

An Examination of the Contributed Major League WAR by an Organization's Top Prospects and
Baseball America Farm System Rank

Landon K Adams

Missouri Southern State University

Table of Contents

Abstract.....	4
Introduction.....	5
Literature Review.....	6
Player Valuation.....	7
Farm System Valuation.....	12
Predictive Variables of Major League Success	13
Conclusions.....	14
Methodology.....	17
Best Subsets Approach to Model Building.....	18
Determine Independent Variables.....	18
Data Collection	21
Possible Models	21
Perform Regression Analysis.....	23
Test Models for Overall Significance.....	24
Evaluate Fit of Model	24
Test for Significance of Slope.....	25
Select the Best Model	25
Results.....	26
Test for Overall Significance	27
Evaluate Fit of Models.....	29
Test for Significance of Slope.....	30
Select the Best Model	32

Conclusion32

Future Research32

References.....35

Abstract

This study created a model that is designed to predict expected aggregate WAR of a farm system in the subsequent eight seasons following a given ranking. The study will consider every farm system from the 2001 to 2004 seasons. The Top 30 prospects will come from the annual Baseball America Prospect Handbook. Each of these prospects aggregate WAR were summed up and tested against several variables using the best subsets approach to model building. The study included 120 observations, six variables and the aggregate WAR of 3,600 prospects. Thirty-two of the sixty-three models were found to be statistically significant. The final model accounted for 9.3% of the variation among aggregate prospect WAR in the subsequent eight seasons.

Today in Major League Baseball there is an uneven playing field. Unlike in the National Football League, where the league prevents teams from obtaining vast competitive advantages, thanks to the markets they play in, baseball operates differently. Baseball doesn't share the revenue like the NFL does. This means that due to differences in local television markets, certain teams have huge economic advantages over other teams. The gap is large now, and it is only widening. As a result, large market teams are able to flex their financial muscle over small market teams, allowing them to attract better players through Free Agency.

Baseball didn't always have free agency. There was a time when ballplayers were treated more like indentured servants than people. However, former big leaguer Curt Flood changed all that when he refused to allow himself to be traded the issue rose all the way to the Supreme Court, eventually leading to the end of the old system and the beginning of Free Agency in Major League Baseball (Barra 2011). Today, a player is only controlled by the team that developed him until he accrues over six seasons of Major League service time (Euston 2007). After six seasons, players are able to attain free agent status. (As a free agent, a player is allowed to open himself up for auction and shop his talents to any Major League team.) For these reasons it isn't hard to imagine why larger market teams have a huge advantage, after all, a larger market means a larger revenue stream. Players that reach free agency often have multiple teams, who are trying to fill the same holes. As a result, contracts often get bloated as a result. Not only do small market teams have the disadvantage of less revenue to spend on their payrolls, but they also wind up having to pay more for free agents, because other teams have bid up their cost.

For the first three seasons of team control, players are only entitled to the league minimum salary. (Euston 2007). In the three subsequent seasons they are entitled to arbitration. Arbitration is a process in which a judge determines a player's salary based off the expected

market value of the player. The player is then awarded 40% value for his first arbitration eligible season, 60% for his second and then 80% for his third. Simply put, arbitration not only restricts market inflation, but it also depresses player's salaries. It isn't hard to see why baseball and its fans have become so obsessed with prospects and farm system ranking. Common sense would dictate expect that if a team can build an impressive farm system, they should then benefit from that farm system in the following seasons. The opposite can be said for farm systems that are not well regarded around the industry.

Obviously fans can't know for sure which prospects are the favorites of each respective team's scouting departments. However, there are publications that focus huge amounts of energy on prospects and their perceived values around baseball. Baseball America is often considered as the cream of the crop in the prospect publication industry. In every season since 2001, Baseball America's annual Prospect Handbook has been highly anticipated by the baseball community. The Handbook includes both a top 30 for each individual team, as well as organizational rankings. While, the Handbook's scouting reports are not compiled by Major League scouts, they do provide an enormous amount of input into Baseball America's Prospect Handbooks. For this reason Baseball America is viewed by most as the premier baseball prospect rating publication. Every offseason baseball fans eagerly await their opportunity to peer into what has become baseball's crystal ball.

Literature Review

Major League success is a function of many factors. The following is an examination of player valuation, farm system valuation, and other factors that affect Major League Baseball success.

Player Valuation

In 2007, Victor Wang, who is currently an Intern of Baseball and Analytics for the Cleveland Indians, published an article entitled, “How Much is a Top Prospect Worth?” in the August newsletter of the SABR Statistical Analysis Committee. Wang’s research was conducted in a time when teams were growing weary of trading their top prospects for veteran players. Wang wondered “whether teams were making the right choice by holding on to these prospects. (Wang 2007).” Quite simply Wang wanted to find out whether statistically speaking it made more sense to hang on to top prospects, or to deal them via trades for players that had already proven themselves at the highest level of competition (Wang 2007).

Wang studied Baseball America’s top 100 Prospect lists’ top twenty-five from 1990-1999 (Wang 2007). In doing so, Wang established a precedent for which prospect ranking system should be used when studying prospects future success in relation to their ranking as a prospect. Unlike draft pick rankings, prospects can remain prospects for several years. In fact the typical Major League timetable for draft picks is two or three years, three to five years for high school selections, and up to six for international signings. Because prospects don’t lose their prospect status until they exhaust their rookie eligibility at the Major League level, they are often ranked by Baseball America multiple times prior to reaching the Major Leagues. For this reason there were several prospects that found themselves on multiple top 100 lists during Wang’s selected time frame. When this occurred Wang elected to only add them to the sample once, and when he did so he went with the highest ranking that the prospect achieved.

