity to be in contact with Bill James has been nothing but admiration from my end. To hear from him will be a memory for the rest of my life. Thank you again to everyone.

Introduction

"The knowledge of who will improve is vastly more important than the knowledge of who is good. Stats can tell you who is good, but they're almost 100 percent useless when it comes to who will improve."

-Bill James, on scouting[1]
April 2008

From 1999 to 2003, the Oakland A's won three division titles with an average winning percentage of .592 per season. This period of success was the central focus of Michael Lewis' *Moneyball*, the story which shed light on the impressive drafting and scouting methods of A's General Manager Billy Beane. The book brought to the forefront the debate between statistical analysis and traditional scouting. The A's in this period were always in the lower quartile of payroll, and thus, Beane's ability to work more with less was glorified. This is not to say the A's did not employ scouts like other teams; they just had fewer. The tension between traditional scouting methods and new knowledge from the area of sabermetrics created cause for concern over the viability of scouts as justified by Beane's track record.

Today, there is significant equalization of technology, process, and information among teams in the scouting area. The increase in these three variables as well as competition to find the best talent has diffused the best scouting practices to teams around the league. Thus, scouting is a function of a much more difficult draft and the return on scouting is questionable. Conversely, Bill James, as depicted in the introductory line, consistently argues that while statistics are important, scouts play and will play a vital role in successful drafting. Does the data hold true to this argument? The strategy to build a team though the draft, and particularly through employment of scouts, is the focus of this paper. While significant research is available

[1] http://freakonomics.blogs.nytimes.com/2008/04/01/bill-james-answers-all-your-baseball-questions/

in the areas of player valuation, I focus on the people who find the players, the scouts, and the variables which affect them.

My paper proposes a framework for understanding how useful the number of scouts really is based on data from the 1987-1999 period. I selected this time period because after 1999, there is a drop-off in the number of players who have made the Major Leagues. I begin by discussing recent trends in order to understand why it is important to study the draft in the first place. I highlight draft characteristics and provide figures to demonstrate how scouting efficiency is not improving, despite significant improvements in technology, communication, and measurability. I continue with important assumptions since this thesis is an experimental analysis on data which has not been analyzed before. Nonetheless, prior literature demonstrates that the draft itself is a topic of significant research volume with particular emphasis on valuing players. This paper uses some of those metrics so it is important to highlight research before me.

Initially, I begin with this question: if a team has more scouts, are they better at drafting because they have more scale and geographical coverage or does it not matter? More scouts should mean more human capital and more geographical scope to find talented players. Once this question is answered I will look for answers to important follow-up questions which affect scouting such as: do teams with high payrolls focus less on scouts and more on free agency? Do teams with high operating incomes (i.e. financially stronger teams) employ more scouts? Wealthier teams should have the resources to afford more scouts and particularly more talented scouts, but this is not true.

In addition to financial metrics, does the General Manager make a difference with regards to scouting philosophy? I test for General Manager effects and look for patterns in particularly successful executives in their approach to scouting. Today especially, with a vast compilation of

detailed statistics and superior technology to measure and find talented players, drafting should be more efficient. Before I analyze efficiency, I begin by asking a simple question: Does the draft matter to a team's success?

Draft Trends

The drafting success and consequent American League Pennant won by the 2008 Tampa Bay Rays is a great start to answer this question. The roster of last year's team included an astonishing 43% of players who were drafted, signed as an amateur free agent, or signed as an international player by the team. Close to half of the team was made up of players who were valued by in-house resources such as scouts, crosscheckers, and consultants. I then decided to analyze the eight most recent winning World Series rosters to gauge how much value the draft added to the team's success:

Table 1: Analysis of Recent World Series Champion Rosters

World Series Winners	Philadelphia	Boston	St. Louis	Chicago (AL)	Boston	Florida	LA Angels	Arizona
Drafted or Signed by Team/Total Players	40.54%	34.21%	30.77%	32.35%	21.62%	23.68%	41.67%	31.71%
Total Drafted/ Total Players	78.38%	76.32%	79.49%	73.53%	83.78%	78.95%	86.11%	85.37%

With the exception of the Boston Red Sox in 2004 and the Florida Marlins in 2003, each championship team's season roster (not including minor September call-ups), included at least 30% of players who were drafted, signed as a free agent, or signed as an international player. This data table, however, may not provide the most conclusive evidence since drafted players take time to develop. As a result, more meaningful data stems from prior periods. For my thesis, I've analyzed the data period from 1987 to 1999 and in this time frame, the championship rosters averaged an astonishing 46.76% of players who were either drafted by the team or signed as an amateur free agent. The 1988 Los Angeles Dodgers and the 1995 Atlanta Braves had more than 62% of their season rosters scouted through organizational resources. No team in this time

period had less than 30% of their team built through the draft. The point is clear; the draft is a valuable and cost-effective resource to develop talent. The next logical question, then, is: do scouts matter? Understanding how the draft is structured is a critical first step in determining scouting value.

The Draft Structure

The Major League Draft officially started in 1965. Prior to this year, teams were free to sign whomever they desired at whatever price they could afford. This led to inequality in team talent and so a draft was implemented to create parity. However, until 1987, the MLB instituted several drafts throughout the year and thus, teams which could afford to scout year long and spend resources for each draft were able to sign more players. Accordingly, in 1987, the draft format was once again changed to its current format: one draft held in June for 50 rounds based on record and loss of free agent players. Consequently, the analysis in this paper focuses on drafts post-1987.

The American League and National League alternate picks annually, starting with the first pick overall, which goes to the team with the worst record. Teams remain in the draft until they pass, until they have finished making their picks, or until the draft is over. Graduating high school seniors, all junior college players, players that have completed their third year of college, and players that have turned 21 years of age within 45 days of the draft are all eligible.[2] The team which selects the player in a given round retains the sole negotiating rights and a contract must be agreed to by August 15. If no contract is reached, negotiating rights are terminated and the player enters the open market as a free agent. Major League teams are compensated for losing free-agent player (ranked by type) with a sandwich pick that falls in between the first and second

[2] http://www.angelfire.com/vt/prospectwatch/index88.html

round. In addition, a pick is also provided for a club which failed to sign their first-round pick from the prior year. These picks are supposed to help small market teams who can not afford to sign players once free agency hits. If these teams employ an effective scouting system, then the results of their draft picks should be significant. The data, however, does not support this theory and the following analysis will determine why.

Assumptions

In my thesis, there are several major assumptions I utilize. First and foremost is distinguishing between the value given to a team's draft class based on scouting ability versus signing ability. While scouts certainly find talent for the team, the organization's monetary capabilities may not match or fulfill the player's wants or needs. J.D. Drew and Mark Prior, for example, were both drafted and went unsigned, and they reentered the draft pool in a later year. For much of this analysis I give credit to scouts who helped find the player the team drafted; this means players may not have signed with the team. To adjust for this, I later analyze just the top 100 picks in each draft class to test the same hypotheses but this time including only players who signed with the team. The argument for leaving unsigned players in the data analysis is that a scout may not have much control over whether a team can sign the player to the bonus the agent asks for. The scout, however, should not be penalized for finding this talent.

On the issue of signing players, Keith Law of ESPN wrote, "Small-market or just plain cheap teams selecting near the top of the draft are hostages of their situations. If the best player on the board wants a bonus well above slot for that teams' position, and they are unwilling or unable to pay, they must select the best player on the board whom they can afford." [3] In the process they have no way to recover the value they lost from bypassing the best player.

[3] http://www.bizofbaseball.com/index.php?option=com_content&view=article&id=2621

The credit to the scout is the major assumption of this paper. Another major assumption stems from the player valuation output. I used Bill James' metric of Win-Shares[4] and took the cumulative value a player earned for a six-year period. This is a major assumption since players are often traded before the six-year mark or are sent back down to the minors. The reason six years is used for the analysis is players are not eligible for free agency until after this period. Thus, a scout should get credit for a player's first six years of value since the team which drafted the player cannot lose him to free agency. This is certainly subject to criticism since a player may get better in a different system, but the scout should nonetheless get credit for finding that player in the first place. More assumptions will be analyzed in the data collection section.

