AMATEUR BASEBALL PLAYERS: WHAT ARE THE RELATIVE MERITS OF ATTENDING COLLEGE VERSUS GOING DIRECTLY TO THE MINOR LEAGUES?

BY

TIFFANY CHERI RICHARDSON

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Recreation, Sport and Tourism
in the Graduate College of the
University of Illinois at Urbana-Champaign 2012

Urbana, Illinois

Doctoral Committee:

Associate Professor Lynn Barnett Morris, Chair
Professor Brad R. Humphreys, Director of Research
Professor Dan Rascher, University of San Francisco
Associate Professor Laura Payne
Associate Professor David Strauser
Lecturer Michael Raycraft

Abstract

The effects of education and on-the-job training on the performance of professional baseball players was studied using a secondary dataset consisting of 112 players and five years of play at the Major League level, classified as hitters from the 1990 through 2000 seasons. The study revealed that higher education significantly shortened the time to reach the Major Leagues, while both education and on-the-job training did not significantly affect the offensive performance as measured by five commonly used measures: batting average (AVG) runs batted in (RBI), on base percentage (OBP), slugging (SLG), and on base + slugging (OPS). The results also demonstrated that a few offensive performance measures significantly affected career length in the major leagues, while other factors (age, education, team drafted, round drafted, performance statistics (AVG, OPS, and years in the minors) did not significantly affect the survival time in Major League Baseball. The findings suggested that the personnel in charge of player evaluation and MLB the general managers should fully explore a player's potential by multiple factors instead of using a generic strategy that is merely based on the generally accepted standard operating process rooted in the history of MLB.

Acknowledgments

I would have never been able to finish my dissertation without the guidance of my committee members, help and support from my friends you are truly my family.

I would like to express my deepest thank you to both Dr. Brad R. Humphreys and Dr. Daniel A. Rascher for their excellent guidance, support and mostly the extreme patience that you have both maintained over these past few years. My research would not have been possible without you both, I appreciate all your support. I would like to thank Dr. Lynn Barnett Morris for championing my cause and being in my corner, when you absolutely did not have to. To Dr. Payne and Dr. Strauser, a special thanks to you both for having the willingness to participate via committee membership at the very last moment.

Michael Raycraft I cannot say enough about you my friend. From my very first footstep on campus you get me, I get you, we got each other, and I appreciate every moment more than you will ever possibly understand.

To all my friends that reminded me constantly that I could complete this astronomical task. The support, the encouragement and the gentle prodding, thank you especially Colleen, Alex and Norm.

Table of Contents

Chapter One: Introduction

The purpose of this dissertation was to examine the relationship between education and player development and the effects of each on the outcomes of the early stages of a professional baseball player's career. The early stages of career for the purpose of this study were defined as those years of service carried out between seasons one and five after being called up to the major league. The intent was to (1) examine the offensive performance statistics of position players who were called up to the major league by conducting a comparative analysis of the differences in human capital between players who made the choice to attend college and those who made the decision to bypass college and immediately enter professional baseball after high school, and (2) how these choices affected their future Major League Baseball (MLB) performance. The primary method of accomplishing this was to utilize two variables that measured one's human capital: first, education, (high school only, junior college or four-year college) and second, on-the-job training. On-the-job training was measured by using the number of years in college and in the minor league. The objective was to determine what types of human capital contributed to the player's professional baseball career once they were called up to the major league and whether the human capital an amateur baseball player received had an effect on any portion of the first five seasons of his major league career while he was still restricted by the reserve clause and not eligible for free agency.

It was outside the scope of this dissertation to examine all players initially drafted and to measure their offensive contribution to their minor league club because only a fraction of these players would ever be called up to begin a MLB career. Therefore, this research only considered those players who were actually selected by a team through the

official draft of MLB and then called up from the minor league to the major league. Moreover, by only considering those players who were called up to the major league, a finer distinction would be drawn between a player's human capital and his contribution to his major league team at varying points throughout the first five years of service to his MLB team.

The empirical analysis incorporated offensive performance data for all Major League Baseball players, with minor exceptions, from the years 1991-2000. This analysis examined the relationship between offensive performance during the first five seasons of service in the major league for players drafted between the years 1989 and 2000 and relevant measures of human capital, including level of education and on-the-job training.

Many people view the game of baseball as a team sport; however, it is truly a game of individual performance, especially in the case of the hitter. A hitter's performance is largely independent of the actions of his teammates but dependent on the defensive skills of the opposing team (Dinerstein, 2007). The primary opportunities to score runs are created by hitters; therefore, the main composition of the data consists of hitters and their offensive statistics.

Finally, the theoretical premise to this research study is Human Capital Theory. The human capital of each player is analyzed – including but not limited to where each individual player received his training, whether he did or did not attend college, and the number of years spent developing the skills the professional baseball player required. These guided the study, generating and answering a myriad of questions regarding various training practices, education, and outcomes.

Human capital Theory was not directly tested; rather, this theory was applied to the research questions, providing direction to identify and analyze certain variables such as education and on-the-job training and how those variables affected the offensive performance of MLB players once they were called up to the major league. According to Human Capital Theory, workers who have a higher level of education and experience tend to command higher wages and tend to be more productive workers. Without examining the individual wages of each player, the objective was to examine the amount of human capital via education and years spent in on-the-job training and the impact on future performance these two types of human capital had on a player, comparing the two paths travelled. Therefore, if the previous research findings are valid and can be applied to MLB, an MLB team should be more inclined to select a college educated player over a non-college educated player, as well as having the expectation that the player will be more productive. This will be discussed in detail in Chapter Two.

Introduction to the Problem

In professional sports, three of the major sports leagues in North America use various methods to ensure that the cost of player development is not solely the burden of its teams. For example, in both the National Football League (NFL) and the National Basketball Association (NBA), teams have typically been beneficiaries of an existing college system that stands on its own financially while providing training at the same time. However, MLB teams have a long tradition utilizing minor league affiliates to provide the majority, if not all, of the training received by its players. These minor league teams are usually subsidized by the major league parent franchise to which they are associated. In order to recuperate player development costs, MLB players must be in

the major leagues for six seasons before they are eligible for free agency. The reserve clause is the period of a player's career when he is under contract to a team and must play for whatever salary that team offers, subject to the league minimum, or not play at all. It is also possible for a player's contract to be purchased before free agency (Depken, 2002). After six years a player is then allowed to offer his services to any team in MLB. Why is this MLB training model so different from that of its counterparts, the NFL and the NBA, in being less dependent on colleges and universities to provide its amateur players with the majority of their player development. There continue to be a number of potential reasons; because there are fewer games played in both the NFL and NBA that leaves time for players to practice; in MLB there are too many games per week for meaningful practices thus they need minor league baseball to provide live games for player to compete therefore college and the minor league are both substitutes for actual practices.

High school players have a choice of two options once drafted, they can enter college and play baseball or enter minor league professional baseball, therefore presenting the buyers with multiple additional screening streams (Class A, Class AA, Class AAA) in which to evaluate talent. This is relevant because, unlike the NFL and NBA, much more time is spent trying to determine which player is going to have success in MLB. In addition, there is a significant cost to running and managing the Minor League system and college programs may have the capability of serving as a reliable minor league training substitute.

Many players enter the draft and then decide either to continue their education and play baseball in college or to sign an MLB contract with a minor league team. When

a player is drafted directly from high school the team that drafted him has an exclusive negotiating right with him. He can choose to not sign a contract with the team that drafted him; however, he cannot sign with another MLB team until he is either redrafted or until he makes it through another draft without being selected. Players who enter the draft but subsequently are not selected become amateur free agents and are free to sign with any interested team. High school players selected in the draft can then either choose to sign with the team that drafted them and lose their eligibility to play in college or choose not to sign with the team and instead attend and play baseball at the collegiate level.

Many players accept the notion that their best time to play professional baseball is immediately after high school while they are still young and possess the ability to continue with the sport they love. Others argue that these amateur athletes have the same opportunity to turn professional after college. If the training received in college is similar or equivalent to that received in minor league baseball, the amateur athlete might continue on to college and diversify his investment in his own human capital by not only receiving on-the-job training as he would in the minor leagues but by adding the education to his arsenal for future use. Investing in an education directly after high school gives the player a greater chance to increase his non-baseball wages for his future, especially if a successful professional baseball career does not materialize. Alternatively, players are also faced with the possibility of suffering an injury in college before they even reach the professional ranks thus giving the player an increased incentive to bypass college and head directly to professional baseball.

In comparison to the NBA and NFL draft rules it is important to note the following differences in draft policy. College basketball has long been utilized by the NBA as its minor league system. College basketball helps young players sharpen their skills and become more productive players. According to Groothuis et al. (2007), college basketball serves as a signaling device and provides more information about a player to NBA scouts. When players leave college early or do not attend they have less experience and a more complicated signal than a player who stays in college. Furthermore, the team that previously made the choice to draft an inexperienced early entrant was taking a risk. In the 2006 National Basketball Association Collective Bargaining Agreement (NBA CBA) the following rule changes were made: the age limit for entering the draft was increased from 18 to 19 years of age; U.S. players had to be at least one year removed from high school and 19 years of age (by the end of that calendar year) before entering the NBA draft; and international player's must have turned 19 during the calendar year of the draft. This age requirement in essence forced amateur basketball players to attend college for at least one year or more, until they reached the minimum age required to be eligible for the NBA draft. Or alternatively, this rule potentially forced them overseas to embark on a new trend, playing in Europe for a year before entering the NBA draft. Another possibility these amateur players might realize is that they were not ready to enter into professional basketball and that staying in school was their best option. Prior to the NBA draft rule changes in 2006, players were drafted once they achieved a certain set of skills. This selection was occurring earlier and earlier in their college careers; for example, both Kobe Bryant and LeBron James were drafted from high school (Goothuis, Hill, & Perri, 2007). However, if they were fortunate enough to be drafted, they signed a

contract and began training with the team immediately, unlike MLB where spending time in the minor leagues was almost guaranteed. It must be mentioned that the NBA does sponsor a development league as well, and according to the most recent NBA collective bargaining agreement, the rules are as follows: during an NBA player's first two seasons in the league (regardless of his age when he entered the league), his team will be permitted to assign him to a team in the NBA Developmental League (NBDL). A player can be assigned to the NBDL up to three times per season. The player will continue to be paid his NBA salary and will continue to be included on his NBA team's roster (on the inactive list) while playing in the NBDL (NBA CBA). Once again, a clear distinction exists in salary – an NBA player in the Development League receives his full salary and does not incur a pay cut or substantial decrease in salary to participate in the Development League; however, in the MLB minor league system the salary scale is minuscule in comparison.