Wang determined the value of Baseball America’s prospect rankings, compared it to an advanced statistic called Wins Above Replacement Player, or WARP for short (Wang 2007) .

WARP was created by Baseball Prospectus writer Clay Davenport who defined the statistic as “the number of wins this player contributed, above what a replacement level hitter, fielder, and pitcher would have done, with adjustments only for within the season (Wang 2007).” WARP provided Wang with a statistic that could be applied to separate positions and player types and give at least a somewhat accurate gauge of how they stack up with one another. The statistic provides a means of determining the contributions of a player relative to their Major League peers. WARP, or WAR as it is now abbreviated, is also easy to access through baseball statistical database sites such as either Baseball Reference, or Fangraphs.

Wang also had to decide what seasons he would compile WARP accumulation in his results. Wang decided that since players are often only under team control for six seasons upon their final promotion to the Major Leagues, he would use the first six seasons of each player’s careers prior to that player reaching free agency (Wang 2007). In doing so, Wang established which seasons should be considered when comparing WARP versus prospect rank. By considering only those seasons, Wang could find specifically the effects each prospect had on the organization that developed them, or the effect each prospect could have had on the organization should they not have traded them or released them prior to the player’s reaching of free agency.

Wang also thought it important to create categories for the various degrees of prospect success. He chose to place prospects into four groups: bust, contributor, everyday player, and star (Wang 2007). Table 1 contains a breakdown of prospect success classifications and WARP qualifications.)

Table 1
Prospect Success Classifications according to Wang

Classification	Qualification	Qualification Per Year
Bust	0-12 WARP	0-2 WARP Per Year
Contributor Everyday	12-24 WARP	2-4 WARP Per Year
Player	24-36 WARP	4-6 WARP Per Year
Star	36+ WARP	6+ WARP Per Year

Wang found that while hitting prospects average about 23.72 WARP and a bust rate of roughly 21%, pitchers were much more unpredictable (Wang 2007). In Wang's study, pitchers averaged 12.91 WARP in their first six seasons, while over half (54%) of the pitching prospects that ranked in the top 100 for Baseball America from 1990-1999 were deemed "Busts" as judged by the criteria Wang created (Wang 2007). In fact over four times as many hitters qualified for the "Star" category as pitchers (Wang 2007).

Wang decided that for comparison sake, he needed to judge the success rates against another study. Since baseball is unique in having an emphasis on a minor league feeder system that is controlled by the team's themselves, it made sense to Wang to compare his study to a study that looked at the success of prospects based on their draft position. The study that Wang, looked at in particular was conducted by Baseball Prospectus writer Rany Jazayerli. In Jazayerli's draft study he attempts to find the value of prospects drafted in the top 100 of Major League Baseball's amateur draft (Jazayerli 2006). Wang, by comparing Jazayerli's study to his own, was able to conclude that "a top prospect has a higher value than a top draft pick. (Wang 2007)"

Wang (2007) clearly established the relationship between Baseball America's prospect ranking system and major league success or top prospects. This implies that a farm system with top prospects provides players that help the major league ball club win games.

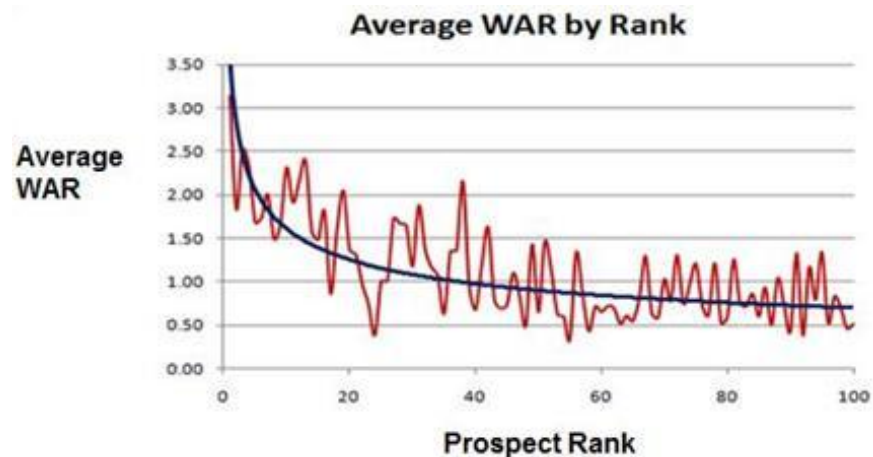
In 2011, blogger Scott McKinney, decided that Baseball America's top 100 prospect lists needed a more in depth study than the one conducted by Wang (2007) just four years earlier. McKinney expanded Wang's study to include every top 100 prospect, quadrupling Wang's sample from four years prior (McKinney 2011). In expanding the sample, McKinney set a precedent, providing the importance studying more prospects in such studies. McKinney also elected to include each multiple instance that a prospect made additional appearances on Top 100 lists because he believed his "goal was to determine the meaning of various rankings by determining their success and failure (McKinney 2011)." McKinney also switched from WARP to the most updated form of the statistic: WAR. Just like the statistic WARP, WAR attempts to account for all factors while giving providing the number of wins that the individual is worth over the course of a season. Using WAR, McKinney broke down the players into categories generally accepted by the baseball community. These can be seen in Table 2.