Prior Literature

My paper is related to several areas of baseball research. Much of the existing academic research on player valuation has stemmed from the legendary Bill James who coined the term sabermetrics and created the Win-Share metric (2002). James focuses his research on a wide array of baseball topics, but particular to the draft he has analyzed adjusting statistics to park factors and the value of college players over high school players. His annual baseball abstracts in the 80s and 90s led to the creation of several more complicated and thorough statistical standards to measure a player's true value to a team. Over the years, the main conclusions from James are:

1. College competition is more difficult to dominate than high school competition. Scouts are bowled over by people who hit .573 and drive in three runs per game; you can't do that in college.
2. A "preference for drafting high school players, however small, might cause college players to be drafted lower than they ought to be. This would cause their rates of return to be higher.

[4] Please see Section B of Data Analysis and Appendix B for explanation

3. Pitchers who have been made very high draft picks (among the first ten players taken) have proven to be quite poor risks.[5]

This analysis is critical for scouts who have since tailored their approach in valuing players by position and level of maturity. James' research is valuable to my output data, but he has not analyzed how effective a team's scouting resources have been over time.

A team's drafting strategy is not limited to scouting reports, but rather a wide variety of factors. University of Iowa Professor Jeffrey Ohlmann (2007) modeled the selection decision as an optimization problem subject to uncertainty induced by imperfect knowledge of competing team decisions.[6] Factors on draft-day decision making includes player valuation, organizational need, budget, player signability, and selection strategy of opposing teams. Interesting observations from his research pertains to the decision-making strategy. As the team becomes more uncertain about what other teams will do, the team will pick more sincerely according to their own valuations. Additionally, as the discrepancy between the drafting team and other teams' player valuations increases there's more opportunity to take advantage.

Similarly, researchers have measured the value of team situational factors in the draft such as slot (Silver, 2008) and Type A free agent draft pick value (Wang, 2009). Interestingly, Victor Wang, a student at Northwestern University, calculates the value of a Type A draft pick between $3-5 million dollars and this certainly would have implications for the importance of a scout's role with this slotted position. If such a pick provides significant value, then a team may allocate more research and scouting resources to find a signable and valuable player to draft with this pick.

While decision-making under certainty is a valuable area of analysis, my paper focuses on the team's ability to utilize scouting resources, not on the optimization of draft strategy. To

[5] http://baseballanalysts.com/archives/2004/11/abstracts_from_20.php
[6] Ohlmann (University of Iowa), Presentation to NYU Stern

summarize, my research is distinct from existing baseball draft research as it focuses on scouting implications rather than statistic and player valuation metrics. I extend the draft literature by measuring scouting efficiency over time through data that has not yet been analyzed.

Data Analysis

A. Data Collection

While the signing issue and player valuation metric delineated above are certainly critical to the central thesis of the paper, data sources must also be scrutinized. I received the entire draft history from The Baseball Cube Register and matched the entries from various other sources such as Major League Baseball and Baseball Reference. The most important data set came from the Cooperstown Hall of Fame research library which holds The Blue Book, a binder of every business transaction from the given year. The four hour journey to Cooperstown was well worth it with the rich baseball literature and statistics available. The Blue Book contains the name and location of every scout for every team in the year. As later analyzed, the volatility in scouting resources is quite particular. Some teams are consistent in the number of scouts they employ whereas others rapidly shift in every other year analyzed. I chose to analyze every other year for two reasons. The task to count each scout for every year was burdensome and I thought unnecessary since too much fluctuation wasn't anticipated. However, by analyzing every other year, the variance in scouting resources emerges. According to the library, some teams may reduce scouts when utilizing the Central Scouting Bureau whereas other teams may increase scouts when they reclassify outside consultants as scouts. Furthermore, I could only utilize this data as given by the teams; the Blue Book tries to verify the accuracy of scouting information, but in international areas few teams may report the exact number of scouts employed.

Finally, I analyzed many metrics of player and team data once the scouting numbers were compiled in a spreadsheet. I utilized Dave Studenmund's Win-Shares Above Average database, a collection of Bill James' adjusted win-shares player value metric. The win-shares metric will be described below, but it is important to note the source of this valuable output data. Furthermore, I found operating income numbers from baseballchronology.com and team payroll figures from baseball-almanac.com. Both sources stem from Major League Baseball archives. I took all this data and compiled the information into one readable spreadsheet in order to correlate certain figures and develop conclusions based on the questions introduced in the beginning.[7] With baseball data, verification of figures is very important as statistics tend to slightly change from one reference to another. My columns of data are unique in the analytical sense, since no researcher I found has analyzed the impact of various scouting figures and team financial data to assess efficiencies and trends in baseball drafting.

B. Use of Win-Shares

In order to best assess how scouts have fared in the twelve year time period, one of many metrics to value a player needed to be selected. Ohlmann (2007) utilized at bats/drafted hitter but this is a relatively weak measure to value a player because at-bats fluctuate wildly and by league. I determined the most thorough and observed metric of valuation currently is a player's win-share.

Bill James invented Win Shares as a simple way to compare baseball players. The idea was to develop a statistic which allows comparison between first basemen and outfielders, starters and relievers, pitchers and fielders etc. The Win Share methodology is extremely complex to calculate, but the output is simple: it is one number that represents the number of

[7] See Appendix A for how data compilation looks visually

wins contributed by that player.[8] Win Shares is the number of wins contributed by that player multiplied by three in order to provide enough meaningful distinction between players. Batting, fielding, and pitching are the three types of Win Shares. Everyday players tend to garner more Win Shares than pitchers, because they are credited with both batting and fielding Win Shares. Win Shares are even more flexible and encompassing, because the measure adjusts for contexts (this is critical for pitchers and batters who benefit from either pitching-friendly stadiums such as Comerica Park or hitter-conducive parks such as Coors). A run is harder to score in Comerica Park than in Coors Field, so hitters receive more credit for what they accomplish in Detroit versus Colorado. An important attribute about Win Shares is if a player plays for a winning team, he won't get credit for more Win Shares than if he had played for a losing team. Consequently, Win Shares is fair to all players and can be used to compare at any point in their careers.

If the reader is interested, Appendix B includes a rough guide to how Bill James calculates Win-Shares. Additionally, I must briefly explain Win-Shares Above Average, as this is the metric Dave Studenmand sent to me in his database. Calculation of win shares above average (WSAA) is relatively straightforward once the player's win shares have been calculated. The general equation looks like this:

(1) WSAA = [CareerWS - AvgWSperX * CarX][9]

where:

WSAA = Win Shares Above Average
CareerWS = Career Win Shares
AvgWSperX = Average Win Shares per Time Analyzed
CarX = Time Frame Analyzed

[8] http://www.hardballtimes.com/main/article/2004-win-shares-have-arrived/
[9] http://members.cox.net/~harlowk22/atgwsobj.html

While this formula seems complicated, it is simply important to note that Win Shares Above Average is what an average player would have produced with the same playing time at his position.[10]

C. Team Scouting Trends (1987-1999)

The question of how scouting has influenced draft success is best observed by first looking at patterns over time. Looking at a simple count of scouts depicts an interesting story of allocation by league.

Table 2: Allocation of Scouting Resources (Top Teams)

Top 5	Average # of Scouts Per Year (Draftable Region)	Range	Cumulative Win Shares of Players Drafted	Average Win Share Per Draft Class
Dodgers	45	37	847	121
Braves	44	19	540	77
Reds	42	22	749	107
Mariners	40	64	1243	176
Marlins	39	10	444	111

With the exception of the Seattle Mariners, the National League has consistently employed more scouts per year on average then their American League counterparts. I included a range column (highest number scouts in one given year minus the lowest) to demonstrate the volatility in scouting resources over time. The second-to-last and last columns incorporate the win-share metric to see how effective the scouts have been for the teams. I also used average

[10] http://dodgerthoughts.baseballtoaster.com/archives/014740.html

win-share per draft class to normalize for teams such as the Diamondbacks, Marlins, Rays, and Rockies who have not been in the league as long as the others.