Identical to the NBA, the NFL does not draft high school players under any circumstance. According to ARTICLE XVI of the 2006 NFL collective bargaining agreement:

> No player shall be permitted to apply for special eligibility for selection in the Draft, or otherwise be eligible for the Draft, until three NFL regular seasons have begun and ended following either his graduation from high school or graduation of the class with which he entered high school, whichever is earlier. For example, if a player graduated from high school in December 2006, he would not be permitted to apply for special eligibility, and would not otherwise be eligible for selection, until the 2010 Draft. (NFLPA CBA, 2006, p.46)

The NFL requires players that desire to enter the draft to play college football because it is used as a substitute for not having a minor league system to get players prepared for the level of competition they will face once they reach the NFL. After a successful college football career, a wealth of information is available on each player that is in essence eligible for hire in the NFL via the draft. The NFL draft is the first opportunity each team gets to select players who have been out of high school for at least three NFL seasons and these players traditionally come from colleges and universities. Players whose high school class did not graduate three or more years before the given draft year are not eligible to play in the NFL. In the NFL draft, the majority of drafted players come directly out of college programs as seniors or juniors; though some underclassmen are eligible. Players who are drafted and sign a contract begin training with the team immediately, once again unlike MLB.

Unlike the NBA and the NFL, Major League Baseball has the option of drafting players one of three ways:

- High school players, if they have graduated from high school and have not yet attended college or junior college;
- College players, from four-year colleges who have either completed their junior or senior years or are at least 21 years old;
- Junior college players, regardless of how many years of school they have completed.

Teams select players in reverse order of the previous year's standings, alternating between the American League and the National League. With very few exceptions, the majority of the players drafted are sent directly to the minor leagues. However, if a

player chooses to go to a four year college or university, he must stay in school for a minimum of three seasons before he can then re-enter the draft. A high school player who chooses to play baseball professionally loses his eligibility to play college baseball. A player who is eligible to be selected and is passed over by every club becomes a free agent and may sign with any interested club until the player enters, or returns to, a four-year college full-time, or enters or returns to a junior college (Winfree & Molitor, 2007; MLB.com).

Attending college does not end the possibility of a Major League Baseball career; attending college means possibly greater earning potential for a non-baseball related career. Not all, but some, of these young men may even have the ability to graduate in three years. Moreover, if the athlete does not graduate, and his MLB career does not materialize, returning to school to earn a degree is an achievable goal because most likely the majority of the credits have been completed. It must be mentioned that within the MLB CBA, MLB does offer the College Scholarship Plan. According to this plan, the major league player for whom there is in effect on or after January 1, 1973 a valid and unexpired scholarship under the College Scholarship Plan, may commence or resume his studies under the Plan at any time within two years after his last day of MLB service. If his college studies have not commenced under the Plan within two years after his last day of major league service his scholarship terminates. Otherwise, his scholarship continues unless he fails to attend college for more than two consecutive years after his last day of MLB service without proper reason as set forth in Major League Rule 3(c)(4)(D).

Problem Statement

There are multiple paths available for amateur baseball players to gain entry to the MLB. For the purpose of this study only those entering via the selection process of the MLB draft that are eligible as outlined by MLB rules were included in the study. These players came from one of the following three areas: high school, junior college or four-year college (Figure 1). The analysis did not include those players who were not drafted, but who signed as amateur free agents and then subsequently made their way into MLB.

After the selection process is complete, MLB invests in the human capital of those drafted by providing additional on-the-job training in the form of the minor league baseball experience. Research is needed to determine how beneficial this investment is. Furthermore, conducting research on player development can assist in a better understanding as to what the benefits are to attending college versus not attending college and entering professional baseball directly after high school. Does the combination of on-the-job training (OJT) and higher education tell us anything about the human capital involved in a successful career in MLB? Finally, does the offensive performance of MLB baseball players during any period of the first five seasons in the major league inform us about the type of training or combination of human capital (education and/or OJT) received prior to being called up? These concerns highlight a need for research on the aspects of human capital and its relationship to the player development of MLB baseball players.

Purpose of the Study

The purpose of this study was first to examine the relationship between a baseball player's human capital and player development, and second to investigate the effect it

had on certain offensive performance outcomes once he was called up to the major leagues. By concentrating on the first five seasons of a player's career we examined how his combination of education and on-the-job training influenced him during the early portion of his career. The first five seasons are crucial because it is during this time that a player is under the reserve clause; it is when a player is trying to establish himself and has not been influenced financially by multi-million dollar contracts, and it is usually before a player is deemed a superstar.

The research built upon an earlier study conducted by Shughart and Goff (1992) in which they specifically addressed the following problem:

> Other things being equal, do professional baseball players with college experience require smaller investments in training and development by Major League Baseball than their counterparts who turn pro after high school? More specifically, do big league ball players who enter the pro ranks from college spend less time in the minors, on average, than other athletes who start their pro careers with only a high school background? In short we seek to measure the value of the college game to professional baseball by computing the marginal reduction in farm system playing time associated with college baseball experience. (p.93)

In an effort to expand on the Shughart and Goff study and present an additional contribution to the area of research that focuses on player development, it is imperative that additional research be conducted. The research questions outlined in the dissertation moved the Shughart and Goff study forward by asking questions about these same players during the next phases of their careers – the major leagues, by specifically including such elements as examining the effects of productivity between amateur

baseball players with and without college experience once in the major leagues. It is important to note that this study was not intended to determine the marginal revenue product of any player; the aim of the study was to look backwards and evaluate the training received by each player regardless of their decision to go to college or not. Also, the study did not consider all players called up to the major league ranks, international players and those not officially drafted during the MLB draft. By analyzing what previous research has identified as the most informative offense performance statistics during the first five years of their MLB career and making a prediction as to who provided their MLB team with better output, a player who attended college could be compared to a player that did not.

Research Questions

RQ1: Do MLB clubs get better production from amateur baseball players taking the path from high school to the Minor Leagues to the major leagues or from those with some collegiate playing experience?

RQ2: Is there any difference in the mean time to reach the major leagues for a baseball player with only a high school education and a player who attended college?

RQ3: What effect does on-the-job training (OJT) have on the following offensive performance statistics of MLB players who did not attend college compared to those who did: batting average (AVG), runs batted in (RBI), on base percentage (OBP), slugging (SLG), and on base+slugging (OPS)?

RQ4: Do Major league baseball players who attended/played at a four-year college perform better during their first four years in major league baseball than do their teammates who do not have any collegiate playing experience?

Significance of the Study

The significance of this dissertation was its contribution to the advancement of sport management research in higher education and the contribution of human capital to player performance. By exploring the relationship between data and theory, in this case data from MLB and Human Capital Theory, this dissertation tested each research question formulated above. These research questions increase our understanding about how sports data reflect the human capital investment made by professional baseball players. In addition, this research assisted in providing an understanding as to how education and on-the-job training affect the most transitional and critical years of an MLB player's career the year after being called up and the year prior to free agency. While contributing to the advancement of sport research, there are three groups specifically that have been identified as potential beneficiaries of the study: MLB teams, amateur baseball players, and colleges.

From the perspective of the MLB teams, the study was significant for numerous reasons, including providing the buyer, MLB teams, with valid information to assist in the decision to draft or not to draft players depending on the level of education achieved by the input, the player. As the buyer of talent, MLB teams are also able to utilize the findings to learn how offensive performance statistics and human capital variables can be used to make a prediction about the player development process. Finally, the findings could provide MLB teams with more information about their investment in a player's human capital, namely, the on-the-job training provided by the minor league system it currently utilizes.

In addition to being significant to MLB teams, the findings also provided amateur baseball players with details about which path to embark upon and whether or not attending college makes sense. The player will also be able to understand from which education level the MLB teams are acquiring the majority of its players, as well as identifying which players are the most successful early in their careers, which will then assist in their decision making process whether to attend college or enter professional baseball immediately following high school.

Chapter Two: Literature Review

The purpose of this chapter is to first provide an overview of professional baseball and to introduce some relevant details of the draft process and the minor league system within the structure of professional baseball. The second is to review Human Capital Theory (HCT) and establish a thoughtful argument as to how HCT is applicable for conducting research on the relationship of the human capital of MLB baseball players and their development as players. Specifically, examining how previous research in the disciplines of labor economics and the economics of education have been guided by HCT will direct us towards identifying the proper variables that will provide a spotlight on the characteristics that a player possesses and which are then labeled as human capital variables. Next, a review of the literature in the area of economics that focuses on sport will assist in identifying an existing gap that illustrates why it is relevant to perform such an analysis of player development.

Background of the Draft Process and Minor League System within Professional Baseball

Major League Baseball established the reverse order amateur draft in 1965. Participants entering the draft are United States or Canadian citizens that usually fall in one of three categories: high school graduates, athletes that have completed their junior or senior season in college, or those from junior college baseball programs. Since 1998, the only draft that MLB conducts occurs in June and is known as The First Year Player Draft. Several studies have been conducted examining the impact of collegiate playing experience and its effect on the draft (Shughart & Goff, 1992; Spurr, 2001; Winfree & Molitor, 2007).

The average minor league baseball career length for players who enter directly after high school into professional baseball is 4.38 years, and for players who do not directly enter into professional baseball and attend college is 1.70 years (Winfree & Molitor, 2007). College baseball players either receive a full or partial scholarship, or no scholarship. The minor league pay structure is at the highest level before the majors. A Triple-A first year player earns a minimum salary of $2,150 a month, and after the first year is no less than $2,150 a month. At the next level down, Class AA; first year player earns $1,500 a month, and after the first year no less than $1,500 a month. Finally, Class A (full season) first year earns $1,050 a month, and after the first year no less than $1,050 a month, while Class A (short-season) first year earns $850 a month, and after first year no less than $850 a month (MLBPA, CBA, 2006-2011).

Human Capital Theory

Human Capital Theory can help to understand how the professional sports labor market rewards the performance attributes of professional athletes (Antonietti, 2006). The basic Theory of Human Capital suggests that individuals and society tend to benefit economically from investments in people. Investments can be evaluated in a number of ways, including health, nutrition, education, or on-the-job training. Three of the seminal contributors to what is Human Capital Theory today are Gary Becker (1964), Theodore Shultz (1963) and Jacob Mincer (1958). Human Capital Theory tends to be studied across many disciplines but traditionally has developed roots within the disciplines of economics and education. As an economic concept, Human Capital Theory has been around for several centuries. There are a number of studies throughout both education and economics that examine the relationship between the level of education attained and

lifetime wages and between on-the-job training and wages. As these studies have focused on wages, there are three primary themes that consistently reoccur in the literature. The first and most utilized is formal education; second is on-the-job training; and last is investment in health (e.g. exercise, health care, personal beautification expenditures, etc.). Workers with higher levels of education and more work experience tend to have higher wages and increased productivity. For the purpose of this research, the investments in formal education and on-the-job training were the primary focus. Human Capital Theory, as outlined by Becker and Mincer, is typically conceived as being carried out by individuals acting in their own best interest; therefore, for the purposes of this study, the focus was on the primary investment of human capital by the individual through formal education and the secondary investment of on-the-job-training because workers with a higher level of education and more work experience tend to have higher wages because they are more productive (Blass, 1992). The argument is not that one is better than the other, but that each provides a contribution to the study. Therefore, applying the Human Capital Model, an MLB team would want to make an investment in a player who has a higher level of education rather than a player that does not.

