An Examination of the Moneyball Theory: A Baseball Statistical Analysis

by U.S. Sports Academy - Sunday, January 02, 2005


Submitted by: Ehren Wassermann, Daniel R. Czech, Matthew J. Wilson & A Barry Joyner

INTRODUCTION

Money is a very important aspect in almost every professional sport. In professional baseball, there are large (New York Yankees) and small (Oakland Athletics) market organizations that make important decisions based on their economic status. For example, many smaller city market teams, must spend their money wisely to ensure the best outcome; whereas, a larger city market team has more income that is expendable (Lewis, 2003). This money spending process originates during the Major League Baseball player draft held each June. The draft process involves fifty rounds of selections by all thirty teams. Each team gathers their general managers, scouts, and professional consultants to decide which players should be drafted. The higher the draftee the more valuable he is believed to the team. Therefore, the procedure to decide which players should be selected earliest is very important (Lewis, 2003). According to Lewis (2003) there are two main theories that are being used to narrow the selection process.

The first theory is generally considered the "old" scouting theory. Scouts venture out and evaluate players all over the country. They do not pay particular attention to statistics, but rather base decisions on the five tools: speed, quickness, arm strength, hitting ability and mental toughness (Lewis, 2003). Each scout goes through "scout school" and is given a pamphlet on what should be looked for in certain aspects of baseball, such as arm strength, fielding, running, and the most important hitting. For arm, strength evaluation, scouts are instructed to look for players exhibiting a "fluid arm action and easy release" (Major League Baseball, 2001 p. 10). Furthermore arm strength evaluation is conducted with the assistance of a radar gun. In the fielding category, "a strong arm and defensive skills can and do carry a player to the major leagues" (MLB, 2001 p. 10). Also, "a live, active lower body, quick feet, agility, instinct, . . . alertness, are some of the qualities that go into the rankings of a major league infielder" (MLB, 2001 p. 10). Running is commonly judged through a timed 60 yard sprint (Baechle & Earle, 2003). Hitting is the "most difficult of all scouting categories of judgment" (MLB, 2001 p. 11). A general list of guidelines that scouts look for is: (1). Strength, (2). Starting the bat, generating bat speed, (3). Full arm extension and follow through after making contact, (4). Head stays on ball, (5). Lack of fear, butt stays up at plate, (6). Short stride, (7). Top hand is evident upon making contact and follow through, (8). Head of bat does not lag, (9). Aggressive, hits first good pitch, (10). Short strokes, yet ball jumps off bat, (11). Bat goes to ball (Not a swing through a certain arc area and the ball happens to be in that zone) (MLB, 2001 p. 11). Scouts are instructed not to scout performance but to "watch for things that are done mechanically that will eventually bring results and success" (MLB, 2001 p. 13). When a scout sees a player he then gives the player a certain grade. "The evaluated grade of five (5) in any respective category portrays the player as having, or will have, an average skill of major league standards, currently or once he reaches major league competition" (MLB, 2001 p. 14)
The second theory is based on the Oakland A's general manager Billy Beane and is illustrated in a novel by Micheal Lewis entitled Moneyball. The Moneyball theory places no emphasis on the body of the athlete or the physical tools that the athlete possess' (Lewis, 2003). This theory illustrates the simplicity of baseball by asking two questions: Does this player get on base? and Can he hit? According to Lewis (2003), Billy Beane (the inspiration of Moneyball) decided to base his drafting of position players/hitters on certain statistics. His main two statistics included on-base percentage (OBP) and slugging percentage. These two stats combined to form a new statistic called on-base plus slugging (OPS). Another differing aspect in Beane's approach was his lack of emphasis on power (Lewis, 2003). Therefore, Beane believed that power could be developed, but patience at the plate and the ability to get on base could not. Moreover, Beane believed in the notion to select college players who are experienced on a different level than the high school "phenom" who needs to be developed into a player. Beane's theory was created based on the works of a sabermetrician named Bill James. "Sabermetrics is the mathematical and statistical analysis of baseball records" (James, 1982 p. 3). James spent years trying to decipher numbers via the Bill James Baseball Abstract, which in turn, resulted in a specific philosophy on hitters.