The range column is interesting, particularly for the Seattle Mariners. In 1987 the Seattle Mariners employed 19 scouts in the draftable area under Dick Balderson, the general manager. In 1988, Woody Woodward became the general manager and the Mariners increased the scouts in their system to 53 in 1989, a 179% increase in a two year time span. This significant increase, however, did not correlate with an increase in draft production. In fact, the reverse occurred as Seattle's draft class fell precipitously from 232 to 51 in Win Share value. On the other hand, teams which employed the lowest scouts on average over the period were:

Table 3: Allocation of Scouting Resources (Bottom Teams)

Bottom 5	Average # of Scouts Per Year (Draftable Region)	Range	Cumulative Win Shares of Players Drafted	Average Win Share Per Draft Class
Rangers	24	11	813	116
Royals	23	7	1002	143
Cardinals	22	10	1196	171
Twins	22	11	1376	197
A's	18	11	867	124

This data table provides very interesting information as we see contradictory results from the top scouting teams. Four American League teams are at the bottom of the average scouting allocation compared to four National League teams who employed significantly more scouts. Furthermore, this group of teams is very consistent in the twelve year span, barely fluctuating in the number of scouts utilized. Most importantly though, are the results of the draft valuation for these teams. Texas, Kansas City, St. Louis, Minnesota, and Oakland almost completely

outperform the top scouting teams. The data indicates that while the National League generally has employed more scouts, the American League has outperformed their counterparts in drafting. Ten of the top fifteen performing teams in the twelve year period come from the American League. Similarly, twelve out of the top fifteen scouting teams come from the National League. Thus, the difference extends beyond the extremes in both cases. Why is this the case?

The absence of the designated hitter may affect a National League's drafting strategy. While American League teams can select great hitters out of college or high school who may have no other positional role than hitting, National League teams may seek more pitching or talented utility-type players who could eventually prove pivotal with double-switching and pinch-hitting strategies.

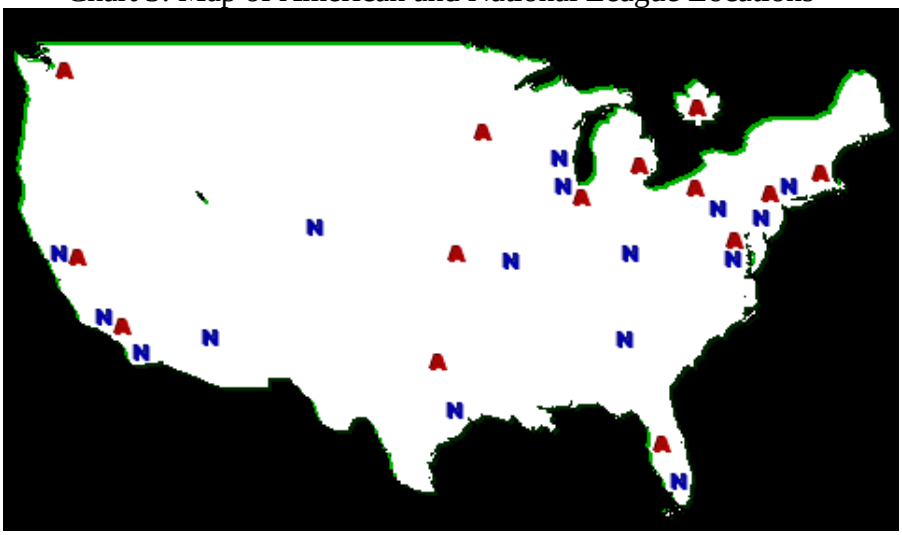
In a 2006 study, a baseball blog[11] compared the differences between the two leagues which shed some light on possible theories as to why this scouting and drafting disparity occurred. The National League utilizes, on average, many more pitchers in a given game particularly in later innings when pitchers are lifted for bench players. Thus, NL teams may draft to fit need whereas AL teams may draft simply for talent. As a result, AL-drafted players could make a more immediate impact based on their position while NL-drafted players may be under utilized in their early years due to strategy. Similarly, the author of the blog observed a striking difference in style and speed for the National League versus American League. In general, AL teams are associated with power and long-ball whereas NL teams play small ball and steal a lot of bases. For example, in the author's study, the NL out stole the AL by nearly 150 bases in the year analyzed. Consequently, the NL may draft speedy players, situationally skilled hitters, fielders, or pitchers which would reduce their win-share contribution. The differing trend in scouting allocation and draft variance has been consistent in the time period. National League

[11] http://progressiveboink.com/archive/avn.htm

teams generally cluster near the top in number of scouts and near the bottom in draft success. This difference may point to why the AL has won many more all-star games and World Series than the NL in the past twenty years.

Drafting superiority keeps development costs low while infusing young talent into starting lineups. I do believe in addition to strategy, location is an important factor in draft analysis and why scouting figures are disparate. Scouting trends may be a function of a team's location since particular areas may foster more or less commitment to scouting in general. A map of Major League teams provides a visual look at this:

Chart 3: Map of American and National League Locations[12]



National League teams are generally located further south than American League teams (Milwaukee was in the American League for a long time as well) and consequently have easy access to scouting warm weather regions. As a result, it is very possible NL teams employ more scouts to analyze the abundance of players in the warm regions. I also wanted to test whether location factors affected draft selections. After all, familiarity and accessibility facilitate drafting local players. Regions referred to here are based on the Little League World Series regions (i.e.

[12] www.ballparks.com/baseball/general/**maps**/current.htm

Mid-Atlantic, New England, Southeast, Southwest, etc.). I tallied the number of players drafted for every team's particular region against all players drafted by that team in the time span. For example, if a team was close to two regions, I included all drafted players from both regions. The percent column represents the number of players drafted in the tangential area over total players drafted. Below is what I found:

Table 4: Home Drafting Bias

Team	Region	%	Team	Region	%
Angels	West	55%	Rangers	Southwest	41%
Marlins	Southeast	51%	Padres	West	38%
Braves	Southeast	48%	Dodgers	West	37%
Rays	Southeast	47%	Giants	West	36%
A's	West	47%	Diamondbacks	West	33%

This is a fairly strong indicator that warm weather teams draft from their local areas which correlates to emphasis on scouting in these areas. The notion of teams overdrafting from their home region may have a business rationale. According to baseball consultant Vince Gennaro, "teams generally believe that local talent has a positive impact on attendance, all other things equal." The correlation in the free agent market was the Orioles' and Nationals' pursuit of Mark Teixeira, as he is from the area. Thus, in summary, warm weather teams have a local bias in drafting and their selections correlate to the effectiveness of their scouting systems.

D. Variable Models of Testing (Explanation of Data)

In order to statistically test the impact of the relationship between number and location of scouts versus player draft value, I compiled a spreadsheet with several important variables. Key variables include: league, first pick position, prior year winning percentage, number of scouts for

draftable players, team payroll, operating income, general manager, and of course, cumulative win-share valuation. The league variable was a simple dummy variable, with 0 representing an American League team, and 1 representing a National League. The first pick position is an important variable to test because one would expect the team drafting first to do best in the draft due to the reverse-order structure. I differentiate between first pick position and prior winning percentage since first pick position could be a function of the team's off-season activities (trades, signings, etc.). Team payroll and operating income are financial metrics which are tested against number of scouts to understand if financially stronger teams utilized their advantage to develop their scouting system. Finally, decision-making under uncertainty is what differentiates strong drafting teams from weak teams. Thus, I tested for any General Manager effects.