Education

Human Capital Theory deals with acquired abilities and capacities that are developed through formal and informal education at school and at home (Antonietti, 2006). Formalized education can be defined as education received at the primary, secondary, and higher levels, and informal education is defined as at home, work, or in specialized vocational training at secondary and higher levels (Sweetland, 1996).

Becker, one of the foremost authorities on Human Capital Theory, conducted many studies on the Theory. One of his preliminary studies compared the personal incomes of college graduates with those of high school graduates and related the difference in income to the cost of attending college (Becker, 1962). Education is the prime investment in human capital. It is therefore appropriate to assume that education has the ability to increase or improve the economic capabilities of people, including athletes (Shultz, 1971; Sweetland, 1996). Sport can be considered as a type of Human Capital investment and Human Capital Theory can be applied to understand how the professional sports labor market rewards performance attributes of players. Human Capital Theory states that workers with higher levels of education and more work experience, such as participating in college baseball programs, tend to have higher wages and higher productivity (Antonietti, 2006).

Applying the Human Capital Model, DuMond et al. (2008) specifically examined the choice process of high school football recruits and the benefits of attending college. Attending college is assumed to be an investment in human capital that would increase the productivity of the recruit or draft choice in the labor market. This investment in human capital may differ from one school to another. In the case of improvements in human capital, there is a difference by school but not because of academic rank of the university but more likely because of the number of former NFL athletes the program has produced (DuMond, Lynch, & Platania, 2008).

The educational component of the Human Capital Model allows for many unobservable characteristics to be correlated with schooling and to be taken into account when considering the addition of a player or worker. Becker (1962) made it clear that a

distinction between schools and firms is not always necessary and in actuality schools can be treated as a special kind of firm and students as a special kind of trainee. It is possible that college students do not only work for necessity of income but rather to gain more marketable skills that are not readily available in the classroom setting (Light, 2001). Individuals may be more apt to take on the role of college student and employee because they lack a desire to be categorized into one or the other. In the case of the NBA for example, the Human Capital Model informs scouting departments that teams can expect fewer minutes will be played per game during the earlier seasons and rise with experience in the league if teams invest in a player's human capital during his earlier years. Players with at least two years experience in college play more minutes than their colleagues with less than one year of college experience (Groothius, Hill, & Perri, 2007). According to Groothuis and colleagues (2007), the results support the argument that players gain human capital with increased experience in the league. Furthermore, NBA teams are more likely to take a risk by drafting those players that are believed to be future all-stars.

Athletics aside, the research has shown that if a year of schooling is worth completing when a worker is older, it is certainly worth completing when a worker is younger (Weiss, 1995). Investing in a formal education when someone is young allows returns to be enjoyed over a longer period of time.

To illustrate the importance of obtaining an education, several examples from a variety of professional sports are available. In the NFL for example, training consists of playing college football in order to reach the NFL draft. Young amateur football players make the decision to position themselves to overcome the barrier of entry into the NFL

by participating in collegiate football. In addition to developing athletic skills, they are investing in human capital by attending college and participating in on-the-job training.

Another way in which Human Capital Theory has been applied is to analyze the human capital transfer they receive from their individual situational circumstances. Laband and Lentz (1990) argued that the sons of major league baseball players are the recipients of valuable human capital transfer from their fathers; these transfers exist in the form of tips and access to scouts, coaches, games, practices, and locker rooms at a young age. Having access to major league baseball experiences such as working on batting stance, receiving fielding tips, and strategies about how to pace through a 162-game season at such a young age provide the sons of MLB players with a wealth of information and knowledge that most other young players would not receive until later in their teenage years (if at all). A son following his father into baseball has intimate knowledge of the game. The quality of information and instruction is far superior to the average little league baseball player or high school coach. According to Laband and Lentz (1990), this career specific human capital is captured at a relatively low cost to the son of the MLB player. The benefit, however, is higher than if the sons of these players were to choose any other profession because of the years necessary for formal education or training.

Labor Training

On-the-job training is primary to the Theory of Human Capital because, according to Becker (1962) and confirmed by Mincer (1974), it clearly illustrates the effect of human capital on earnings, employment, and numerous other economic variables. Defining on-the-job training as "the process that raises future productivity and differs

from school training in that an investment is made on-the-job rather than in an institution that specializes in training," Becker (1962, p.11) made a strong argument for general training and the portion of the cost employers are willing to pay in order to provide it. In contrast, he made the prediction that workers themselves are willing to pay for general training by accepting a reduced wage during the training period, an example provided by the salary structure of minor league baseball.

Specifically discussing MLB, training occurs during the player's time in the minor league system. Krautman et al. (2000) noted skills taught to players in the minor leagues are transferable to other teams, therefore, the initial team that provided that training must be able to recoup its investment in general training. This recovery occurs while a player's team to team movement is restricted under the reserve clause by paying them [players] less than their marginal revenue product. As published in the Handbook of Labor Economics (1986), marginal revenue product (MRP) can be defined as the change in revenue that results from the addition of one extra unit when all other factors are kept equal. The marginal revenue product is used in marginal analysis to examine the effect of variable inputs, such as labor, and typically follows the law of diminishing marginal returns. As the number of units of a variable input increase, the revenue generated by each additional unit decreases at a certain point. It is calculated by taking the marginal product of labor and multiplying it by the marginal revenue of a firm. In Human Capital Models, wages rise with the length of time spent on a job regardless of productivity because that is the environment in which workers tend to learn. Learning tends to occur via two streams either on-the-job training or learning by doing (Blass, 1992; Weiss, 1995).

Training can be categorized as general or firm specific. General human capital includes skills that are learned and have value to many, thereby affecting returns to experience, regardless of whether that experience happens to occur at one job or many jobs. The sacrifice on behalf of the employee is receiving wages below his current productivity. However, the benefit would be a future increase in wages because of participating in the general training. In contrast, firm-specific training takes place at one employer. Workers with a specific type of training unique to a certain industry or occupation are more likely to remain on the job because their skills would have less value in the general labor force.

For example, Singell (2001) argued that the amount of playing experience a Major League baseball manager has establishes a direct correlation to his team's ability to win games and to the individual performance of his players. In this two-step process, the researcher first examined what portion of the managers direct contribution can be attributed to human capital and less to what is typically studied – wages. Second, the researcher evaluated the direct and immediate influence a manager has on a player's performance. In order to evaluate the contribution a manager has on a player, the author looked at the player's performance pre- and post-trade and looked for improvement to outputs.

Another relevant study of Human Capital Theory was conducted by Horowitz and Sherman (1980). The authors measured the relationship between experience and training in the productivity of workers, but in the industrial sector of the economy and not specifically in sports. As an alternative to examining wages as a direct measure of productivity, the authors substituted output at the work group level as a direct measure of

productivity. The findings of this study suggested that it is possible to use a different measure of productivity other than current wages. In sectors of the economy where physical output is unobserved, the productivity of maintenance workers was measured by the condition of the equipment it was their duty to care for.

It is frequently argued that individuals continue to invest in themselves after college by choosing to enter into professions that require some type of general training. In accepting this path, wages are once again sacrificed in exchange for the higher salary that can be commanded once the training has been completed. This argument mirrors what amateur baseball players embark on down their dichotomous training path.

Previous Research

This section provides a review of previous and current research areas relevant to investing in player development. The section concentrates on four areas of research: education and in-school employment, training cost/investment in player development, the draft, and performance. First, many studies have focused on the advantages and disadvantages of working while in school. Even if athletes are not being paid to play, the argument is still made that with the amount of hours spent on their sport, it is essentially work (Light, 2001; Long & Caudill, 1989; Parsons, 1974). Second, the investment of player development sets the tone for this research study by indentifying how previous studies have examined training cost and player development. Subsequently, the draft is an important de facto measure of talent at a point in time and can also be used as a control factor. For example Staw and Hoang (1995) looked at the NBA draft and the order in which players were selected to predict playing time, trades, and career length in the NBA. Finally performance establishing what the players are worth at different stages

in their career makes sense, but having the ability to determine productivity prior to introducing the large quantities of money once a player reaches free agency lends to the creditability of this study. The review of these four areas encapsulate the topic of player development providing support in identifying the existing gap in the literature for this study and its relevance.

Employment While in School

A large body of literature has addressed the issue of how athletics and work experience while in college affect academic performance, completion rates, post college labor market success, and the decision to enroll in graduate school (DeBrock, Hendricks, & Koenker, 1996; Ehrenberg, 1984; Ehrenberg & Sherman 1986; Gee, 1984; Haley, 1973; Light, 2001; Long & Caudill, 1991; Rumberger, 1979).

Ehrenberg and Sherman (1986) made the following two assumptions: first, the less time spent working the more time the student will have to study and the better he will do in college; and second, gaining work experience while enrolled in college helps to develop work habits and attributes such as high motivation that make potential employees more attractive to employers. Consequently, once someone has successfully completed the requirements for a college degree, their wages should be positively correlated to hours worked while enrolled in college. The authors' findings suggested that the number of hours worked, which is typically less than 25 hours per week, had no adverse affect on the grade point average of students attending either a four-year or two-year college. However, the authors did find that the hours worked had a positive effect on the dropout rate and an adverse affect on a student's ability to graduate on time. The authors provided a plausible explanation for these effects: first, students have a reduced

perception as to the value of a college degree; and second, there is an inability to enroll in the required number of courses per year due to the amount of hours worked. The Ehrenberg and Sherman study did ignore the following: the effects of loan burdens, occupational choice, the choice of college major, and the decision to enroll in postgraduate education. Consistent with the finding of Astin (1975), working off-campus was found to have a positive relationship with increased levels of dropping out and failure to graduate on time. Additionally, working on campus was associated with higher graduate school enrollment. In contrast, Long and Caudill (2001) examined the effects of participation in college athletics at the varsity level. By examining males and females that attended college during the early 1970's, not only were the graduation rates of athletes during this time period higher than those of non-athletes, the findings also suggested that athletic participation may enhance the development of discipline, confidence, motivation, a competitive spirit and other subjective traits that encourage success. In addition, when measuring participation in athletics against post-college wages, the authors found that males received an estimated 4% higher annual income than their non-athlete counterparts. These studies are consistent with the argument that students participating in college athletics spend just as many hours on their chosen sport as those students that chose to be employed while in school and garner many, if not more, of the same benefits.