James' idea on hitters differs from the draft process of Billy Beane, but Beane adopted his views from James' ideology. When putting together a lineup, managers must decide the best order in which the team has the best chance of winning. To win the game one must score more runs than the opposing team. This thought provokes the question as to why such great importance is placed on batting averages? "People are in the habit of listing their teams offensive statistics according to batting averages rather than in order of runs scored" (James, 1984 p.10). James believes that "a hitter's job is not to compile a high batting average, maintain a high on-base percentage, create a high slugging percentage, get 200 hits, or hit home runs" (James, 2001 p. 329). However, part of a hitter's job from a coach's perspective, is to hit homeruns, singles, doubles, get on base, drive in runs, and steal bases (James, 2001). James believes the job of a hitter is to create runs. "The essential measure of a hitter's success is how many runs he has created" (James, 2001 p. 330). James then developed a formula that allows one to establish created runs:

\[(\text{Hits} + \text{Walks}) \times \text{Total Bases} / \text{At-bats} + \text{Walks} \]

This formula works 90 % of the time and gives a total of the team's actual scored runs within 5 % (James, 2001). From this philosophy, Beane developed his theory. The only way to score runs is to get on base and since walks are such a vital part of the created runs formula, on-base percentage should be closely monitored. Even though this formula is very accurate, additional steps can be taken to improve the accuracy. This new formula accounts for the more minute aspects of meaningful baseball statistics. It works off the simple formula:

\[(A \times B)/ C \]

The A variable adjusts the "on-base" aspect of baseball.

\[ A = \text{hits} + \text{walks} + \text{hit batsmen} - \text{caught stealing} - \text{ground into double play} (H + W + HBP - CS - GIDP) \]

The B variable takes into account the advancement of the player.

\[ B = \text{total bases} + .26 \times \text{hit batsmen and non-intentional walks}, + .52 \times \text{stolen bases, sacrifice hits, and flies} (TB + .26(TBB - IBB + HBP) + .52(SB + SH + SF)) \]

The C variable accounts for opportunity.
C = at-bats + total walks + sacrifice hits and flies + hit batsmen (AB + TBB + SF + HBP) (James, 1984 p. 14)

James believed that "figuring the number of runs created is a great tool to evaluate hitters since a hitter's job is to create runs" (James, 1983 p. 5). Therefore, Beane also placed a major emphasis on what had to be done to create runs and drafted players accordingly.

The difference between these two theories leads to the following questions, what are the optimal attributes of the ideal draft pick? Are young high school prospects with the ideal 5 physical tools more advantageous to draft than the seasoned college player with high offensive Moneyball statistics?

The purpose of this investigation was to answer the question of whether there is a significant difference in on base percentage, slugging percentage and on base + slugging percentage (OPS) between high school and college drafted position players performing at the professional level? It is hypothesized that because of more experience, more rich statistical data, and better competition at the college level, the college baseball players will have better offensive Moneyball statistics than the high school players.

**METHODS**

**Participants**

The participants in this study were 60 professional baseball players. More specifically, thirty high school and thirty college players from the 1997 major league professional amateur draft were selected for participation in this study. The age range of the participants was 18 to 23 years of age. The mean age of the high school players was x=18.3 and the mean age of the college players is x=20.9. The mean age for the entire participant sample is 19.6 years of age.

**Procedure**

A comprehensive internet search was conducted to locate the high school and college players from the 1997 amateur draft. The authors felt that four years was enough time to examine a drafted player's moneyball statistics, as four years is the time when many players move to their highest level of play. By use of the following website (www.sports-wired.com), draft information i.e. the top thirty drafted position players from high school and college Moneyball statistics were obtained. Each player's professional (Major and Minor League) Moneyball statistics (slugging percentage, on-base percentage, and on-base plus slugging) from their rookie year to their 4th year of playing professionally were utilized. Slugging percentage was calculated as (Total Bases divided by At Bats). On base Percentage was calculated as (Hits + Base on Ball + Hit By Pitch) divided by (At Bats + Base on Balls + Hit by Pitch + Sacrifice Flies)

**Results**

Descriptive statistics included the means and standard deviation ranges overall and as a function of both major league and minor league slugging percentage, on base percentage, and OPS. A score was calculated, comparing college and high school players, for each variable using the SPSS 12.0 statistical package. An independent samples T-test was utilized to compare differences between collegiate and high school players. An alpha level of .05 was used for all statistical tests.
The mean and standard deviation for the college and high school player's performances in the major and minor leagues is illustrated in Table 1. An independent T-test revealed a significant difference between college and high school minor league slugging percentage. No significant differences were found when comparing college and high school on base percentage and OPS.

DISCUSSION

The purpose of this study was to compare the top collegiate and high school drafted baseball player's professional offensive Moneyball statistics—slugging percentage, on base percentage, and on base plus slugging (OPS) over a four year period. It was hypothesized that college drafted players would have significantly higher Moneyball related offensive statistics than the high school players. The results did not support the hypothesis in that the only significant difference was between college and high school minor league slugging percentage. These results may contradict some of Beane's Moneyball theory (Lewis, 2003).