E. Summary Statistics

Important summary statistics of my database highlight basic characteristics about the draft and how important scouting is. First, I compiled total win-shares by round to address the simple question of how often a productive player really gets selected after the first few rounds. The results are depicted in the following table:

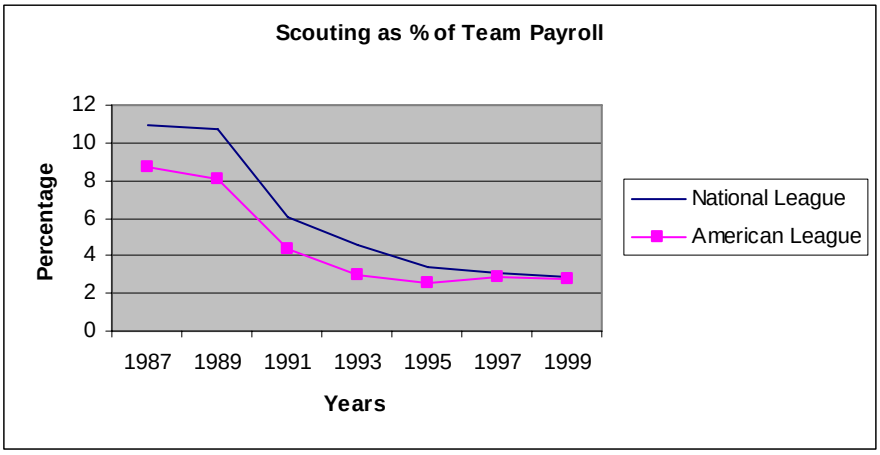
Table 5: Total Win-Shares vs. Round Drafted

Round	WS	Round	WS	Round	WS	Round	WS
1	5608	6	909	11	768	16-25	280
2	2046	7	678	12	240	26-35	155
3	1529	8	795	13	898	36-45	229
4	1024	9	768	14	293	46-55	65
5	1031	10	511	15	370	56+	14

After round 16, win-share values fall consistently, so I took the average of each round for the purpose of relative comparison. Although cumulative win-share values do fall in general, certain later rounds generate variation which is interesting. For example, the 13th round has a significant outlier, since the St. Louis Cardinals drafted Albert Pujols (total win-share value of 216). Similarly, in the 15th round, Jake Peavy of the San Diego Padres and in the 18th round, Bobby Higginson of the Detroit Tigers were selected. I believe this is very important because each player makes up a large proportion of the total win-shares for that round, indicating scouts have the ability to find gems in later rounds though not consistently.

The ability to find talent in later rounds is pivotal since as mentioned in the introduction, information on talent is becoming more and more symmetric. Thus, I hypothesize the data would depict scouting as a function of team payroll should remain consistent so teams could seek value beyond the information available, but this is not the case. Scouting budgets as a percentage of team payrolls have dropped significantly over time as evidenced by the following chart:

Chart 4: Scouting Budgets as a Percentage of Payroll over Time



The National League consistently spent more on scouting as a percentage of team payroll which supports the fact that the league utilizes more scouts, but the strategy doesn't necessarily equate to successful drafting. Scouting budgets as a percentage of payroll fell significantly in twelve

years and as a result, teams may be shifting capital allocation to other variables such as stadium development, free agency, and marketing. Is there statistical significance for the number of scouts on win-share valuation then?

F. Impact of Number of Scouts

A simple regression testing number of scouts in the draftable area against win-share valuations without control of other important variables can be found in Appendix C. The low R^2 implies the model explains only 3% of the variance in win-shares, but with a t-statistic of -2.696, the variable does have significant inverse explanatory power in the regression. Teams with more scouts should be more effective at covering larger areas of the region to evaluate players. Therefore, the more players they evaluate, the probability of efficiently and effectively selecting draft picks should increase, not the other way around. This model is only introductory to understand the impact of the number of scouts on win-shares in isolation. To verify the initial significance, I created a dummy variable for specific ranges of scouts. For example, teams which employed scouts in the range of 11-20 were coded as 1, 21-30 as 2, and so on. The resulting regression is also found in the same appendix and the t-statistic is -3.08 for the ranges of scouts. Higher ranges seem to perform worse than lower ranges, which again confirm the earlier assessment that NL teams employ more scouts, but yield less from the draft. I also want to mention here, with traditional standard regression procedures, variables are assumed to be perfectly reliable without noisiness from measurement error. Since I selected data from every other year rather than each year, not incorporating measurement error causes a reduction in statistical power for detecting relationships among variables. Thus, while the variable is negatively significant, ultimately I believe the statistical analysis done here is inconclusive because of the nature of the data. However, the number of scouts for draftable players must be

tested alongside other important factors to win-shares in order to control for variation and correlation. Two variables which directly impact the number of scouts are payroll and operating income.

G. Payroll and Operating Income Effects

> "Dad, do you know what your problem is?"
> "No, son, what is my problem?"
> "You're just too poor to get rich."
>
> —Bill James conversation with father[13]

An important philosophy grounding Bill James' assessment of draft spending stemmed from the above conversation with his father. James argues that some baseball teams are operating very close to margin to have the flexibility and freedom to make long-term investments. They are, in a sense, too poor to get rich. This comment directly correlates to an important measure of understanding investment in scouting. Which of the leagues or specific teams could afford to sign the best players? I analyzed the financial standing (adjusted for inflation) of teams in both leagues to see if either league was in a better position to sign more talented players, or if particular teams utilized their financial success well. Additionally, I hypothesized teams with high payrolls focus more on free agency and less on scouting resources.

Framing the financial question begins with an understanding of existing philosophy. J.P. Ricciardi, current general manager of the Toronto Blue Jays and one time scout for the Oakland A's claims, "The rich clubs can squash other teams. They can overpay the kids. Conversely small market teams can't afford to be as patient as those with deeper pockets and a slower time table."[14] His statement implies revenue disparity between teams creates an uneven playing field in the draft. Essentially, people who share his opinion believe small market teams are dominated

[13] http://freakonomics.blogs.nytimes.com/2008/04/01/bill-james-answers-all-your-baseball-questions/
[14] www.hsbaseballweb.com/pro-scouting/scouts_dig_**deep**.htm

by larger market teams when it comes to competing for stars who have bargaining leverage. Does this assessment hold true in the time period I analyzed?

Payroll differences were marginal between the two leagues in the twelve year span, with an average payroll of $56.7 mm for the AL compared with $56.5 mm for the NL. Additionally, operating income average for American League teams was $6.5 mm versus $7.0 mm for the National League. The Yankees and Braves were the worst performing drafting teams and they had high payrolls throughout the 90s relative to peers. On the contrary, small market teams such as the Mariners, Twins, and Blue Jays were consistently successful in drafting relative to their rich counterparts. Statistically, I ran simple regressions testing payroll and operating income against win-shares.[15] Then I ran a multiple regression testing both payroll and operating income against win-shares because they have a weak negative correlation (-.15). In all three regressions, the variables have a significant negative t-statistic, suggesting inverse explanatory power.

Teams with higher operating incomes and higher payrolls are not spending on draft resources. Instead, clubs must be allocating to one variable or the other (i.e. spend on free agency or focus on the draft). As mentioned earlier, there is significant variance in the relationship between payroll and number of scouts. Small market teams such as the Reds and Mariners have employed many more scouts relative to peers despite low payrolls whereas the A's, Twins, and Royals, also with low payrolls, have employed a very low number of scouts throughout the years. The financial analysis suggests while disparity exists within teams, the effects are not limited to the free agent market. Clubs which are financially strong seem to allocate resources to spend in the market rather than develop and find low-cost, talented labor.

[15] Please see Appendix D for statistical results

H. General Manager Testing

Spending stems not just from team ownership but the general manager's execution of strategy and capital allocation. This section analyzes whether the general manager has any measurable effect on drafting strategy and successful selections. In a 2003 Baseball America article analyzing general manager prospects, writer Josh Boyd observed the role of the top executive has continually changed over time, "The job description of a GM is changing as economics become more and more of an element of the game. Today's general managers have varying backgrounds from scouting (Brian Sabean), major league playing experience (Billy Beane) or Ivy League educations (Theo Epstein)."[16] The best general managers have been described as possessing great leadership, ability to motivate staff, and most importantly having the skill to evaluate and acquire talent. In a unique study analyzing general managers from 1995-2005, Haverford College student Douglas Black evaluated GMs based on the performance of the teams they create. His thesis concluded an increase in team wins and making the playoffs all decrease the likelihood of firings while dollars spent per team, GM tenure, and GM experience increase the likelihood of a General Manager being fired.[17] My focus here is not evaluating general managers on their likelihood to be fired, but rather understanding draft patterns and influences which affect success and failure with similar variables.