Other researchers have utilized an earnings or wage function to estimate the implications of student employment and the relationship between schooling and post school wages (Ahituv, Tienda, Xu, & Hotz, 1994; Coleman, 1984; Light, 1998, 2001; Michael & Tuma, 1984). Light (2001) suggested that students might choose to work in

order to acquire marketable skills that are different than skills gained in the classroom alone. Participating in baseball while in college provides not only the opportunity for a MLB career but also some additional of skills not obtainable while in the classroom: time management skills, leadership skills, discipline, as well as myriad others. Additionally, by examining the causal effects of post school wages for those in school and time spent working at the same time as being enrolled in school, and also controlling for work experience, the researchers were able to identify the value of those skills acquired by taking part in things outside the classroom.

Employment while in school was found to have a substantial impact on post school earnings. Light's main finding was that estimated schooling coefficients were 25-44% higher when in-school work experience was ignored from the earnings function. Simply stated, when others choose to omit in-school work experience there is a significantly overstated wage effect for schooling. These findings are consistent with Hotz and colleagues, where a positive relationship between post school wages and in-school employment was eliminated when unobserved factors were taken into account (Hotz, Xu, Tienda, & Ahituv, 2002; Light, 2001). Hotz et al.'s (2002) findings indicated that when examined alone, schooling tends to have a greater influence on wages than when school and employment while in school are combined. An additional finding of Hotz et al. (2002) was that wages earned after college and employment while in school tended to both be determined by unobserved observations which included ability and family background. These findings are consistent with findings from a previous study by Rumberger (1979), that college graduates had an economic advantage over high school graduates in the labor market.

Building on Light (2001), Molitor and Leigh (2005) examined the effect of employment experience gained while in school on the returns to education when looking at each level and type of educational institution attended. The findings were consistent with other research findings in that when there was a failure to take into account experience gained from working while in school the results in returns to schooling tended to be overestimated. Furthermore, by examining the different levels of education, such as attending a two-year college where the receiving of a degree may not be the primary reason for attending, this overestimation was more likely to occur.

Training Cost/Investment in Player Development

Player development and training costs have not been widely studied and the following section reviews the studies (Shughart & Goff, 1992; Winfree & Molitor, 2007). Much of the literature is found in economics where baseball salaries have been estimated and revenue sharing and competitive balance have been heavily researched (Fort & Quirk, 1996; Krautmann, Gustafson, & Hadley, 2000; Scully, 1974, 1989, 1993). This section goes on further to review and make note of where the gap in the literature exists in providing the rational for this research study.

The following studies set the foundation to extensively examine the topic of player development. It seems to be that research has started to become inquisitive as to how players are trained but then moves immediately on to examining the salary structures of professional sports teams and players. Researchers have had remarkable findings as to estimating the time frame in which different players are exiting the minor league system and the choices they make, but have failed to examine and compare the productivity of

college and non-college players to answer the question: "Is there any benefit to spending less time in the minors?"

Winfree and Molitor (2007) used salary and return to education estimates to analyze the decision that drafted high school baseball players must make – whether to attend college or sign a MLB contract – utilizing a sample of 7,800 high school baseball players drafted between 1965 and 1980. The results of the study suggested that if a player was drafted during one of the earliest rounds of the draft and decided to attend college, he had a higher probability of being redrafted and receiving a signing bonus; however, the findings offer no recommendation as to the amount of the bonus they could expect. The most significant finding related to a player's lifetime earning potential, findings showed that high school players drafted in the earlier rounds could maximize their earnings if they signed a professional contract immediately, but if they were drafted after the 11th round, the potential to maximize their lifetime earnings were achieved by attending college. Furthermore, a player drafted in the earliest rounds was better off entering professional baseball regardless of whether the player was from high school or college. In contrast, the team that drafted him had no effect on a player and his decision to attend college.

While the previous research study addressed the financial value of attending college, Shughart and Goff (1992) focused on the length of time spent in the minor leagues after the decision was made, the impact of college baseball playing experience on player development time in professional baseball's minor leagues. The findings were consistent, however, the approach was different. This study analyzed data on all major league baseball players from one season (715 observations), the 1989 regular season,

whereas Winfree and Molitor (2007) chose to analyze multiple seasons. Additionally, the Winfree and Molitor (2007) study only included conclusions as to players choosing to attend four year colleges and excluded players that chose to get playing experience from junior colleges. This study also utilized a sample of 7,800 high school baseball players drafted in the June regular draft between 1965 and 1980 choosing to examine a player's entire career. Shughart and Goff (1992) found that amateur players who chose to attend and play baseball at a four year college advanced to the big leagues faster than their counterparts having attended only high school or junior college. Moreover, they supported the argument that players choosing to extend their playing careers by attending college benefit major league teams by providing additional screening information as well as conserving on the expenses associated with player development. They determined that college baseball provides an important input into the production of major league baseball players, and leaves the option open to further analysis of the output of major league baseball players based on level of playing experience achieved. As stated by Rosen and Sanderson (2001),

> In the United States, the supply of athletic talent for professional individual and team sports comes from two main sources: minor league 'farm' systems or tours and colleges and universities. From there outset, professional sport leagues have instituted a variety of arrangements or restrictions among owners that affect the employment of players and their distribution across teams and impact the supply side of labor markets. (p. 19)

Additionally, other economists have performed analyses on the reserve clause, salaries, and the effects on training costs (Daly, 1992; Krautman, Gustafson & Hadley, 2000; Miceli & Scollo, 1999; Rottenberg, 1956). Examining the effects of the reserve clause on training costs began with Simon Rottenberg. He offered the following

argument regarding the reserve clause and invariance proposition. Without the reserve clause teams would be unwilling to invest in player development and training because devoid of the ability to suppress wages, how a team would recover their investment from those players eventually making it to the majors and for those that never make it to the majors is unknown. Rottenberg (1956) later reversed his argument by stating that even without the reserve clause owners would simply discount wages (owners no longer set the salary) so in effect the players would then pay for their own training cost. This would be an efficient method for the players as they would find it beneficial because their lifetime earnings potential would increase. Minceli and Scollo (1999) offered the following in support of the reserve clause and its relationship to training costs and player development: if players are interested in achieving the greatest amount of increase in salary and expected wealth, they will leave a limited form of player reservation in place as this is in their best interest. The argument is that the reserve clause allows for players to honor a commitment to the owners that provided the training for a brief time, allowing the owners to not be hesitant in investing in the training and thereby alleviating the players from covering their own training cost.

Krautman et al. (2000) addressed training costs and player development from a more traditional approach, that of wages. The researchers examined monopsonistic exploitation, the underpayment of players, as a justified attempt by owners to recuperate the expenses of general training. By outlining a method for estimating the surplus extracted from those players restricted by the reserve clause they then determined how the surplus was divided among different players. There were two different types of players identified as existing under the reserve clause: players with less than 3 seasons of

major league service who were bound to their teams with no recourse when it came to their salary; and players with between 3 and 6 years of major league service also bound to their team but who were eligible to submit salary disputes through final offer arbitration. The results suggested that owners extracted the largest amount from the most productive and indentured players -- those with less than 3 years of service and under the reserve clause. They estimated this amount to be $3 million dollars per year. Additionally, the results indicated that for those players still under the reserve clause but eligible for final offer arbitration (FOA) they were being overpaid. They offered an explanation that owners would rather overpay than go to FOA as an attempt to offer a settlement in order to shorten the time that a player spends arbitration eligible, between seasons three and six.

The Draft

The MLB draft is the primary source of identifying talent for selection into major league baseball. Baseball America tracks each draft class with latest numbers since 1965. During that period, 64.9% of Round 1 selections eventually played at the major league level and 41.6% of second-rounders. Including sandwich picks between the opening two rounds, the combined figure was 53.5%. The percentage dropped nearly 10% in each of the following rounds to only 12% for Round 5 choices. Finally, rounds 6-10 combined was only 15.5%. First-round picks usually guarantee instant millionaire status. Bonuses ranged from $845,000 to $3.96 million in 1999, $400,000-$5.3 million in 2000 and $900,000-$5.15 million in 2007 (baseball-reference.com).

College baseball has also been proven to provide additional screening or signaling to major league teams about amateur players up to an additional four years. These

additional years of observation are beneficial to both the athlete as well as to the major league club owner, the former provides the clubs more information about his skill and the latter is twofold: it saves the major league club the expense of providing the training for four years and allows a risk to be avoided if the player in fact failed to live up to the drafting team's expectations.

Predicting the best draft choice between high school players and college players has also been a topic of research. Spurr (2000) conducted an analysis focusing on the demand side of the baseball draft and the ability to find talent. The objective of the study was to determine if a player's ability to reach the major leagues varied by the position he played or the level of schooling he attained. The findings suggested that the drafted player's potential for success (ability to reach major leagues) was substantially greater if he attended college.

The analysis also included an evaluation to determine if one team had a more superior draft selection process over another and if that process was able to predict future major league talent. The author's findings suggested that no one team has any advantage over any other within MLB. However, he did find that there is a greater propensity for a player with college playing experience to make it to the majors more frequently than a player with only a high school background, which is consistent with Shughart and Goff (1992). Once again, nothing was taken into consideration in regard to how the players performed once they reach the majors, thus leaving the option open for further analysis.

Performance

In 1975, a contractual change in major league baseball occurred: veteran players not under contract were given the rights to sell their services to other teams within MLB

(Scully, 1989). After six years of service to their major league club, major league baseball players become eligible for free agency. A number of studies have been conducted evaluating the correlation between pay and performance (Fort, 1992; MacDonald & Reynolds, 1994; Maxcy, 1997; Scully, 1974; Zimbalist, 1992).

Scully's (1974) initial study on MLB involved a two step approach: 1) how different measures of a player's offensive performance such as runs scored (RUNS), runs batted in (RBI), and strike-to-walk ratio and the market size in which a team was located affected a team's winning percentage; and 2) what effect the team's revenue had on the team's winning percentage. The findings suggested that players labeled as average received roughly 20% of their marginal revenue product and those players labeled as stars about 15% during the days of the reserve clause.

In additional studies, Scully (1989) and Zimbalist (1992) revisited and updated Scully's (1974) calculations. By examining players labeled as stars during the 1987 season, the new results nearly tripled the original study's calculations in that star players were paid between 29% and 45% of their MRP and no longer held to the restrictions of the reserve clause but free agency.