Table 2
Prospect Success Classifications According to
McKinney

WAR	Contribution	Success vs Bust
Under 0.50	Very poor	Bust
0.50 - 1.49	Below Average	Bust
1.50 - 2.49	Average	Success
2.50 - 3.49	Good	Superior
3.50 - 4.49	Very Good	Superior
4.50 or higher	Great	Superior

McKinney's results were similar to Wang's findings from four years prior. Roughly two out of every five position prospects in the top 100 succeed, however that number drops to just one in four, when it comes to pitching prospects (McKinney 2011). A larger discrepancy exists for star expectancy. Nearly one in four position prospects become stars, while only one out of every ten pitchers achieves star status (McKinney 2011). By studying the entire top 100 list McKinney also provides us with a graph that displays the average WAR by prospect rank. See Figure 1.

Figure 1



The graph shows that the higher the prospect is ranked, the higher the Average WAR that is produced by said prospect. Notice that the difference between prospect rank diminishes the farther down the rankings one goes. The difference between the #1 ranked prospect and #20 ranked prospect is huge, while the difference between the #40 ranked prospect and #60 ranking prospect is not.

Thanks to McKinney's follow up of Wang's study there is much more information regarding the success rates of top prospects. But more importantly in the context of this paper,

McKinney built upon a foundation that had already established the importance of prospect ranking systems, specifically Baseball America's annual rankings. McKinney also established the importance of the expansion of data sets as more information becomes available. An individual prospect's success rate is a great piece of knowledge to have, but ultimately the question is 'how are seeking out how these prospects are contributing at the Major League level?' The advanced statistic WAR answers this question.

Farm System Valuation

In 2010, Sports Illustrated columnist Sky Andrecheck examined farm system ranking. He asked "Does a good (organizational) ranking actually lead to future success, and if so, how much (Andrecheck 2010) Andrecheck (2010) used linear regression to examine the relationship between Baseball American rankings and team winning percentage.

Andrecheck found that a given year's farm system rankings were most accurate when predicting the next season's results and the season's results four years in the future (Andrecheck 2010). This establishes the farm system as having predictive value in wins for the parent team. In seasons occurring two years and seven years ahead, the rankings show a statistically noteworthy predictive power, and in all other seasons a team's farm system ranking has no predictive power whatsoever (Andrecheck 2010).

In his current analysis, Andrecheck combined four years worth of data for the farm system's he was studying. He found that by combining these rankings there is a strong correlation with winning percentage. Andrecheck believing that he was on to something, decided to average the four years of Baseball America's team rankings and ran those numbers in his

original multiple regression model. Andrecheck used his model to predict extra wins for each farm system. Table 3 shows these predictions for the highest and lowest farm systems of 2010.

Table 3
Andrecheck's Net Wins Effect Model

Year	#1 Rangers NetWins Effect	#30 Astros Net Wins Effect
2010	1.2	-3.3
2011	2.6	-4.4
2012	5.2	-6.0
2013	5.6	-5.6
2014	4.9	-4.9
2015	3.8	-3.8
2016	3.3	-3.3
2017	3.1	-3.1
2018	2.1	-2.1

Andrecheck's conclusions based off of the above findings are as follows. The Rangers, who are expected to net themselves thirty-one extra wins over the next nine seasons due to their farm system, are in prime position to be an extremely competitive organization. The opposite can be said for the Houston Astros, whose farm system presents them with a huge disadvantage moving forward. Without the influx of top prospects or other young talent, Andrecheck concluded that the organization would either need to be extremely lucky or extremely successful in trade and the free agent market (Andrecheck 2010). Obviously, the success of an organization's farm system is a critical factor on Major League success.

Predictive Variables of Major League Success

Andrecheck realized that there would be more factors in a team's success (measured by winning percentage) than just the farm system's ranking. In addition to farm system ranking,

Andrecheck used five other variables to help explain team success. They are winning percentage from the previous season, winning percentage from two seasons ago, salary from this season, salary from two seasons ago, and market size. Andrecheck's multiple regression analysis indicates that previous season's winning percentage and salary as the primary variables in predicting a team's season success (Andrecheck 2010).

Conclusions

Wang broke the ice in 2007 when he studied Baseball America's top 25 prospects, McKinney followed by expanding the research to the full Top 100. McKinney's study established the importance of expanding the study of prospects to encompass a larger sample. Both studies utilized an advanced statistic called Wins Above Replacement or WAR for short. This statistic was an obvious choice because it is already normalized by taking into consideration several factors.

In Andrecheck's study, evidence suggests the Major League farm system plays a critical factor in a team's winning percentage. However, how can an individual prospect's success be measured? Luckily, prior to Andrecheck's study Wang and McKinney answered that question in each of their respective studies that were previously discussed. Thanks to these studies a precedent exists to measure prospects success using WAR. By simply combining the thought process of the aforementioned studies there is the following idea. Farm system success is a critical factor for a Major League Baseball team. Farm systems are comprised of many individual prospects. As Wang and McKinney illustrated, individual prospect's success can be measured in WAR. As a result, an entire farm system's success could be measured by the aggregate WAR of each of the farm system's top 30 prospects.

For the first time in history, there is now the opportunity to evaluate Baseball America's top 30 prospects for each organization. The reason for this is that Baseball America began their annual prospect handbooks in 2001. Consider that eight seasons worth of data are necessary to aggregate the WAR of these prospects, there haven't been many Baseball America Prospect Handbooks published that are old enough to study. At this moment, there is four years worth of rankings that can be examined.