Using a simple linear regression, I tested whether a general manager who played professionally or not had any effect on cumulative win-share valuation. The result, found in Appendix E is a t-statistic of significance. The data suggests general managers who played in the majors are significantly worse drafters than general managers who did not. As a result, the

[16] http://www.baseballamerica.com/today/features/031209gmprospects.html
[17] http://triceratops.brynmawr.edu/dspace/handle/10066/584

analysis indicates a bias toward hiring less qualified general managers if they have playing experience.

On the other hand, the top performing general managers (Hank Peters, Gord Ash, Walt Jocketty, Andy MacPhail, Harry Dalton) all have significant experience with the team covering scouting and operations, not necessarily playing. Using the Baseball America Executive Database, I coded the General Managers by the number of years spent in player development and scouting.[18] The coefficient proves significant with a t-value of 2.96 (Appendix E), indicating the more years an executive participates in a scouting system or personnel development role, the more successful he is at drafting. This is an interesting result as it proves while an executive doesn't need to play professionally, a background in scouting and development is vital in order to draft talent. Consequently, the strategy and success of scouting is influenced by a general manager's prior experience in the development aspects of an organization.

The General Manager has significant influence on draft strategy, player personnel, and operational roles. While some executives use differentiated tools of analysis, much of this information is proprietary and hard to tally. The statistical analysis performed in this section indicates the importance of experience in development and scouting. Over this time period though, many general managers were former players, suggesting a bias to hire these types of executives. The data depicts that ex-player general managers are not as successful in drafting compared to peers who have spent time in scouting. More experience in a system, specifically a scouting system, often does help a General Manager in his draft performance over the long run.

[18] Baseball America Executive Database

I. Changes in Win-Loss Percentage

The impact of a team's performance in the year prior to the draft should be an important variable because of the reverse-order draft structure described earlier. Teams with worse records have higher picks in the next year's draft and should have the best players available to choose from. Furthermore, teams which not only have poor win-loss records, but have low payrolls are subject to lose players to free agency. In return, they are compensated with sandwich draft picks which are offered between the first and second round. As a consequence, a low winning percentage should give a team a boost in drafting. The regression, found in Appendix F, provides an insignificant t-statistic of -.97. The result indicates teams with weak results are not taking advantage of their draft position either because signing bonuses are particularly detrimental to drafting better players or their scouting systems are very weak. The results seem very counterintuitive from a drafting perspective. If a club suffers a losing season and knows they have a high pick in the draft, then they should focus their off-season on allocating resources to scouting. Teams in this position can sign young talent for a lower market price than free agent players while securing their services for an extended period of time. Data over this period, though, suggests winning percentage and team performance in prior years have no measurable influence in the successive draft. From the period of 1987 to 1999, bonus spending was a significant deterrent to securing top-notch talent. I believe teams who had higher picks either could not afford to secure players or did not want to pay top dollar for the best players. As a result, teams did not utilize their advantageous position in the draft.

J. Relationships between Key Variables

After analyzing all of these variables, I tested for the relationship between these important factors against cumulative win-shares. The resulting regression can be found in

Appendix G. While the model does not have significant explanatory power, three variables are significant with a t-value greater than -1.64. League, first pick position, and team payroll have negative explanatory power in the regression. This result supports earlier conclusions. The American League drafts better than the National League, teams with higher first picks draft poorer compared to peers, and team payrolls also have an inverse relationship with draft success. This suggests teams are either spending on free agents or on the draft, but not both. When all predictor variables are considered against win-share valuation, the number of scouts for the draftable area is now insignificant. This is to say, significant factors such as league, first pick position, and team payroll influence draft valuation, but when considered with these control variables the number of scouts has no explanatory power against draft success.

K. Adjusted Valuation

All the analysis above is based on the key assumption that signability is irrelevant versus scoutability. In other words, cumulative win-share valuation is subject to relevance in this paper. Is it correct to give win-share credit to a team who does not sign a player? As I argued in the assumption section, baseball readers and analysts certainly have the right to criticize this methodology. I gave credit to the scout for finding the player, not necessarily facilitating the financial transaction needed to sign him. However, to fairly assess this criticism, I adjusted the win-share valuation in this section to include players within the top 100 picks of each draft who signed with the team that drafted them. This way I capture both scouting success and signability.

The regression in Appendix H yields similar results as the regression against cumulative total win-shares. League, first-pick position, and team payroll are significant variables and the number of scouts is still insignificant. Accordingly, the corresponding interpretations do not

change. The adjusted valuation regression supports my analysis and so the criticism is addressed.

The Central Scouting Bureau

The analysis suggests data on scouting return is inconclusive, though statistically, the variable is insignificant. What are the implications of these results for the teams? One effective alternative to devoting resource to scouting is to outsource the function to the Central Scouting Bureau. Established in 1974, the Central Scouting Bureau has been and still is a major player in draft scouting. Rather than spend significant resources on in-house scouts and consultants, Major League teams can utilize scouting information provided from one central location for a fraction of the price of having full-time scouts of their own.

Data above indicates tremendous volatility and disparity in the number of scouts teams employ. The Central Scouting Bureau (CSB) has historically employed between 20 and 30 full time scouts (34 full-time scouts today) and various part time scouts (13 currently) across the United States, Canada and Puerto Rico. According to Frank Marcos, Assistant Director at the Bureau, teams do not employ CSB scouts or use their reports when they believe they have sharper tools of analysis, larger regional scope, or asymmetric information to skillfully "outdraft" their competition.

The Bureau ranks position players on certain characteristics including: hitting, power, speed, arm strength and fielding. A scale of 2-8 in each category is utilized to grade players and an Overall Future Potential metric or OFP is created. Forty is the minimum score in order to be considered a Major League prospect and these players are grouped into fringe, average, and definite prospect categories.[19] Furthermore, pitchers are graded on fastball, curveball, slider and

[19] http://mlb.mlb.com/mlb/official_info/about_mlb/scouting_overview.jsp

other, but a scout is given leeway to grade higher if potential to develop another pitch is likely. Overlapping qualities between pitchers and hitters are aggressiveness, instinct, and work ethic. The generic qualities that apply to both position players and pitchers are things like aggressiveness, instinct, dedication, work ethic. The organization's goal for every year is successful evaluation of prospects for the draft in the given year. Before and after the draft, scouts focus on younger players to generate visibility for teams so they can plan ahead.

The Bureau is also very proactive in scouting in all parts of the country. Each Bureau coverage region has showcase events and tryout camps which help players who don't have the chance to be seen. For the price and quality of the scouting bureau, it is surprising why we still see so much variance in the number of scouts.

International Development

No more do we see variance then in the realm of international scouting. From 1987 to 1999, scouting resources abroad increased more than 679%. While countries such as the Dominican Republic and Venezuela have been scouted for decades, the late 90s/early 00s saw a dramatic shift into all regions of the world, from Japan to Australia to Europe. International scouting plays a key role in today's roster dynamics as more than 29% of MLB players come from abroad.[20] Analyzing international development poses a few interesting hypotheses. Is there a relationship between drafting success and international resources? Do teams with more scouts abroad focus less on the draft? Do teams with financial advantages have more scouts abroad? The answer to these questions ties in directly with the hotly contested issue of should there be a worldwide draft?