According to Szymanski and Zimbalist (2005), the new system helped to increase player salaries as well as increase the gap between similarly skilled players. Zimbalist (1992) compared two groups of baseball players, those eligible for free agency and those not eligible for free agency in 1989. During this season, players with less than three years of service were not eligible for salary arbitration or free agency. He found thst the ratio of salary actually paid to the marginal revenue product was .38 times what it was for those eligible only for salary arbitration and .18 times for those eligible for free agency. The

authors noted that these results only provided for the effects on revenue by evaluating a player's performance statistics effects on winning.

However, the data show a very close correlation between estimated MRP and wages for experienced players that have achieved free agency and have the ability to negotiate with multiple teams. Contrary to popular research, the majority of players do not have a record breaking season the year before they are eligible to become a free agent. Research has shown that those MLB players in the year before their free agency year typically did not show improvements prior to becoming a free agent. On the other hand, in the same study there was evidence that during the first year of the new contract, performance statistics were actually significantly lower than in the year prior to the negotiating year, suggesting that if MLB teams pay free agents solely based on the performance of the free agent year, they might not be satisfied with the results (Ahlstrom, Si, & Kennelly, 1999).

MacDonald and Reynolds (1994) examined how the changes to free agency and final offer salary arbitration have drawn the salaries of MLB players more in line with their current marginal revenue product. The marginal revenue product (MRP) of a particular player is the extra price that a spectator is willing to pay multiplied by the number of people that are attracted, in person or via television (Rosen & Sanderson, 2001). Using publically listed salary data on players on major league rosters from August 1986 to August 1987, the findings offered substantial evidence that players with more major league playing experience are paid in accordance with their estimated career marginal revenue products. However, younger players, those labeled as rookies, tended to be paid less than their actual marginal revenue product. This runs parallel to whether

or not they have met the necessary requirements qualifying them for final offer arbitration or free agency.

Continuing the investigation of pay for performance economist examined why owners were willing to sign players to multiyear multi-million dollar contracts after one season of spectacular performance. Krautman (1990) compared each player's performance in the period right before and right after signing a long term contract and found no evidence that the proximity to contract negotiations have any adverse effect on performance. Maxcy (1997) examined players over eight seasons that were currently in long term contracts to determine the effects of long-term contracts had any effect on player effort or performance. Results are consistent with those in the previous study, which indicate the players with long term contracts that are in the season before contract renegotiations does not influence performance. Further, the results showed minimal evidence that there was any increase in performance when a player was not in the last year of his current contract.

Harder (1989) also examined performance during the walk year by measuring performance using two statistics that are more unconventional than most, runs created and total average claiming that these two statistics best measured a player's contribution to his team's success. A walk year is defined as the season prior to becoming a free agent. Using data from four seasons, 1976-77 and 1987-88, he concluded that a player participating in his contract year had no effect on total average runs created during the 76-77 seasons and a negative effect in 1989. For total average he found that there was no statistical significance indicating that total average increased during either of the test periods. Therefore, he concluded that a player had no reason to believe that an increase

in statistical performance would have any bearing on a higher salary in future years. However, Blass (1992) tested the theory that workers were initially underpaid relative to productivity and overpaid later in an analysis of MLB, and found that the wages of MLB players increased with experience independent of productivity, thereby suggesting that MLB players needed to stay consistent to enjoy the wage gains down the line.

After a thorough review of the literature and the theoretical premises guiding this research, it is apparent how Human Capital Theory is applicable to the current research project. Human Capital Theory has directed us toward certain variables that can be used to examine the relationship between the theory and the data. The general conclusion that is drawn is that Human Capital Theory points to education and on-the-job training as key components of a workers background because worker's with higher levels of education and experience on-the-job tend to be more productive and contribute more to society as a whole than those that do not have such a background. This study set out to determine the various outcomes that education and training have on the productivity of major league baseball players during the first five years after being called up to the major leagues. Human Capital Theory therefore informed this study by acknowledging and affirming the importance of education and on-the-job training for people as individuals. What can be done to take these same principles and apply them to MLB players and find the measureable impact of these two variables on the early stages of their careers. The review of previous research has identified an existing gap to support the argument of moving forward with the current research project. Salary and performance have both been extensively researched; however, the development of players has been virtually ignored. Though previous studies have compared the length of time spent in the minor

league system by college players and high school players, no study that has examined what happens once they are in the major leagues. Even if college players spend less time in the minor league system, it is possible to go further to compare the productivity of these players versus high school players to make a determination as to player development. If it is found that high school players are more productive than college players, then maybe college players need to spend more time in the minor league system working on the development of their game.

Chapter Three: Method

The goal of this study was to analyze how a player's human capital influenced his major league baseball playing career, specifically his offensive contribution to his MLB team, by examining player performance over the first five years after being called up to the major leagues. What type of education a player received (high school, junior college, or four year college) prior to being called up to the major leagues was informative about this training. It is possible that the source or the path of his training may affect its quality. This was determined by analyzing performance in the major leagues as a measure of the training quality, controlling for other factors. The method employed was secondary data analysis utilizing two primary sources: mlb.com and baseball-reference.com. The main objective was to document and analyze a player's level of education, draft year, age, location, information regarding time spent in the minor leagues and performance (via hitting statistics of all MLB players from the 1990 season through the 2000 season). All information on the players and their performance was based on regular season games.

Data

The sample consisted of amateur baseball players selected in the MLB June Regular Draft over eleven seasons covering the years 1990 through 2000. The data used in this analysis came from two primary sources: the official website of Major League Baseball (http://www.MLB.com) and Baseball-Reference.com (www.baseball-reference.com). Secondary data already exist and require only retrieval and the selection of the analytical technique. Secondary analysis is extremely versatile in that it can be applied to studies designed to understand the present or the past, to understand change,

make comparison, or to replicate and extend previous studies (Hyman, 1972; Kiecolt, 1985; Stewart & Kamins, 1993).

Secondary Data

Many researchers choose to analyze secondary data over collecting primary data, for a number of reasons, including but not limited to the fact that the data already exists and in this case is constantly updated. In addition, by using secondary data one can access large amounts of information that might be difficult to obtain through primary data gathering methods such as surveys and focus groups. With the ever evolving reach of the World Wide Web and technology, we are afforded with real time up to date access to reliable current information.

Keicolt and Nathan (1984) noted how independent data collection by an individual investigator has become increasingly more difficult. As educators and researchers in a climate of declining resources, the state of the economy and educational funding has declined in recent years, using secondary data requires less money, less people, less time. Unfortunately, collecting data first-hand requires time, planning, in some cases travel, and access. Secondary data avoids primary data collection problems when relying on human subjects. By using secondary data it is not necessary to be concerned about whether or not the test subject will cooperate, obtaining sufficient participants, or the amount of noise generated by the experiment or data collection procedure. Secondary data allows for a better represented sample, access, and can be more readily subject to a variety of research designs and software programs. Research projects conducted using secondary data for analysis are more readily able to be

duplicated and expanded on. When conducting studies requiring data from the past it could be nearly impossible to obtain the original data in some circumstances.

In contrast to the advantages of secondary data, there are several disadvantages as well. The availability of the necessary data can sometimes be difficult to locate or a fee might be required to access it. Depending on what the area of research is, the study might be based on some general topics that are more readily available than some more specific areas the researcher may want to analyze. Time can be both an advantage as well as a disadvantage, time is required to collect the data from various archives and this can be a very arduous process. In addition, it is not uncommon for incomplete and inconsistent information to exist as it may not be on file in its entirety or there is the possibility that it might be inaccurate. Inaccuracies can occur because there might be errors that occurred in the initial recording of the data or the data quality can create a poorly designed study, a poorly coded study, a lack of competent procedures, and missing variables in addition, the exact variable a researcher is interested in investigating may not have been recorded in the original data. These are all reasons why primary research is conducted and preferred. It must be noted that the advantages far outweigh the disadvantages in terms of the necessary information and lack of access needed to undertake this particular project.

Why Use Secondary Data for This Study

For this particular study the use of secondary data was a suitable choice due to the lack of access to MLB and because of the historical nature of the MLB data needed to accurately address the research questions. Therefore utilizing the "Official Site of Major League Baseball" MLB.com and trusting the statements put forth on the home page of

the website, "MLB.com is committed to you, our fans, to the source of baseball. We want to bring you the most up-to-date, informative and exciting Baseball information on the web", a statement that puts the reliability of the data in perspective. In addition to access, the complicated process, the number of games and the number of players the data set encompasses would have required years and many people to collect. With the technological advances of the internet, secondary data was an attractive alternative for this study. MLB.com and Baseball-Reference.com made it possible to conduct a limited number of searches for the data needed. Both sites were readily accessible and fairly simply to navigate. Finding complete information for all draft years used in the study, all statistics and up to date biographical information did not present an insurmountable challenge.

Research Design

Data Description. It is important to note that the data set consisted of regular season statistics only and no post-season information. In order to be included in the study, a player must have met all of the following conditions: first, he could not play the position of pitcher; second, he had to be drafted instead of being signed as an amateur free agent; and third, he must have been called up to the major leagues.

Pitchers were omitted from the study because of their lack of contribution to scoring. Only in the National League do pitchers hit and in the American League pitchers do not hit, except on very rare occasions. The primary way teams win games is by scoring runs and runs are produced by the offensive output of players. Offensive players are utilized because there are more of them in comparison to pitchers and more detailed performance data. In addition, players signed as amateur free agents (meaning

that they were not drafted) also were not utilized in this study. Information about a player's level of education and his age at the time of the draft and during each season played was a critical part of this research. Unfortunately, this information was not readily available for players signed as amateur free agents.

While collecting the data a pattern emerged: those players listed as having entered the league as amateur free agents primarily resided outside the United States, its territories, or Canada and therefore were not eligible for the draft as outlined by the rules of MLB. According to Chiba (2004), from 1995-1999 approximately 90% of foreign-born players in MLB were of Latin American origin. It is widely accepted that in addition to operating and supporting minor league teams, many teams also operate baseball academies in foreign countries such as the Dominican Republic, Venezuela and Japan at a cost upwards of 60 million dollars. There is much more untapped, cheaper and much easier to sign talent in these countries (Chiba, 2004). In 2003, major league baseball had representation from 16 countries, consisting of more than 200 players born outside the United States on 25-man rosters and the disabled list at the beginning of the season (SportsIllustrated.com, 2003). Amateur free agents and baseball academies from around the world were outside the scope of this dissertation. Finally, the criteria to play minor league baseball are not nearly as difficult as playing in the major leagues. This study made a comparative analysis of player development by utilizing a combination of variables, and one of those variables included total seasons spent playing in the minors.