Of course, even with Andrecheck's attempts to diminish any extraneous variables affects on the study, it would be impossible to totally achieve this when evaluating prospects affects on team winning percentage. What are the contributions of top prospects on the Major League level? Wang and McKinney have established a precedent for the study of individual prospect success, however, they were only able to look at the top twenty-five and then eventually the top 100 prospects overall. This data set wouldn't allow the study of how each ranking position's prospects contributed at the Major League level. Luckily in 2001, Baseball America began publishing the Prospect Handbook.

In the prospect handbook Baseball America published both the overall team rankings in terms of prospect wealth, as well as the top 30 prospects for each individual organization. Thanks to these lists, for the first time in history, it would now be possible to judge an entire farm system and its respective team ranking, based off of the prospects it was judged on. In order to achieve this, Wang's study would need to be expanded yet again. (Keep in mind, McKinney has already established the importance of study expansion in 2011, when he studied all of the top 100 prospects instead of just the top twenty-five.) In order to evaluate farm system's Major League contributions, the study would need to be as comprehensive as possible. Thanks, to the annual Baseball America Prospect Handbook, each ranking position could now be evaluated

based off the Major League performance of each of its top 30 prospects. Thanks, to the timing of this study, it is now the first ever opportunity to conduct such a comprehensive study of Baseball America's Farm System Rankings.

Another issue with Andrecheck's study is sample size. Andrecheck combined four years worth of prospect rankings in order to find any substantial results. Thanks to the Prospect Handbook, each of the 2001, 2002, and 2003 Farm System Rankings can now be analyzed. The reason that the study cannot be extended to include 2004, 2005 and so on, is because after 2003, there aren't enough subsequent seasons worth of data to analyze the contribution at the Major League level of the prospects.

Finally, a defined method of evaluation must be used. In Wang and McKinney's studies they were researching the contribution of top prospects at the Major League level and then assigning a success rate for each prospect. However, in terms of an entire farm system, both the contribution of the prospects at the Major League level, and the systems success rate for producing various types of Major Leaguers could be deemed critical pieces of information. In order to find each of the pieces of information the other piece will come hand in hand, due to the nature of the research involved. Because farm systems will be examined on the whole, just as Andrecheck did in his 2010 study, the numbers of each individual prospect will need to be boiled down into one number to represent the total contribution of the farm system. Once this final contribution is realized it can then be easily compared against other farm systems of the same season, or any of the other seasons that will be examined.

The statistic that would need to be used in this study needs to be easily obtainable, as well as widely accepted as a means of evaluation. At the time of Wang's study the statistic of choice

was WARP (Wins Above Replacement Player), several years later and more recently McKinney elected to upgrade to WAR (Wins Above Replacement). This again is what McKinney had to say regarding this decision: “While [WARP] is conceptually similar to WAR, it is calculated using significantly different methodology, and WAR is widely regarded as a superior measure. While WARP has improved in recent years, the 2007 version of the statistic was significantly flawed.” Since McKinney’s study was conducted most recently, and because WAR can be obtained easily compared to WARP, WAR is the clear choice as the statistic that will be used to boil down the contributions of a farm system into one number.

Wang, McKinney, and Andrecheck all studied methods to evaluate major league baseball prospects, farm systems and their effect on major league ball club’s success. There is evidence that a relationship exists between farm system ranking and winning at the major league level (Andrecheck 2010). Additional research (Andrecheck 2010) indicates that Major League team success is a function of six variables: winning percentage from the previous season, winning percentage from two seasons ago, salary from this season, salary from two seasons ago, and market size (Andrecheck 2010). Note that most of the variables in Andrecheck’s model are arguably related to the market in which the team plays.

Methodology

The purpose of this research is to develop the best regression model to explain the aggregate WAR of an organization’s Top 30 prospects in the eight subsequent seasons following their prospect ranking. The best subsets approach for building linear regression models was used. This section contains a discussion of the best subsets approach followed by sections outlining: determining the independent variables, data collection, possible models, performing the

regression analysis, test models for overall significance, evaluate fit of each model, test for overall significance, and selecting the best model.

Best Subsets Approach to Model Building

The best subsets approach to model building was selected as this study requires a comprehensive approach to determine the best regression model to explain WAR. The best subsets approach to model building looks at every possible regression model for a set of independent variables (Levine, Stephan, Krehbiel & Berenson, 2008).

To be considered, a model had to be statistically significant, and have at least one slope parameter that was significantly different than zero. Models also need to explain as much variation as possible as measured by adjusted R^2 .

Determine Independent Variables

Six variables were determined to be relevant enough to include. These were: Baseball America farm system rank, Major League payroll rank, Major League winning percentage, Minor League winning percentage, rank among organizations in Minor League winning percentage, position of the top 30 prospects.

Baseball America farm system rank. According to Andrecheck (2010), Baseball America's farm system rankings have a "statistically significant effect on the winning percentage for the next eight years." The statistic Wins Above Replacement (WAR) attempts to quantify the wins contributed to a team by a single player. If a farm system significantly impacts the winning percentage of a team for the eight subsequent season and winning percentage is a reflection of WAR, then one would expect that farm system ranking has an impact on the WAR accumulated

by an organization's top 30 prospects over the following eight seasons. Since an organization's farm system ranking is a direct reflection of the perceived quality of an organization's top 30 prospects, one would expect these two variables to be tied together.

There are several publications that rank prospects and farm systems, however since it has been established Baseball America as the premier prospect ranking publication, and since the study is aggregating the WAR of Baseball America's Top 30 prospects, Baseball America's farm system rankings are the most relevant to the study.