[20] http://sports.espn.go.com/mlb/news/story?id=282429

A simple linear regression (found in Appendix I) testing international scouts versus cumulative win-share valuation proves insignificant with a low t-statistic. Interestingly enough, team payroll and number of scouts for non-draftable players are positively correlated (.23) and operating income and number of scouts are very weakly positively correlated (.004). Finally, the number of scouts for non-draftable players and draftable-players are highly correlated at .44. This suggests teams with higher payrolls and higher operating income are not necessarily shifting resources from domestic scouts to international scouts. International scouting, though, should be an area where we see rapid growth while domestic scouts should see a drop-off based on valuation analysis. Until international recruiting is either capped or restricted to a draft, this is the area to find very low cost gems. Teams with small resources for scouting development may be able to seize valuable talent abroad, but higher signing fees seem to be blocking this development. Thus, the argument for a worldwide draft is significant and in his article "The Worldwide Draft," legal adviser Arturo Marcano contends, "A worldwide draft would bring some centralization to Latin recruiting; but the extent of centralization depends on the structure, substance, and scope of the worldwide draft."[21] Marcano advocates for democratization, centralization, harmonization, specialization, and implementation if a worldwide draft were instituted. Similarly, Craig Calcaterra of the Hardball Times recently observed, "Increasingly, there is a call for baseball and its union to adopt an international draft in the next collective bargaining agreement in 2012 as a way to streamline and clean up the way players from outside North America are acquired, particularly in talent-rich places like the Dominican Republic and Venezuela. Moreover, supporters say a draft would provide an equal playing field among teams that recruit in Latin America."[22] If the worldwide draft were implemented, scouting would most

[21] http://baseballguru.com/articles/analysismarcano01.html
[22] http://www.hardballtimes.com/main/shysterball/article/worldwide-draft/

likely follow suit and become a centralized means of recruitment. The lack of centralization in the international scouting system fosters club-controlled academies which has led to corruption and increased financial incentives. The next round of labor talks occurs in 2012, so all eyes will be watching for the impact of such a decision.

Conclusion

The fall of 2008 displayed the significance of scouting as instrumental in postseason success. Just as I began this paper with the stark contrast of the Yankees and Rays, the world champion Phillies built their franchise with drafted players like Ryan Howard, Chase Utley, Jimmy Rollins, Cole Hamels, Carlos Ruiz, and Brett Myers. This core group was developed from within just as their opponents in the World Series were.

Even the Boston Red Sox, generally in the top 5 in league payroll over the past decade, have developed a winning team with players such as Dustin Pedroia, Kevin Youkilis, Jed Lowrie, Jacoby Ellsbury, Jon Lester, and Jonathan Papelbon. And as Peter Gammons mentioned in a recent article on drafting and scouting, the Red Sox developed all of these players for less than the Giants were paying Barry Zito.

Similarly, Brian Cashman tried to convince his Yankee superiors that if you fill your team with free agents and keep adding on players in their 30s, it eventually will catch up with you and the only way to replace the aged is to go spend on another generation of players paid for what they did in the past, not their futures.[23] The benefit of scouting, while not necessarily tied to volume, is certainly tied with cost. The best scouts in the age of disparity have to help teams not only find the players who fit within the system and are talented, but players who will sign with the team as well. If successful, a club can secure four to six years of service for $2-3 million. If

[23] http://sports.espn.go.com/mlb/preview09/columns/story?columnist=gammons_peter&id=4030641

a club can judge talent, it can sign players like Evan Longoria, Pedroia and Lester to long-term contracts and pay them in the prime years of their careers, before they turn 30.[24]

After an era of free agency splurging, general managers and scouts a like have realized the value in development and in-house production. As Theo Epstein recently commented, "If they (drafted players) are properly developed and coached, integrating young players into the mix creates an energy that the veteran players feed off."[25] One quick look at SportsCenter highlights from the past two World Series champs shows a young Jon Lester and Cole Hamels performing masterfully in clinching games. Supporting Epstein is Indians General Manager, Mark Shapiro, who advocates "We have to focus on our young players. We know what our budget restrictions are. We have to scout and develop and make good decisions."

The Indians, Marlins, Twins, and Brewers have all built successful and potentially successful teams in the past few years through scouting. As Gammons emphasized in his article, the ability to develop talented, low-cost players might be more important than it has been in any recent year. The dynamic of scouting continually changes and today, scouts spend more time off the diamond with parents and friends to understand a player better. Scouting evaluation has moved beyond understanding how a prospect throws, hits, fields, and runs to understanding how players think and react in certain situations. Unlike basketball or football scouts who evaluate mostly college players, baseball scouts are often trying to project the career upside for a 17-year-old, who hasn't yet attended his senior prom.[26]

Few secrets remain in the draftable region today as everyone knows who the best players are and where they play. How will scouts differentiate themselves in the future? As J.P. Ricciardi observes, "The scout who does his homework can walk away from a potential problem.

[24] http://sports.espn.go.com/mlb/preview09/columns/story?columnist=gammons_peter&id=4030641
[25] Ibid
[26] http://www.hsbaseballweb.com/pro-scouting/scouts_dig_deep.htm

That's what separates the really good scout from all the others."[27] As financial resources and scouting dynamics continue to evolve, it will be interesting to gauge the role of scouting in a few years. The number of scouts in a given organization is a critical financial resource, but not necessarily the most optimal variable for drafting success. This analysis indicates scouts must be utilized successfully in particular regions and especially outside of the draft area to prove valuable. As the data proves, the number of scouts is not a significant variable, but the draft is an important foundation for a team's strategy. The rationale that scouting adds no value could be for a number of reasons based on the analysis I performed. Scouts may not contribute much in the era of data, but I do not believe this is the reasonable conclusion. I do believe the best scouting techniques have diffused to the most efficient teams and this is a logical explanation for why the number of scouts is not significant. Additionally, my regression analysis must be considered with a measurement error since every other year of scouting numbers are left off. My decision to look at variance over a two period rather than a single period leaves me to believe that the specific scouting analysis is truly inconclusive.

Furthermore, international scouting is becoming a fixture in capital budgeting. As a result, analysis in this paper offers a reasonable conclusion that teams should outsource scouting to the Central Scouting Bureau and focus on finding talent abroad. Will domestic scouts ultimately be eliminated? I do not believe so because of their inherent fixture within the game. At a fundamental level, the basic challenge will always be the same for a scout: find talent and maximize every pick in every round. If this is accomplished, scouts will indeed have an impacting role in the draft.

[27] http://www.hsbaseballweb.com/pro-scouting/scouts_dig_deep.htm

Works Cited

"American League vs. National League." 16 May 2006. <http://progressiveboink.com/archive/avn.htm>.

"Baseball America Executive Database." Baseball America. 16 Apr. 2009 <http://www.baseballamerica.com/execdb/>.

Black, Douglas. "Major League Baseball general managers : their turn for the public microscope (1995-2005)." Thesis. Haverford College. Dept. of Economics, 2006. <http://triceratops.brynmawr.edu/dspace/handle/10066/584>.

Boyd, Josh. "Top 10 General Manager Prospects." 09 Dec. 2003. Baseball America. <http://www.baseballamerica.com/today/features/031209gmprospects.html>.

Calcaterra, Craig. "Worldwide Draft." 06 Mar. 2009. <http://www.hardballtimes.com/main/shysterball/article/worldwide-draft/>.

Dubner, Stephen. "Bill James Answers All Your Baseball Questions." Freakonomics: The Hidden Side of Everything. 01 Apr. 2008. The New York Times. <http://freakonomics.blogs.nytimes.com/2008/04/01/bill-james-answers-all-your-baseball-questions/>.

"First Year Player Draft." 03 Mar. 2009 <http://www.angelfire.com/vt/prospectwatch/index88.html>.

Ford, Jerry. "Where Do the Best Baseball Players Come From?" 03 Mar. 2009. Perfect Game USA. <http://www.pgcrosschecker.com/Articles/DisplayArticle.aspx?article=628>.

Gammons, Peter. "The road to success begins with youth." 01 Apr. 2009. ESPN. <http://sports.espn.go.com/mlb/preview09/columns/story?columnist=gammons_peter&id=4030641>.

Harlow, Kevin. "Baseball's All-Time Best Non-Pitchers (Win Shares Based)." 05 Mar. 2004. <http://members.cox.net/~harlowk22/atgwsobj.html>.