The data from mlb.com yielded 5,408 observations. The variables obtained from mlb.com included: position played, games, at bats, runs, hits, doubles, triples, homeruns, runs batted in, total bases, base on balls, strike-outs, stolen bases, caught stealing, on base

percentage, slugging, batting average, sacrifice flies, hit by pitch, intentional walks, total plate appearances, extra base hits, stolen base percentage, and on base percentage plus slugging (terms are defined in Appendix A) for each player each season. The next step in this process was to augment the data from mlb.com with information from the second data source, baseball-reference.com which is an online clearinghouse of up-to-date sports statistics and history that goes beyond baseball. The information gathered from baseball-reference.com included: birth year, draft year, draft round, education, first and last year in the major leagues, players still active and total career length. Baseball-reference.com provided substantial information; however, areas of missing data were also brought to the forefront. Missing data was identified for, last year in the major leagues, which occurred because a player was currently still active and his career had yet to conclude, which coincided with the missing data for total years in the major leagues. Baseball-reference.com was unable to provide complete information on education for some players but the missing information was available at mlb.com. The information included: player full name, draft pick number, in which round they were drafted, and the name, location, and institution they were selected from. This information provided the ability to fill in all missing data for level of education. Because this study was only interested in MLB players drafted between 1990 and 2000, all players drafted prior to 1990 were omitted for this study, leaving approximately 1600 observations. Over the course of the eleven seasons of offensive performance data collected, many of the players were not drafted between the identified years of 1990-2000 because they were active players during those seasons their statistics were in the data, and those 3808 were eliminated.

Many of the players competing in MLB were not drafted. In accordance with the rules set forth by MLB with regard to the draft, a player is eligible for the draft if he is a resident of the United States, its territories or Canada, has never signed a major league or minor league contract, or is in high school in the U.S regardless of where he was born. A player is labeled as an amateur free agent if he enters the draft and goes unselected. He may then sign with any club as an undrafted free agent. Additionally, those players who are not eligible for the draft because they do not meet the criteria to enter the draft, for example they are residents of other countries, were, are not part of this study. These players were identified in the data set as AFA meaning they signed with the club as an amateur free agent and were not selected in the draft. Approximately 600 players were identified as AFA and subsequently not chosen for this study, leaving the data set to be analyzed containing 1384 observations. Observations were collected each year a player was active in the major leagues between 1990 and 2000. Since 1994 was a strike year, all observations in that year were omitted, totaling approximately 47. Finally, there were players who were in the major leagues for less than five years during the sample period. These players were either injured, released, retired, sent back down to the minors, still playing beyond the year 2000, quit, or deceased leaving, a sample of 560 observations for this study (N=112).

In preparing the dataset for analysis one final elimination of variables occurred leaving the dataset with the following performance variables: games (G), at bats (AB), runs (R), hits (H), doubles (2B), triples (3B), home runs (HR), runs batted in (RBI), total bases (TB), base on balls (BB), strike outs (SO), caught stealing (CS), on base percentage (OBP), slugging (SLG), batting average (AVG), sacrifice flies (SF), hit by pitch (HBP),

intentional walks (IBB), grounded into a double play (GDP), total plate appearances (TPA), extra base hits (XBH), stolen base percentage (SB%) and on base plus slugging percentage (OPS). The data set was reduced from an initial 29 performance variables to 23. The six variables excluded were: fly out/ground out ratio (GO/AO), fly outs (AO), ground outs (GO), number of pitches (NP), sacrifice hit (SH), and stolen bases allowed (SB). These variables were excluded because they did not contribute to a players' offensive performance as it related to this study.

Sample Characteristics. The sample included information about all MLB players classified as hitters from the 1990 through 2000 seasons. The sample was restricted to those drafted between the 1990 and 2000 seasons, and for the purposes of this study, only the first five years of a player's career were examined. The choice was made to only examine the first five years of a player's career because of the rules about free agency. Player mobility is restricted in MLB; there is a contractual reserve clause that prohibits players from negotiating contracts for their services with any league member except their current team which was a mandated provision within a player's contract until 1976 (Maxcy, 2002). Starting with the 1977 Basic Agreement, and outlined in the current Basic Agreement, players with 6 or more years of MLB service are eligible to become free agents and negotiate with other teams at the expiration of their current contract without any restrictions or qualifications (2007-11 CBA, MLBPA). This study asserted that after five years in MLB any prior training or player development (minor leagues or college) would no longer be influential to the players they became in year six. In year six players take on the traits and learned behaviors of being in the league and begin to become more predisposed by free agency and money.

The sample consisted of all MLB players who were drafted between 1990 and 2000 and were called up to the major league level. These years and their corresponding numbers represented how many players were in the study during the eleven year time period. For example, there were approximately 50 players drafted in 1990 and 53 players drafted in 1991, however the first year a player from the 1990 cohort was called up to the major leagues was 1992 and there were four: Bret Boone, David Hulse, Dan Wilson, and Kevin Young. Therefore, Figure 1 reflects no observations for the years 1990 or 1991. The number of players by education who were called up to MLB from 1990-2000 is shown in Figure 2.

Figure 1. Number of Players Called Up to Major Leagues in Draft between 1990-2000

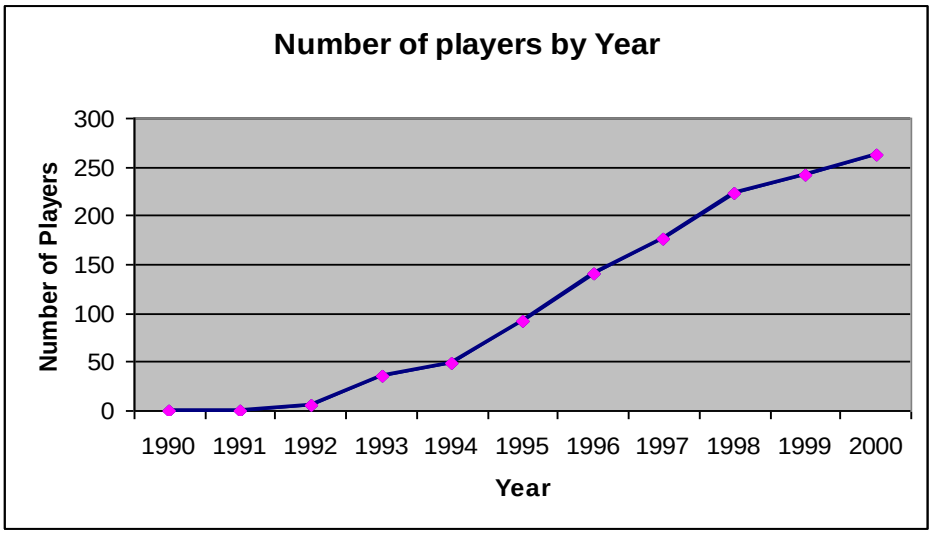
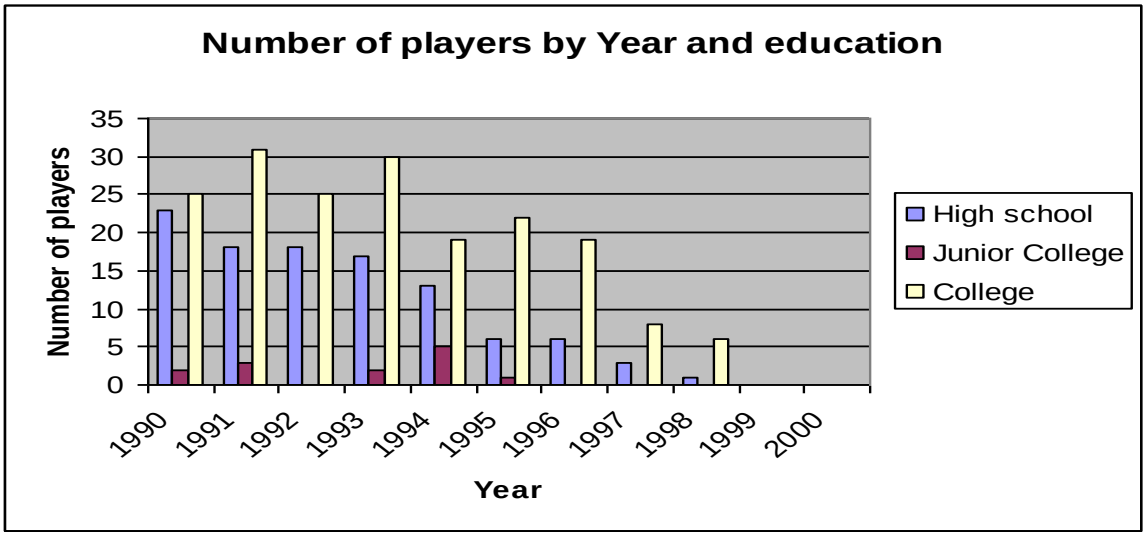


Figure 2. Number of Players by Education and Called Up to MLB from 1990-2000



Based on the information gathered from the literature review there were two human capital characteristics identified for the study: level of schooling obtained and years of on-the-job training. One measurement of human capital for the players in the sample was level of education and it consisted of three categories: high school only, junior college, or four-year college. Players labeled as high school only are players at the time of their draft year who were selected with a high school education only. Many players enter the draft immediately after high school to see where they are selected and then make the decision to go to college. For example, Barry Bonds was drafted in the second round of the 1982 amateur draft by the San Francisco Giants coming out of high school, and he did not sign. Instead he attended Arizona State University for three years and reentered the draft in 1985 and was selected number six in the 1st round by the Pittsburgh Pirates (MLB.com, 2007).

The next measurement of human capital, on-the-job training (OJT), was defined by two variables. The first was the previously identified education variable - specifically the four-year college category. When an athlete decides to attend and play baseball in college under the rules of MLB he must stay for at least three seasons, therefore the college variable can be utilized for both measurements of human capital. The second variable (MiLB), was the total number of seasons each player in the sample spent in the minor league system before getting the call up to the major leagues.

This data set spans eleven seasons of major league baseball performance statistics, 1990 – 2000; the data is both time series and cross-sectional. Data characterized as time series data are measurements of a variable taken at regular intervals over time. Cross-sectional data is parallel data collected on many individuals at a given time. Panel data contains multiple observations over multiple time periods. The data set for this study was in actuality an example of an unbalanced panel data set because individual characteristics including demographics (age, education attainment, employment status), performance statistics, team, draft information (team, year, round), and player information (position and career length) were collected for different persons and different years and not all players were observed in the same years. The data is right censored because not all observations have an end point. Many of the players are still active and were still playing in MLB after the 2000 season.

Data Coding. It was necessary to code six variables. Level of education was initially coded as discrete where ‘1’ stood for high school, ‘2’ for junior college and ‘3’ for four year college, and ‘4’ for unknown. However, it seemed far more efficient to create indicator variables for each level of education, therefore, three new indicator, or

dummy, variables were created: high school = 1 vs. non high school = 0, collegetyr = 1 vs. not collegetyr = 0 and collegefy = 1 vs. not collegefy = 0. The variables *Team* and *Position* were qualitative and coded on a nominal scale, since no team was considered better than the other nor was one position of greater magnitude than the other (hence, these are not ordinal variables). For the team variable, the teams were arranged in alphabetical order then assigned a number 1-30 representing the number of teams in MLB during the study timeframe (Table 1).