Major league payroll rank. In Andrecheck's "The Value of a Good Farm System," it was determined that Major League payroll is a significant factor on a team's winning percentage (the measure of success used by Andrecheck). If Major League payroll affected the success of a team in terms of winning percentage (Andrecheck 2010), then one could surmise that it would affect the success of a team in terms of the WAR accumulated by an organization's top 30 prospects.

In Andrecheck's study, the Major League Payroll was averaged between the previous season and the season in which he conducted the study. Payroll will be approached in the same manner and the year's payroll prior to the Baseball America Prospect Handbook and the year's payroll the handbook was released will be averaged.

Major league winning percentage. When determining the contributing factors of team success in terms of winning percentage, Andrecheck determined that the previous two season's winning percentages were critical in predicting future success (2010).

Since an organization's winning percentage also determines the draft position of Major League baseball teams, hypothetically it would factor into the quality of prospects in an

organization. In this study, the winning percentage from each of the previous three seasons will be averaged.

Minor league winning percentage. Since Andrecheck (2010) found that previous Major League winning percentages were critical in predicting future team success, and because this study is looking at the success of Minor League prospects, it would make sense that the success of Minor League prospects at the Major League level could be predicted by the success in winning percentage that the prospects achieved at the Minor League level. The minor league winning percentage used will not be an average of the previous three seasons, but instead will just be the collective winning percentage of all minor league teams in a given farm system from the previous season.

Rank among organizations in minor league winning percentage. This variable is another representation of the Minor League winning percentage of an organization in the previous season. This variable puts minor league winning percentage into a context that allows it to be compared to other organizations across seasons.

Position of the top 30 prospects. In 2007, Wang established that the higher the ranking of a prospect, the greater the WAR that prospect would produce at the Major League level. In the same study, Wang also found that the expectations should vary depending on whether or not the prospect was a hitter or pitcher. Wang also found that hitting prospects of different positions achieved higher/lower success rates.

In order to consider prospect's positions in this study a numerical scale was developed. Here is the numerical scale that was used: Pitcher = 1, Catcher & Shortstop = 2, Second Baseman & Third Baseman = 3, Outfielder = 4, First Baseman = 5 . The spectrum was then

assigned a number 1-5, the lower the number the farther left the player fell on the defensive spectrum. The values were then averaged, giving each organization one number to describe the average position of its top 30 prospects.

Data Collection

The Baseball America Top 30 lists, as well as the minor league winning percentage, minor league winning percentage rank and prospect positions were all available in the Baseball America Prospect Handbook. The Wins Above Replacement data was collected from FanGraphs.com and put into a table for each organization in the study. The Major League winning percentages were collected from BaseballReference.com and placed into a database to find the average of the previous three seasons. Finally, the payroll information was collected from SteveTheUmp.com and placed into a database.

Possible Models

In the best-subsets approach to model building, each model must be ran to determine which models best explain the independent variables (Levine, etc. 2008). Each model has a different combination of the six variables used in the study. There was one model utilizing all six variables, six models utilizing five variables, 15 four variable models, 20 three variable models, 15 two variable models, and 6 one variable models. There were a total of sixty-three models developed in order to complete the best-subsets approach to model building. The models developed are shown in Table 4.

Table 4
The Variables Used in Each Model

Model	Sig Model	x1	x2	x3	x4	x5	x6
1	Yes	x	x	x	x	x	x
2	No		x	x	x	x	x
3	Yes	x		x	x	x	x
4	Yes	x	x		x	x	x
5	Yes	x	x	x		x	x
6	Yes	x	x	x	x		x
7	Yes	x	x	x	x	x	
8	No			x	x	x	x
9	No		x		x	x	x
10	No		x	x		x	x
11	No		x	x	x		x
12	No		x	x	x	x	
13	Yes	x			x	x	x
14	Yes	x		x		x	x
15	Yes	x		x	x		x
16	Yes	x		x	x	x	
17	Yes	x	x			x	x
18	Yes	x	x		x		x
19	Yes	x	x		x	x	
20	Yes	x	x	x			x
21	Yes	x	x	x		x	
22	Yes	x	x	x	x		
23	No				x	x	x
24	No			x		x	x
25	No			x	x		x
26	No			x	x	x	
27	No		x			x	x
28	No		x		x		x
29	No		x		x	x	
30	No		x	x			x
31	No		x	x		x	
32	No		x	x	x		
33	Yes	x				x	x
34	Yes	x			x		x
35	Yes	x			x	x	
36	Yes	x		x			x
37	Yes	x		x		x	
38	Yes	x		x	x		
39	Yes	x	x				x
40	Yes	x	x			x	

41	Yes	x	x		x		
42	Yes	x	x	x			
43	Yes	x	x				
44	Yes	x		x			
45	Yes	x			x		
46	Yes	x				x	
47	Yes	x					x
48	No		x	x			
49	No		x		x		
50	No		x			x	
51	No		x				x
52	No			x	x		
53	No			x		x	
54	No			x			x
55	No				x	x	
56	No				x		x
57	No					x	x
58	Yes	x					
59	No		x				
60	No			x			
61	No				x		
62	No					x	
63	No						x

Note.

Variable	Variable Nickname	
Baseball America Farm System Rank		x1
Average MLB Payroll Rank		x2
Average Major League Winning Percentage		x3
Previous Minor League Winning Percentage		x4
Previous Minor League Winning Percentage Rank		x5
Position of Top 30		x6

Perform Regression Analysis

Data Analysis capabilities of Microsoft Excel were used for the regressions. Microsoft Excel provides reasonably complete regression output that allows one to easily test for

significance of regression, evaluate fit of data to the model and determine if slope parameters are statistically significant.