James, Bill. "Regional Valuation." E-mail to the author. 05 Apr. 2009.

Lederer, Rich. "Abstracts from the Abstracts." The Baseball Analysts. 14 Nov. 2004. <http://baseballanalysts.com/archives/2004/11/abstracts_from_20.php>.

"Major League Baseball draft Q&A." 03 June 2003. USA Today. <http://www.usatoday.com/sports/baseball/draft/2003-06-03-draft-qanda.htm>.

Marcano, Arturo. "Worldwide Draft." The Latin Insider. <http://baseballguru.com/articles/analysismarcano01.html>.

Marcos, Frank. "Major League Baseball Scouting Bureau Q&A." Major League Baseball. <http://mlb.mlb.com/mlb/official_info/about_mlb/scouting_overview.jsp>.

McAdam, Sean. "Scouts Must Dig Deeper Than Tools." The High School Baseball Web. 12 Apr. 2009 <http://www.hsbaseballweb.com/pro-scouting/scouts_dig_deep.htm>.

Munsey & Suppes. "Map of current MLB ballparks." Map. 02 Apr. 2009 <http://www.ballparks.com/baseball/general/maps/current.htm>.

Ohlmann, Jeffrey. Optimization of Player Selections in MLB Draft. 16 July 2008. University of Iowa. <http://pages.stern.nyu.edu/~jsimonof/Ohlmann.pdf>.

Press, Associated. "National pastime takes on international flavor." 03 Apr. 2007. ESPN. <http://sports.espn.go.com/mlb/news/story?id=2824295>.

Studeman, Dave. "Win Shares have Arrived." The Hardball Times. 16 May 2004. <http://www.hardballtimes.com/main/article/2004-win-shares-have-arrived/>.

Toms, Pete. "RULE IV DRAFT -- What is the Future of MLB Player Development?" The Biz of Baseball. 16 Nov. 2008.

<http://www.bizofbaseball.com/index.php?option=com_content&view=article&id=2621>.

Weisman, Jon. "No Rumors - Just Some Lo Duca Thoughts." Dodger Thoughts. 30 July 2004. <http://dodgerthoughts.baseballtoaster.com/archives/014740.html>.

Appendix A: Data Compilation (Example)

Year	Team	Change in Scouts From Prior Year	Change in Win-Shares from Prior Year	League (0-American, 1-National)	Region Location	First Pick Position	First-Pick Win Shares
1987	Atlanta	None	None	1	5	6	29
1987	Baltimore	None	None	1	2	7	0
1987	Boston	None	None	0	1	26	5
1987	Chicago Cubs	None	None	1	4	4	27
1987	Chicago White Sox	None	None	0	4	5	80
1987	Cincinnati	None	None	1	4	18	22
1987	Cleveland	None	None	0	4	47	88
1987	Detroit	None	None	0	4	21	1
1987	Houston	None	None	1	6	22	115

Prior Year Winning Percentage	Number of Scouts for Draftable Players	Team Payroll (adjusted for inflation)- 2008 USD	MLB Players from Draft	Total Years Played to Date-Signed Players	Cumulative Six Year Win-Shares Draft Value
0.447	37	$27,332,247	5	44	104
0.451	23	$25,526,301	10	64	295
0.590	22	$25,102,559	9	20	88
0.438	44	$24,639,249	8	41	225
0.444	27	$17,076,089	8	35	99
0.531	44	$16,219,350	9	40	171
0.519	17	$15,114,975	7	23	140
0.537	22	$23,859,974	8	36	144
0.593	29	$21,865,905	8	42	236

Appendix B: Win-Shares Calculation[28]

- First, you divide responsibility for a team's wins between the offense (batting and baserunning) and defense (pitching and fielding). You do this by calculating the team run differential through a method James calls Marginal Runs. You first calculate the average number of runs scored per team in the league. You next adjust your team's runs scored and runs allowed for the ballpark in which they played half their games (i.e. home games). Then you add together two figures: all runs scored over 52% of the league average (credited to the offense), and all runs allowed less than 152% of the league average (credited to the defense). This is total marginal runs.

- Next, you take the percent of marginal runs contributed by the offense, multiply it by the number of wins times three. This is the total number of offensive Win Shares. You do the same thing for defensive Win Shares.

- Next, you attribute offensive Win Shares to individual players. This is done through two key metrics: Runs Created and Outs Made. Runs Created is a formula built by James and refined over the years. It starts with the basic equation of OBP times total bases and then adds player credit for other factors, including stolen bases, caught stealing, grounding into double plays, batting average and home runs with runners in scoring position and the kitchen sink. Runs Created is calculated for every single batter, including pitchers (if they're in the National League).

- Next, you subtract the league "background" Runs Created (52% of the league average) from each player's Runs Created based on the number of Outs Made by that batter, adjust it for ballpark, and credit each player with the result; essentially individual marginal runs created. Add these up for all players and use each player's percentage of the whole to allocate offensive Win Shares to each. Note that any player whose Runs Created are less than 52% of the league average runs created per out is credited with no Win Shares. This doesn't happen very often (except for pitchers).

- That was the easy part. Now you've got to deal with the defense. The first step is to divide defensive Win Shares between pitching and fielding. This done through a complicated formula that accounts for FIP elements that can be attributed only to pitchers (home runs, walks and strikeouts) as well as a team's DER (Defensive Efficiency Ratio, adjusted for the ballpark) and other fielding statistics such as passed balls, errors and double plays. Typically, about 70% of defensive Win Shares are credited to pitching, and 30% to fielding. The Win Shares system is bound so that pitching never is credited with less than 60%, or more than 75%, of defensive Win Shares.

- Next, you allocate pitching Win Shares to individual pitchers. This is accomplished through an even more complicated formula that starts with each pitcher's marginal runs not allowed (same approach as team marginal runs not allowed), wins, losses and saves. Special consideration is given to relievers by estimating the number of high-leverage

[28] http://www.baseballgraphs.com/main/index.php/site/details/

innings they pitched (ninth innings with one-run leads are more important than first innings with no score) and something called "Component ERA" which is essentially ERA re-calculated according to the actual underlying run elements.

- Finally, pitchers are deducted Win Shares if they are absolutely lousy hitters. Call this the "Dean Chance" factor. All these elements are then mixed together in a complicated formula to allocate pitching Win Shares to individual pitchers. As in offensive Win Shares, any pitcher who gives up more than 152% of league-average Runs Scored (adjusted for ballpark) does not receive any credit for pitching Win Shares.

- One note: responsibility for unearned runs is split 50/50 between pitching and fielding.

- Which leads us to the next, most complicated step: allocating fielding Win Shares to fielding positions, and then to individual fielders. The calculations differ for each position. Essentially, James has selected four defensive statistics to evaluate positions. Here they are by position, listed in order of importance:
    - Catchers: Caught Stealing, Errors, Passed Balls and Sacrifice Hits Allowed
    - First Basemen: Plays Made, Errors, Arm Rating and Errors by third basemen and shortstops
    - Second Basemen: Double Plays, Assists, Errors and Putouts
    - Shortstops: Assists, Double Plays, Errors and Putouts
    - Third Basemen: Assists, Errors, Sacrifice Hits Allowed and Double Plays
    - Outfielders: Putouts, Team DER, Arm Elements and Assists and Errors

- Lots of things to note about the fielding calculations.
    - First, the statistics are adjusted based on the number of innings a lefthander pitches for the team, which has an impact on which side of the field batters hit the ball to.
    - Second, these stats are calculated as a proportion of the team's total, divided by the league-average proportions of the total. In other words, if a shortstop has 50 assists and his team has 100 assists in total, he receives just as much credit as the shortstop who has 100 assists and plays on a team with 200 assists in total. This is important, because it adjusts the fielding stats for the fact that fielders may be playing behind pitchers with certain tendencies such as giving up more ground balls vs. fly balls.
    - Third, double plays are only factored in as a proportion of potential double plays. If teams don't have a lot of runners on first, they have less of a chance to turn double plays, and Win Shares takes this into account.
    - Fourth, team DER is used to credit outfielders with fielding Win Shares because it is James' observation that outfielders have a much larger impact on DER than infielders. James acknowledges that there is some "circular logic" here.
    - Fifth, there is a final element included in the formula to allocate fielding Win Shares to individual fielders. This element is called "Range Bonus Play." It particularly impacts outfielders in the following manner: if one outfielder handles more opportunities per inning played than the other outfielders on the team, he will be credited with more fielding Win Shares.