Table 1. Teams in Each MLB League

National League	**American League**
Arizona Diamondbacks	Baltimore Orioles
Atlanta Braves	Boston Red Sox
Chicago Cubs	Chicago White Sox
Cincinnati Reds	Cleveland Indians
Colorado Rockies	Detroit Tigers
Florida Marlins	Kansas City Royals
Houston Astros	Los Angeles Angels
Los Angeles Dodgers	Minnesota Twins
Milwaukee Brewers	New York Yankees
New York Mets	Oakland Athletics
Philadelphia Phillies	Seattle Mariners
Pittsburgh Pirates	Tampa Bay Rays
San Diego Padres	Texas Rangers
St. Louis Cardinals	
Montreal Expos*	

* The Montreal Expos entered Major League Baseball as an expansion team in 1969 and relocated to Washington D.C., December 3, 2004 and were renamed the Washington Nationals.

The team that drafted the player into MLB was also recorded and followed the same nominal scale as the team variable. The variable position was coded in accordance with the scoring rules of major league baseball and after elimination of the position of pitcher, the combination of the three outfield positions and the addition of the designated hitter, 2-8 remained. The nominal scale was as follows; catcher (2), first baseman (3), second baseman (4), third baseman (5), short stop (6), outfield (7) and designated hitter (8) (Table 2).

Table 2.

Positions by Coded Number

MLB Positions	Number	StudyPosition	Number
*Pitcher	1	Catcher	2
Catcher	2	First Base	3
First Base	3	Second Base	4
Second Base	4	Third Base	5
Third base	5	Short Stop	6
Short Stop	6	*Outfield	7
*Left field	7	DH	8
*Centerfield	8		
*Right Field	9		

The last variable that was coded was an indicator variable for active players. The active variable indicated if the player was still currently an active major league baseball player as of the year 2000 and was coded as ‘1’ for active and ‘0’ for no longer active. A complete list of all variables is presented in Appendix A.

Data Analysis

Data were analyzed using the PAWS statistical software package (version 18.0: SPSS, Chicago, IL, U.S.A.). Before the actual analyses, the data were explored for accuracy of entries, missing data, and statistical assumptions. Data analysis procedures included basic descriptive statistics, two independent sample t-tests or Mann-Whitney tests, linear regression, life table, and Cox proportional hazards regression. Two independent sample t-test or Mann-Whitney test (depending on normality of the variables as checked by Shapiro-Wilk test at 1% level) was performed to compare the differences in performance between subjects with high school education and some college education.

Linear regression was used to determine whether education, OJT and other independent variables (such as age, team drafted, round drafted, certain offensive performance statistics, years in the minors for an amateur baseball player once he was called up to the big league, position dummy variables) predicted the offensive production (AVG, RBI, OBP, SLG, and OPS) in MLB and the mean time to reach the major leagues. Position were grouped into catcher (number 2), infielder (number 3,4, 5, 6) and outfielder (number 7) and further dummy coded into three new variables. The two (infielder and outfielder) of the three new variables was included in the linear regression model with catcher was used as the reference category. Life tables were used for time-to-event (end of MLB career) comparisons between high school players and players with some college education. Cox regression analysis was used to determine the effect of various factors (age, education, team drafted, round drafted, certain offensive performance statistics, years in the minors for an amateur baseball player once he is called up to the big leagues, position) that were thought to be associated with failure of MLB career.

Research Questions

This chapter asserted that through the combination of the human capital variables (on-the-job (OJT) training while in college and the amount of schooling attained) the more productive a MLB player is going to be as measured by his offensive performance. In order to test this assertion six hypotheses were developed and were tested using statistical methods. The study focused on the following research hypotheses.

H1:

H_0: $\beta 1 = 0$ *Education (high school vs. some college) does not significantly affect offensive performance (AVG, RBI, OBP, SLG, and OPS) in MLB*

H_1: $\beta 1 \neq 0$ *Education (high school vs. some college) significantly affects offensive performance (AVG, RBI, OBP, SLG, and OPS) in MLB*

H2

H_0: $\beta 2 = 0$ *On the job training (OJT) does not significantly affect offensive performance (AVG, RBI, OBP, SLG, and OPS) in MLB*

H_1: $\beta 2 \neq 0$ *On the job training (OJT) significantly affects offensive performance (AVG, RBI, OBP, SLG, and OPS) in MLB*

H3:

H_0: $\beta 3 = 0$ *Education (high school vs. some college) does not significantly affect the mean time to reach the Major Leagues for a baseball player*

H_1: $\beta 3 \neq 0$ *Education (high school vs. some college) significantly affect the mean time to reach the Major Leagues for a baseball player*

H4

***H_0**: $\beta i = 0$ for all independent variables (i.e., age, education, team drafted, round drafted, certain offensive performance statistics (AVG, RBI, OBP, SLG, and OPS), years in the minors) on the survival of MLB.*

***H_1**: at least one of the $\beta i \neq 0$ for all independent variable (i.e., age, education, team drafted, round drafted, certain offensive performance statistics (AVG, RBI, OBP, SLG, and OPS), years in the minors) on the survival of MLB.*

H5

***H_0**: The overall survival is not different between high school players and players with 4 year or junior college*

***H_1**: The overall survival is different between high school players and players with 4 year or junior college*

Assumptions and Limitations

In linear regression it is assumed that the data meet the following assumptions: homoscedasticity, linearity and normality. The assumption of homoscedasticity states that the variability for one continuous variable is roughly the same at all values of another continuous variable, which was tested by Breusch-Pagan test. The assumption of linearity is important because regression analysis only tests for a linear relationship between the independent variables and the dependent variable. If the relationship between the variables is not linear, the linear regression will not be able to identify a non-linear relationship between the DV and IV(s).

In regression analysis it is sometimes assumed that the unobservable equation error term is normally distributed. A normality test check can be done in multiple ways.

The two most popular, examination of the histograms of the dependent variables or checking the Q-Q plot. The residuals for survival data are different than for other types of models because of censoring. Though the Cox regression model is non-parametric to the extent that no assumptions are made about the form of the baseline hazard, there are still a number of important issues which need be assessed before the results can be safely applied in this setting. First is the issue of non-informative or random censoring. To satisfy this assumption, the design of the underlying study must ensure that the mechanisms giving rise to censoring of individual subjects are not related to the probability of an event occurring. The second key assumption in the Cox model is that of proportional hazards. This means that the survival curves must have hazard functions that are proportional over time (i.e. constant relative hazard), which can be evaluated using "log-log" plots in the two-sample comparison case. The major limitation of all regression techniques is causality. Regression analysis ascertains statistical relationships and not the underlying causal mechanism.

Chapter Four: Results

This study first examined the relationship between a MLB baseball player's human capital and player development, and second investigated the effect a player's human capital and education level had on certain offensive performance outcomes once he was called up to the major league. By concentrating on the first five seasons of a player's career the following was examined how a player's combination of education and on the job training have influenced him during the early portion of his career. Four research hypotheses were developed. The independent variables were education track, OJT and other potential predictors which were thought to impact the offensive performance outcomes and the mean time to reach the major league. In addition, the study aimed to explore what influenced MLB career longevity. Two additional research hypotheses were developed in order to examine whether the overall survival in MLB depended on various factors such as age, education, team drafted, round drafted, certain offensive performance statistics, years in the minors for an amateur baseball player once he was called up to the big league and whether players with some college had a higher chance of survival than high school players.

The organization of this chapter is as follows, first the data demographics are presented describing the information gathered on all MLB players included in the study. Second, the summary of findings for each of the research questions and hypotheses was provided in detail and finally the conclusion of what it all means is offered at the end of the chapter.

Demographic Characteristics of the study population

The study consisted of 112 male MLB players drafted between 1990 and 2000. All of these players were aged 18 – 23. The educational background of the population by year they were drafted is presented in Table 3. Of the 112 MLB players, 35 (31.3%) players were drafted with only a high school education and 77 (68.8%) players were drafted from a four year or two year college or university then to minor league baseball before being called up to the major league. The average time to reach the major league appeared to be 3.29 years. The average round a player was drafted in was 7.53. The average years in minor league was 3.29 with a minimum of 1 year and maximum of 6 years. The average OJT (years in minors plus college) was 3.87 years with a minimum of 0 year and maximum of 9 years. The average total seasons played during1990-2000 were 6.29 seasons with a minimum of 4 seasons and maximum of 9 seasons. The descriptive statistics for various performance variables is shown in table 4. All playing positions with the exception of pitcher are represented in the sample.

Table 3

Education background of the population by year they were drafted (N=112)

Draft Year	High School	College	Total
1990	14	18	32
1991	13	26	39
1992	5	18	23
1993	3	8	11
1994	0	6	6
1995	0	1	1
Total	35	77	112

Table 4

Descriptive statistics for various performance variables (N=112)

Variable	Mean	SD	Median	Minimum	Maximum
Games	80.146	30.836	81.400	13.000	149.400
At Bats	254.148	133.009	237.100	28.000	581.800
Runs	36.241	22.803	31.400	3.000	97.600
Hits	69.239	41.037	60.900	5.000	162.400
Home Runs	6.786	5.500	5.500	0.200	23.400
Runs Batted In	32.227	19.984	27.900	1.800	87.200
Total Bases	106.673	64.238	93.100	7.800	279.400
Base on Balls	24.323	15.605	20.000	2.400	66.600
Strike Outs	46.246	22.376	45.000	8.200	103.800
On Base Percentage	0.323	0.039	0.325	0.204	0.454
Slugging	0.387	0.070	0.385	0.220	0.601
Batting Average	0.255	0.036	0.260	0.155	0.370
Total Plate Appearances	285.038	149.741	262.300	32.800	656.800
On Base plus Slugging	0.710	0.100	0.704	0.424	1.055

Hypotheses testing

The data were then analyzed using various statistical methods to test the six research hypotheses. The following sections provide findings for these hypotheses.

Descriptive statistics was performed for offensive performance by education (Table 5). The results demonstrated that the players taking the path from high school to the minor league to the big league had slightly higher average offensive production scores (AVG, RBI, OBP, SLG, and OPS). Box plots displaying the relevant information of offensive production is shown in Figure 4. The box plots showed that the distribution of AVG, OBP, SLG and OPS was very similar for high school players and players with some college education. The distribution of RBI

is more spread for high school players than players with some college education.