Test Models for Overall Significance

$$H_0: \beta_i = 0$$

$$H_0: \beta_i \neq 0$$

Each model was tested for overall significance using an F test. The data analysis capabilities of Excel provide enough regression output for one to easily determine if the model is statistically significant or not.

The p-value criterion was used to determine if the model was used to determine if the model was statistically significant. If the p-value associated with the overall F statistic is below the level of significance one rejects the null hypothesis and concludes that there is sufficient evidence to conclude that the model explains a significant amount of variation. All tests were conducted at the .05 level of significance.

Evaluate Fit of Models

Once the statistically significant models were determined adjusted R^2 was used to evaluate the fit of the model to the data.

R^2 , or the coefficient of multiple determination, is a measure of the proportion of variability in a data set that is accounted for by a statistical model (Levine, etc. 2008). R^2 can assume values between zero (no variation explained) to one (100% of variation explained). The closer that R^2 is to one the better the model fits the data.

R^2 can become inflated based on the number of dependent variables included in the model. For this reason, adjusted R^2 was used. Adjusted R^2 takes into account the number of dependent variables and enables the comparison of models with vary numbers of variables included (Levine, etc. 2008). Excel's regression output contains adjusted R^2 . Adjusted R^2 will be recorded for each model that is statistically significant and is interpreted the same as R^2 .

Test for Significance of Slope

If the model is a significant model, each slope parameter will be tested for significance slope. This will be accomplished via the t- test. The hypotheses for this test are as follows (Levine, etc. 2008):

$$H_0: \beta_i = 0$$

$$H_0: \beta_i \neq 0$$

If the null hypothesis is rejected, there is evidence of a linear relationship. The p-value criterion will be used for this hypothesis test.

Select the Best Model

The best model will be statistically significant based off of the test for overall significance. The best model will then have a high adjusted R^2 (relatively speaking), will have statistically significant slope parameters, and be a parsimonious model.

Results

Test for Overall Significance

Each model was tested for overall significance using the F test. The p-value criterion was used to determine if the null hypothesis was rejected or not rejected. If the p-value was below the .05 level, the null hypothesis was rejected and the model was determined to be significant. The basic form of the test for significance of the overall regression models is shown below (Levine etc. 2008).

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{at least one } \beta_j \neq 0 \ (j = 1, 2, \dots, k)$$

k = number of independent variables in the regression model.

Test Statistic: F.

The results of this analysis are presented in Table 5

Table 5

The Models and their p-values

Model	Sig Model	x1	x2	x3	x4	x5	x6	p-value
1	Yes	X	x	x	x	x	x	0.026
2	No		x	x	x	x	x	0.535
3	Yes	X		x	x	x	x	0.016
4	Yes	X	x		x	x	x	0.013
5	Yes	X	x	x		x	x	0.014
6	Yes	X	x	x	x		x	0.015
7	Yes	X	x	x	x	x		0.020
8	No			x	x	x	x	0.392
9	No		x		x	x	x	0.400
10	No		x	x		x	x	0.391
11	No		x	x	x		x	0.421
12	No		x	x	x	x		0.455
13	Yes	X			x	x	x	0.008
14	Yes	X		x		x	x	0.008
15	Yes	X		x	x		x	0.008
16	Yes	X		x	x	x		0.010
17	Yes	X	x			x	x	0.007

18	Yes	X	x		x		x	0.007
19	Yes	X	x		x	x		0.009
20	Yes	X	x	x			x	0.007
21	Yes	X	x	x		x		0.011
22	Yes	X	x	x	x			0.012
23	No				x	x	x	0.271
24	No			x		x	x	0.249
25	No			x	x		x	0.274
26	No			x	x	x		0.299
27	No		x			x	x	0.256
28	No		x		x		x	0.285
29	No		x		x	x		0.309
30	No		x	x			x	0.720
31	No		x	x		x		0.305
32	No		x	x	x			0.356
33	Yes	X				x	x	0.003
34	Yes	X			x		x	0.004
35	Yes	X			x	x		0.004
36	Yes	X		x			x	0.003
37	Yes	X		x		x		0.005
38	Yes	X		x	x			0.005
39	Yes	X	x				x	0.003
40	Yes	X	x			x		0.004
41	Yes	X	x		x			0.005
42	Yes	X	x	x				0.005
43	Yes	X	x					0.002
44	Yes	X		x				0.002
45	Yes	X			x			0.002
46	Yes	X				x		0.002
47	Yes	X					x	0.001
48	No		x	x				0.794
49	No		x		x			0.210
50	No		x			x		0.170
51	No		x				x	0.604
52	No			x	x			0.196
53	No			x		x		0.162
54	No			x			x	0.511
55	No				x	x		0.170
56	No				x		x	0.162
57	No					x	x	0.141
58	Yes	X						0.000
59	No		x					0.787
60	No			x				0.508

61	No	x	0.080
62	No	x	0.061
63	No	x	0.370

Note.

Variable	Variable Nickname
Baseball America Farm System Rank	x1
Average MLB Payroll Rank	x2
Average Major League Winning Percentage	x3
Previous Minor League Winning Percentage	x4
Previous Minor League Winning Percentage Rank	x5
Position of Top 30	x6

Thirty-two models were statistically significant at the .05 level. There were two additional models found to be statistically significant at the .10 level. It should be noted that all models significant at the .05 level contained the independent variable Baseball America farm system ranking. The two models significant at only the .10 level were Model 61 and Model 62. Model 61 contained the variable previous minor league winning percentage and Model 62 contained the variable previous minor league winning percentage rank. Neither model contained Baseball America farm system rank.