Appendix C: Number of Scouts on Win-Shares Regression

Number of Scouts versus Win-Shares

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.191009856
R Square	0.036484765
Adjusted R Square	0.031466456
Standard Error	77.89260586
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	44110.97541	44110.98	7.270331188	0.007633259
Residual	192	1164913.545	6067.258		
Total	193	1209024.521			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	175.642553	17.5716196	9.995809	3.31486E-19	140.9843529	210.3007531	140.9843529	210.3008
Number of Scouts for Draftable Players	-1.474890873	0.546994287	-2.696355	0.007633259	-2.553780462	-0.396001284	-2.553780462	-0.396001

Dummy Ranges Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.216687008
R Square	0.04695326
Adjusted R Square	0.041989475
Standard Error	77.4683026
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	1	56767.64	56767.64	9.459164	0.002407
Residual	192	1152257	6001.338		
Total	193	1209025			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	173.397793	14.94747	11.60048	6.37E-24	143.9155	202.8801	143.9155	202.8801
Dummy Ranges	-17.17463038	5.584201	-3.075575	0.002407	-28.18889	-6.160373	-28.18889	-6.160373

Appendix D: Financial Effects on Draft Valuation

Payroll versus Win-Shares

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.201118
R Square	0.040448
Adjusted R Square	0.035451
Standard Error	77.73223
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	1	48902.95	48902.95	8.093432	0.004925
Residual	192	1160122	6042.3		
Total	193	1209025			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	157.9265	11.07051	14.26551	5.92E-32	136.0911	179.7619	136.0911	179.7619
Team Payroll (adjusted for inflation)- 2008 USD	-4.81E-07	1.69E-07	-2.844896	0.004925	-8.14E-07	-1.47E-07	-8.14E-07	-1.47E-07

Operating Income (1989-1999) versus Win-Shares (1989-1999)

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.146817
R Square	0.021555
Adjusted R Square	0.015661
Standard Error	79.07463
Observations	168

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	1	22866.6	22866.6	3.657019	0.057556
Residual	166	1037964	6252.798		
Total	167	1060831			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	131.7434	6.634836	19.85632	1.04E-45	118.6439	144.843	118.6439	144.843
X Variable 1	-0.737124	0.385458	-1.912333	0.057556	-1.498155	0.023908	-1.498155	0.023908

Multiple Regression (Payroll and Operating Income)

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.241872
R Square	0.058502
Adjusted R Square	0.04709
Standard Error	77.802
Observations	168

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	2	62061.03	31030.52	5.126341	0.00692
Residual	165	998770	6053.151		
Total	167	1060831			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	162.4418	13.71703	11.84234	8.28E-24	135.3582	189.5253	135.3582	189.5253
Team Payroll	-4.8E-07	1.88E-07	-2.544611	0.011857	-8.52E-07	-1.07E-07	-8.52E-07	-1.07E-07
Operating Income	-0.879741	0.383373	-2.294738	0.023008	-1.63669	-0.122791	-1.63669	-0.122791

Appendix E: General Manager Regression

Played Professionally or Not

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.163809
R Square	0.026833
Adjusted R Square	0.021765
Standard Error	78.28176
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	1	32442.13	32442.13	5.294053	0.022472
Residual	192	1176582	6128.033		
Total	193	1209025			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Jpper 95.0%*
Intercept	139.6947	6.839509	20.42466	5.05E-50	126.2044	153.1849	126.2044	153.1849
GM-Played in MLB or Not?	-27.61529	12.00205	-2.300881	0.022472	-51.2881	-3.942487	-51.2881	-3.942487

Tenure in Player Development/Scouting Roles

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.20896
R Square	0.043664
Adjusted R Square	0.038683
Standard Error	77.60186
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	1	52791.1	52791.1	8.766303	0.003455
Residual	192	1156233	6022.049		
Total	193	1209025			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Jpper 95.0%*
Intercept	111.4145	8.578264	12.98801	4.31E-28	94.4948	128.3343	94.4948	128.3343
Tenure in Development	4.099102	1.38446	2.960794	0.003455	1.368397	6.829806	1.368397	6.829806

Appendix F: Win-Loss Records Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.070098
R Square	0.004914
Adjusted R Square	-0.000269
Standard Error	79.15846
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	1	5940.775	5940.775	0.948088	0.331432
Residual	192	1203084	6266.061		
Total	193	1209025			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	158.6605	29.24581	5.425069	1.73E-07	100.9762	216.3449	100.9762	216.3449
Prior Year Winning Percentage	-57.03881	58.57957	-0.973698	0.331432	-172.5809	58.50334	-172.5809	58.50334

Appendix G: Multiple Regression All Variables

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.325176666
R Square	0.105739864
Adjusted R Square	0.08195635
Standard Error	75.83515269
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	5	127842.1	25568.42	4.445931	0.00075
Residual	188	1081182	5750.97		
Total	193	1209025			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	193.9825474	32.3215	6.001656	9.86E-09	130.2231	257.742	130.2231	257.742
League (0-American, 1-National)	-24.73999995	11.38788	-2.172486	0.03107	-47.20444	-2.275564	-47.20444	-2.275564
First Pick Position	-0.714310631	0.391471	-1.824682	0.069636	-1.486551	0.05793	-1.486551	0.05793
Prior Year Winning Percentage	15.702914	61.47044	0.255455	0.798651	-105.5575	136.9634	-105.5575	136.9634
Number of Scouts for Draftable Players	-0.832539168	0.563514	-1.477407	0.14124	-1.944161	0.279083	-1.944161	0.279083
Team Payroll (adjusted for inflation)- 2008 USD	-3.49884E-07	1.79E-07	-1.956635	0.051871	-7.03E-07	2.87E-09	-7.03E-07	2.87E-09

Appendix H: Multiple Regression-Adjusted Win-Share

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.356189841
R Square	0.126871203
Adjusted R Square	0.103649692
Standard Error	40.43144175
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	5	44656.13	8931.225	5.463521	0.000102
Residual	188	307323.9	1634.701		
Total	193	351980			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	66.7871227	17.23218	3.875721	0.000147	32.79384	100.7804	32.79384	100.7804
League (0-American, 1-National)	-15.43754097	6.071435	-2.542651	0.011807	-27.41443	-3.460647	-27.41443	-3.460647
First Pick Position	-0.349029392	0.208713	-1.672297	0.096129	-0.760749	0.06269	-0.760749	0.06269
Prior Year Winning Percentage	7.935291628	32.77291	0.24213	0.808944	-56.7146	72.58519	-56.7146	72.58519
Number of Scouts for Draftable Players	-0.024535619	0.300437	-0.081666	0.934999	-0.617196	0.568125	-0.617196	0.568125
Team Payroll (adjusted for inflation)- 2008 USD	-3.1686E-07	9.53E-08	-3.323567	0.001068	-5.05E-07	-1.29E-07	-5.05E-07	-1.29E-07

Appendix I: International Scouts Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.087092
R Square	0.007585
Adjusted R Square	0.002416
Standard Error	42.65353
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regression	1	2669.79	2669.79	1.467463	0.227235
Residual	192	349310.2	1819.324		
Total	193	351980			

	Coefficients	*tandard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Upper 95.0%*
Intercept	40.72214	3.995454	10.19212	8.97E-20	32.84152	48.60276	32.84152	48.60276
Number of Scouts for Non-Draftable Players	-0.879148	0.725735	-1.211389	0.227235	-2.310586	0.55229	-2.310586	0.55229