Beane postulated in Lewis' (2003) that college players would perform better than high school players. This hypothesis is due to several factors. First, college players are more mature physically, mentally, and emotionally than high school players. This maturity would enable them to handle the stresses that are involved in minor league baseball such as, long bus rides, the occasional slump, and unfamiliarity with surroundings. Secondly, college players play against stronger and more advanced competition more often than high school players. This allows for more experience which may provide a better preparation for professional play. Finally, college players play a longer schedule and usually practice year round. This consistent playing allows for skills to be refined and mastered. Using these facts, Beane decided that college players are a better investment than high school players (Lewis, 2003).

The results may not have supported the hypothesis because both groups of athletes had to make adjustments to professional baseball. The high school players may adapt more easily to new changes because they are younger and may have had less influence from other less experienced coaches; however, college players may have developed a certain approach to hitting from college that contradicts a new approach at the professional level. Therefore, the college players may take a longer time to alter their approach to hitting and thus hindering their productivity at the plate. Another factor may be due to the notion that high school players are usually placed in lower levels of professional baseball than college players, which in turn may even the offensive statistics. Lastly, college baseball players may have the opportunity to gain more experience with the wooden bat when competing in collegiate summer leagues.

The rest of baseball has seemed to take notice of the Billy Beane philosophy of drafting. In the 2003 First-Year Player Draft, more than 70 % of the players drafted through the first twenty rounds were from a four-year college or a junior college (Mayo/MLB.com, 2003). This percentage was "a marked increase compared to the last three years" (Mayo/MLB.com, 2003, p.1). Even though this significant increase in drafting college players seems to be the trend, "there [has been] little statistical data to support doing that" (Newman/MLB.com, 2003, p. 2). Baseball America researched the 1990s draft and announced that 2,115 players signed in the first ten rounds between 1990-97 (Newman/MLB.com, 2003). "The group includes 1,024 collegians, 398 of whom (38.9 %) reached the Majors" and "920 prepsters, 259 (28.2 %) did the same" (Newman/Mlb.com, 2003, p 2-3). It was noted that most of the differences amounts to only limited time in "The Show". However, "further research noted that 90 college players (8.8 %) and 77 high school players (8.4 %) became Major League regulars for at least a few seasons" (Newman/MLB.com, 2003, p.
3). These last numbers correlate with the findings of this study illustrating little difference between the productivity of college players versus high school players.

It is important to note that there were limitations to this study. For example, one relevant limitation was the number of participants used in the study. A more significant result could have been established utilizing the entire draft. With more participants and more statistical data, a better idea of the purpose could have been allocated. Another limitation that needs to be noted is the speed at which certain players are promoted. Some high draft picks (top ten rounds) are quickly promoted to a higher level, regardless of their success at the current level. This is due to the amount of money invested in the athlete. For example, a fourth round shortstop may get a signing bonus of 450,000 dollars while the 38th round shortstop may only get 1,000 dollars.

Consequently, the organization has a tremendous amount of money invested in the fourth rounder and they need him to develop faster (Lewis, 2003). Hence, even though this player may not be physically and mentally ready, the organization wants to see a quick return on its investment. Finally, a major limitation is the amount of playing the athlete does. Each year when the regular season ends, many players face the decision of playing winter ball (Lewis, 2003). Many believe that rest is needed to help the body recover from a long, strenuous season; however, others believe that winter ball allows them to gain an extra advantage over their competition. No matter the limitations there is significant evidence against the Billy Beane philosophy.

What this study attempted to illustrate was how an organization with a low budget produces quality baseball players using a new philosophy unorthodox to the norm of baseball (Lewis, 2003). From a financial standpoint, the authors believe there are two mindsets regarding the lack of significance. Because of the minimal significant differences between college and high school players' "moneyball" statistics, many MLB teams might want to disregard the notion that cheaper "moneyball" college drafted players are better investments because they do not do as well as their high school drafted counterparts. However, even though the comparison is not significant statistically, the statistics may be significant to an organization/coach, which is playing the Moneyball way of baseball. A small market organization may want to pay less for college players who average .432 (slugging percentage), .344 (on base percentage) and .776 (OPS) than pay more for high school players who average .396 (slugging percentage), .332 (on base percentage), .728 (OPS) over a four year time period. Even though slugging percentage is the only significant difference, the college players have better statistics from a baseball playing perspective. This difference may be the rationale as to draft cheaper players based on the Moneyball statistics and play the Moneyball way of baseball, especially for small market teams. More research, both qualitative and quantitative needs to be completed before making a conclusion regarding the Moneyball way of drafting and playing professional baseball. If the Moneyball method is proven as significant, it could revolutionize the baseball industry. The importance of this theory is not only relevant monetarily, but it could institute a new theory to the selection of baseball players. Future research should examine if other organizations are using Beane's philosophy and if they are how this will affect the Oakland organization. Moreover, future research should analyze OPS and Runs Created.

REFERENCES