Evaluate Fit of Models

Adjusted R^2 was used to assess the fit of the statistically significant models to the data.

Table 6 shows all of the statistically significant models ranked in descending order by adjusted R^2 .

Table 6
The Models' adjusted R^2

Model	Sig Model	p-value	R^2 adj.
47	Yes	0.001	0.093
58	Yes	0.000	0.093
39	Yes	0.003	0.091
46	Yes	0.002	0.089
36	Yes	0.003	0.088
33	Yes	0.003	0.088
43	Yes	0.002	0.088
44	Yes	0.002	0.087
45	Yes	0.002	0.087
34	Yes	0.004	0.087
17	Yes	0.007	0.085
18	Yes	0.007	0.084
35	Yes	0.004	0.084
40	Yes	0.004	0.083
20	Yes	0.007	0.083
37	Yes	0.005	0.082
14	Yes	0.008	0.082
41	Yes	0.005	0.082
13	Yes	0.008	0.082
15	Yes	0.008	0.081
38	Yes	0.005	0.081
42	Yes	0.005	0.080
19	Yes	0.009	0.079
4	Yes	0.013	0.078
16	Yes	0.010	0.077
5	Yes	0.014	0.077
6	Yes	0.015	0.076
21	Yes	0.011	0.076
3	Yes	0.016	0.075
22	Yes	0.012	0.074
7	Yes	0.020	0.071

1	Yes	0.026	0.070
62	No	0.061	0.021
61	No	0.080	0.017

Notice that Models 47, 58, and 39 have adjusted R^2 over .9. Also, Models 46, 36, 33, 43, 44, 45, 34, 17, 18, 35, 40, 20, 37, 14, 41, 13, 15, 38, 42 have adjusted R^2 over .8.

Tests for Significance of Slope

The slope parameters of each statistically significant model were also tested to determine if they were significantly different than zero (Levine, etc. 2008).

$$H_0: \beta_i = 0$$

$$H_0: \beta_i \neq 0$$

Variables with statistically significant slope are in bold type in Table 7.

Table 7

The Variables Used in Each Model

Model	Sig Model	x1	x2	x3	x4	x5	x6	p-value	R ² adj.
47	Yes	x					x	0.001	0.093
58	Yes	x						0.000	0.093
39	Yes	x	x				x	0.003	0.091
46	Yes	x				x		0.002	0.089
36	Yes	x		x			x	0.003	0.088
33	Yes	x				x	x	0.003	0.088
43	Yes	x	x					0.002	0.088
44	Yes	x		x				0.002	0.087
45	Yes	x			x			0.002	0.087
34	Yes	x			x		x	0.004	0.087
17	Yes	x	x			x	x	0.007	0.085
18	Yes	x	x		x		x	0.007	0.084
35	Yes	x			x	x		0.004	0.084
40	Yes	x	x			x		0.004	0.083
20	Yes	x	x	x			x	0.007	0.083

37	Yes	x	x	x			0.005	0.082
14	Yes	x	x	x	x		0.008	0.082
41	Yes	x	x	x			0.005	0.082
13	Yes	x		x	x	x	0.008	0.082
15	Yes	x	x	x		x	0.008	0.081
38	Yes	x	x	x			0.005	0.081
42	Yes	x	x	x			0.005	0.080
19	Yes	x	x		x	x	0.009	0.079
4	Yes	x	x		x	x	0.013	0.078
16	Yes	x		x	x	x	0.010	0.077
5	Yes	x	x	x		x	0.014	0.077
6	Yes	x	x	x	x		0.015	0.076
21	Yes	x	x	x		x	0.011	0.076
3	Yes	x		x	x	x	0.016	0.075
22	Yes	x	x	x	x		0.012	0.074
7	Yes	x	x	x	x	x	0.020	0.071
1	Yes	x	x	x	x	x	0.026	0.070
62	No					x	0.061	0.021
61	No			x			0.080	0.017

Note.

Variable	Variable Nickname
Baseball America Farm System Rank	x1
Average MLB Payroll Rank	x2
Average Major League Winning Percentage	x3
Previous Minor League Winning Percentage	x4
Previous Minor League Winning Percentage Rank	x5
Position of Top 30	x6

Note. Variables with Significant Slope Coefficients are in bold script.

Each statistically significant model had at least one statistically significant slope.

Baseball America farm system ranking had statistically significant in all models. No models significant at the .05 level or higher had more than one variable with statistically significant slope. Variable x4 (previous minor league winning percentage) in Model 61(p-value=.080) was significant at the .10 level. Variable x5 (previous minor league winning percentage rank) in Model 62 (p-value=.061) was also significant at the .10 level.

Select the Best Model

There were two models statistically significant at the .05 level with adjusted R^2 of .093. Model 47, which included Baseball America farm system ranking and the Position of the Top 30 Prospects and Model 58, a one variable model only included Baseball America farm system ranking. However, Model 58 is the parsimonious model, because it only utilizes one variable, while Model 47 required two variables. Model 58 is shown below:

$$WAR = 93.714 - 1.17 \text{Baseball America Farm System Rank}$$

Conclusion

The results indicate that Baseball America Farm system rank accounts for roughly 9% of the variation in aggregate WAR, this means is that around 91% of the variation in aggregate WAR is not explained by its relationship with Baseball America Farm System Ranking. This leads one to believe that variables other than those considered in this study account for the majority of variation in WAR. One is left to hypothesize what variables could explain the remaining 91%. Perhaps, the coaching, player development, and training staff likely factor into the remaining 91%. However, there may be no easy way to quantify any of the aforementioned factors. Finally, the most crucial factor could simply be talent change randomness. Sometimes prospects simply don't develop into the players that scouts predicted.

Future Research

One way in which an individual could quantify scout or coach quality could be to find the salaries of individuals employed by an organization. One would surmise that the market has dictated the salaries of professionals in the field and so the higher quality the individual's talents,

the higher the individual would be paid. It would then be possible to see how the sum of Minor League coach salaries, or the sum of scouting expenses explain the aggregate WAR of an organization's Top 30 prospects. It is possible that this information paired with Baseball America farm system rank could explain more than the 9% that is currently explained by Model 58.

In continuation of the study, it would be interesting to create individual models for each subsequent season. Not only could one find the expected sum WAR for the following 8 seasons, but instead the expected WAR in each individual season. This would be a more effective measure in projecting the future success of a farm system. If the majority of the aggregate WAR all occurs late in the eight season stretch, perhaps teams should be aware that prospects need a certain amount of time in order to truly produce as expected.

Another way in which the data that was accumulated for this study could be analyzed is on an individual prospect basis. The success rates for certain levels of prospects could be tested and models created to explain the expected value of a prospect at a given ranking in an organization's Top 30 prospects. This information could be critical for Major League baseball teams. Every season teams make decisions regarding prospects on the trade market, wouldn't it be much more effective if teams had a statistical method of projecting individual prospects? The answer of course, is that a statistical expectancy, while beneficial would be one tool of many used in the evaluation process. However, the more tools available the more educated the Major League organization would be.

Another obvious step for the data would be to create a model considering various prospect farm system rankings. Baseball America isn't the only publication that creates annual farm system rankings. Baseball Prospectus is another example of a brand that has gained

notoriety for its work on baseball's farm systems. There are also prospect gurus such as Keith Law or John Sickels who comprise their own organization rankings on an annual basis. It would be interesting if these rankings collectively could explain more than the 9% that is explained by Baseball America's rankings. Would the addition of more ranking systems allow a consensus and more accurate predictions? Or would the variables lack additional explanatory power?

References

Andrecheck, Sky (2011, November 2). The Value of a Good Farm System. Durham, NC:[blogPost]

Retrieved from http://baseballanalysts.com/archives/2010/01/how_does_the_fa.php

Angeli, E., Wagner, J., Lawrick, E., Moore, K., Anderson, M., Soderlund, L., & Brizee, A. (2010, May

5). General format. Retrieved from <http://owl.english.purdue.edu/owl/resource/560/01/>

Barra, Allen. (2011, July 12) How Curt Flood Changed Baseball and Killed His Career in the Process.

Retrieved from <http://www.theatlantic.com/entertainment/archive/2011/07/how-curt-flood-changed-baseball-and-killed-his-career-in-the-process/241783/>

Baseball Reference. (2011) Retrieved from <http://www.baseball-reference.com/>

Callis, J, & Lingo, W. (Eds.). (2001). *Baseball America Prospect Handbook 2001*. Durham, NC:

Baseball America Inc.

Callis, J, & Lingo, W. (Eds.). (2002). *Baseball America Prospect Handbook 2002*. Durham, NC:

Baseball America Inc.

Callis, J, & Lingo, W. (Eds.). (2003). *Baseball America Prospect Handbook 2003*. Durham, NC:

Baseball America Inc.

Callis, J, & Lingo, W. (Eds.). (2004). *Baseball America Prospect Handbook 2004*. Durham, NC:

Baseball America Inc.

Euston, Jeff (2007, January 30). Service Time. Retrieved from

http://bizofbaseball.com/index.php?option=com_content&task=view&id=662&Itemid=75

Fangraphs. (2011) Retrieved from <http://www.fangraphs.com/>

Jazayerli, Rany (2006, June 2). Doctoring the Numbers. Retrieved from

<http://www.baseballprospectus.com/article.php?articleid=5152>

James, Bill. (1988). *Bill James Baseball Abstract*. New York, New York: Ballantine Books

Levin, D.M., Stephan, D.F., Krehbiel, T.C., Berenson, M.L. (2008). *Statistics for Managers* (5th ed.).

Upper Saddle River, NJ: Pearson Prentice Hall.

Minor League Baseball Timeline. <http://www.minorleaguebaseball.com/milb/history/timeline.jsp>

Scott McKinney. (2011, February 14) Success and Failure Rate of Top MLB Prospects [blog post]

Retrieved from <http://www.royalsreview.com/2011/2/14/1992424/success-and-failure-rates-of-top-mlb-prospects>

Scott McKinney. (2011, March 16) How Long Does It Take For Top Prospects to Succeed In the

Majors? [blog post] Retrieved from

<http://www.beyondtheboxscore.com/2011/3/16/2054577/how-long-does-it-take-for-top-prospects-to-succeed-in-the-majors>

Steve the Ump. (2011) Retrieved from: <http://www.stevetheump.com/Payrolls.htm#2000payroll>

Victor Wang. (2007, August) How Much is a Top Prospect Worth? [online newsletter] Retrieved from <http://www.philbirnbaum.com/btn2007-08.pdf>

Wil McDonald. (2011, January 25) The Meaning of the Number One Ranked Farm System [blog post] Retrieved from <http://www.royalsreview.com/2011/1/25/1956088/the-meaning-of-the-royals-having-the-number-one-ranked-farm-system>