Two independent sample t-test or Mann-Whitney test depending on normality of the variables (checked by Shapiro-Wilk test at 1% level) was performed to compare the differences in performance between subjects with high school education and some college education. AVG, OBP, SLG, and OPS were normally distributed ($p>.01$) while RBI was not normally distributed ($p<.01$). Therefore, two independent sample t-test was applied to AVG, OBP, SLG and OPS while Mann-Whitney test was applied to RBI. The results demonstrated that the players taking the path from high school to the minor league to the major league had slightly higher average offensive production scores (AVG, RBI, OBP, SLG, and OPS) (Table 5). However, significant difference was only found in RBI ($U=1705.00$, $p=.025$). There were no significant difference in AVG (t (110) =.55, $p=.58$), OBP (t (110) =.28, $p=.78$), SLG (t (110) =1.08, $p=.30$), and OPS (t (110) =.84, $p=.41$) between players with high school education and those with some collegiate education. The results supported the null hypothesis for AVG, OBP, SLG and OPS while were in favor of alternative hypothesis for RBI.

Multiple regression was performed to the test the first two hypotheses. Multiple regression tested the effects of the two main independent variables: education and OJT and other potential predictors (age, team drafted, round drafted, years in the minors, position dummy variables) on dependent variables offensive production (AVG, RBI, OBP, SLG, and OPS). Backward elimination procedure was used in order to obtain a reduced model containing only meaningful variables. No variables were entered into the model for OBP, SLG, and OPS. Therefore, only the full model is presented for these three dependent variables. For AVG, the reduced model contains only the variable infielder, $\beta=.012$, t (109)=3.26, $p=.07$. This suggested that players labeled with the

infielder position had higher AVG than catchers. For RBI, age significantly predicted the RBI, β=-2.846, *t* (108)=-2.45, *p*=.016, which suggested that as age increased by one year, the RBI was expected to decrease by 2.846 unit . The infielder position also significantly predicted the RBI, β=10.802, *t* (108=3.03, *p*=.003), which suggested that players labeled with the infielder position had significantly higher RBI than catchers. The estimate of coefficients and regression statistics is shown in Tables 6-12. The results supported the null hypothesis that neither education nor OJT significantly affected the offensive performance.

Table 5

Descriptive Statistics for Offensive Production by Education (high school (N=35) vs. some collegiate (N=77)

Variable	Education	Mean	SD	Median	Minimum	Maximum
AVG	High School	0.258	0.041	0.261	0.155	0.370
	Some collegiate	0.254	0.034	0.259	0.162	0.320
RBI	High School	38.869	21.747	33.200	1.800	82.800
	Some collegiate	29.208	18.497	25.200	3.200	87.200
OBP	High School	0.325	0.045	0.326	0.225	0.454
	Some collegiate	0.323	0.036	0.325	0.204	0.403
SLG	High School	0.398	0.077	0.405	0.226	0.601
	Some collegiate	0.383	0.066	0.379	0.220	0.558
OPS	High School	0.721	0.115	0.716	0.480	1.055
	Some collegiate	0.704	0.093	0.703	0.424	0.924

Figure 3. Boxplot for Offensive Production by Education (high school (N=35) vs. some collegiate (N=77))

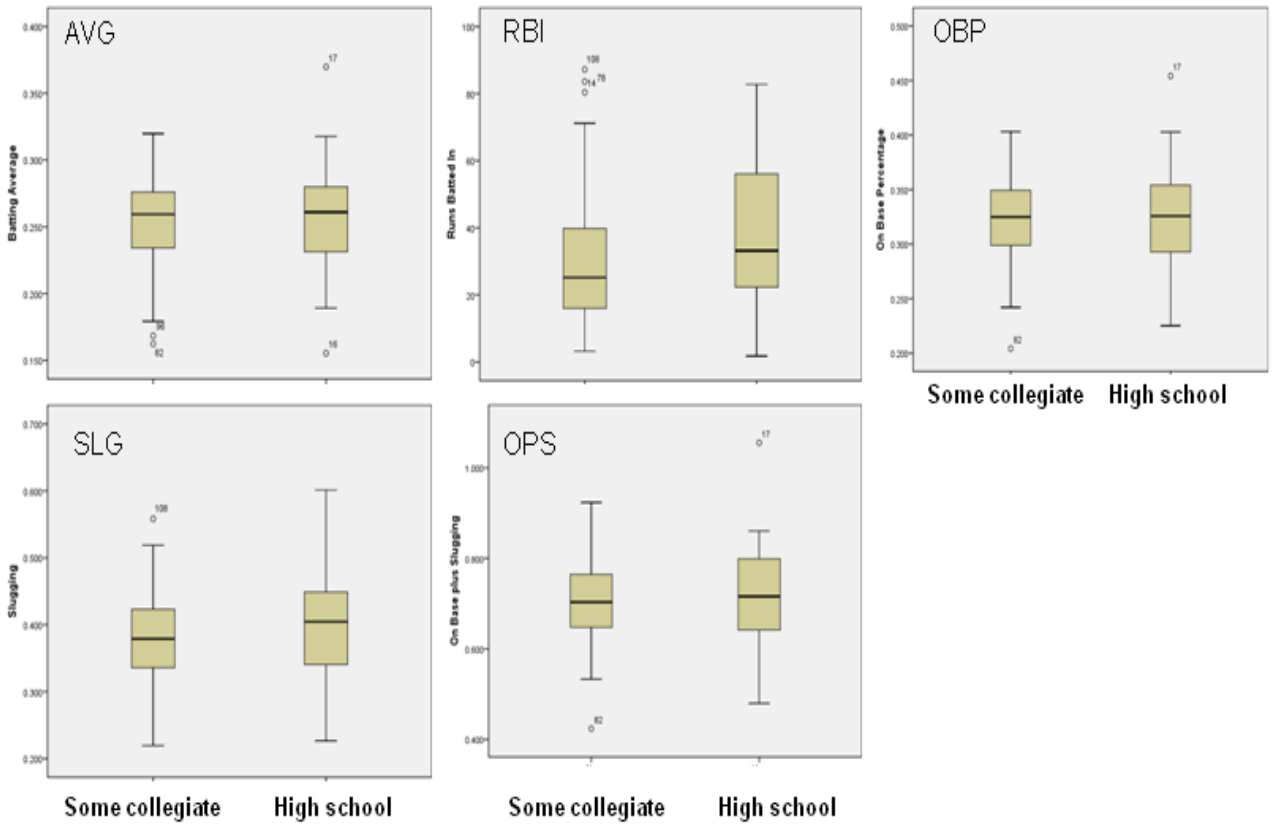


Table 6

Multiple Regression Coefficients for Various Predictors on Batting Average (AVG) (full model)

	B	*SE*	T	*p*-value
(Constant)	0.273	0.136	2.004	0.048
Age	-0.001	0.007	-0.149	0.882
Education	0.009	0.019	0.481	0.631
OJT	0.001	0.003	0.213	0.832
Years in the Minor Leagues	-0.005	0.004	-1.195	0.235
Team that drafted player	0.000	0.000	-0.587	0.559
Round Players was drafted in	0.000	0.000	1.045	0.299
Infielder	0.019	0.010	1.812	0.073
Outfielders	0.008	0.011	0.748	0.456

Note. R2=.063, F (8,102)=.86, p=.55

Table 7

Multiple Regression Coefficients for Various Predictors on Batting Average (AVG) (reduced model)

	B	*SE*	t	*p*-value
(Constant)	0.249	0.005	54.084	0.000
Infielder	0.012	0.007	1.807	0.074

Note. R2=.029, F (1,109)=3.26, p=.07

Table 8

Multiple Regression Coefficients for Various Predictors on Runs Batted In (RBI) (full model)

	B	*SE*	T	*p*-value
(Constant)	61.011	70.177	0.869	0.387
Age	-1.333	3.403	-0.392	0.696
Education	3.630	9.784	0.371	0.711
OJT	-0.526	1.504	-0.350	0.727
Years in the Minor Leagues	-2.059	1.956	-1.053	0.295
Team that drafted player	-0.149	0.228	-0.653	0.516
Round Players was drafted in	-0.048	0.228	-0.210	0.834
Infielder	14.411	5.381	2.678	0.009
Outfielders	5.258	5.574	0.943	0.348

Note. R^2 = .165, F (8, 102) = 2.53, p =.02

Table 9

Multiple Regression Coefficients for Various Predictors on Runs Batted In (RBI) (reduced model)

	B	*SE*	T	*p*-value
(Constant)	85.057	23.923	3.555	0.001
Age	-2.846	1.163	-2.448	0.016
Infielder	10.802	3.563	3.032	0.003

Note. R^2 = .13, F (2, 108) = 8.16, p =.001

Table 10

Multiple Regression Coefficients for Various Predictors on On Base Percentage (OBP)

	B	*SE*	T	*p*-value
(Constant)	0.381	0.148	2.571	0.012
Age	-0.003	0.007	-0.356	0.723
Education	0.000	0.021	-0.001	0.999
OJT	0.000	0.003	0.043	0.966
Years in the Minor Leagues	-0.003	0.004	-0.793	0.429
Team that drafted player	-0.001	0.000	-1.138	0.258
Round Players was drafted in	0.001	0.000	1.038	0.302
Infielder	0.015	0.011	1.364	0.176
Outfielders	0.007	0.012	0.599	0.550

Note.R^2 = .044, F (8, 102) = .59, p =.79

Table 11

Multiple Regression Coefficients for Various Predictors on Slugging (SLG)

	B	*SE*	t	*p*-value
(Constant)	0.552	0.263	2.103	0.038
Age	-0.007	0.013	-0.587	0.559
Education	0.018	0.037	0.502	0.617
OJT	0.003	0.006	0.533	0.595
Years in the Minor Leagues	-0.011	0.007	-1.553	0.124
Team that drafted player	-0.001	0.001	-0.752	0.454
Round Player was drafted in	0.000	0.001	0.508	0.612
Infielder	0.023	0.020	1.133	0.260
Outfielders	0.010	0.021	0.463	0.644

Note.R^2 = .055, F (1, 109) = .75, p =.65

Table 12

Multiple Regression Coefficients for Various Predictors of On Base + Slugging (OPS)

	B	*SE*	t	*p*-value
(Constant)	0.955	0.378	2.529	0.013
Age	-0.011	0.018	-0.610	0.543
Education	0.015	0.053	0.291	0.771
OJT	0.003	0.008	0.411	0.682
Years in the Minor Leagues	-0.014	0.011	-1.362	0.176
Team that drafted player	-0.001	0.001	-0.947	0.346
Round Player was drafted in	0.001	0.001	0.740	0.461
Infielder	0.039	0.029	1.333	0.186
Outfielders	0.015	0.030	0.516	0.607

Note. R^2 = .054, F (8,102) = .73, p =.66

The time to reach the major league was computed as the time taken by a player to reach the major league from the year he was drafted (i.e., year player was called up-year drafted). Descriptive statistics was performed for the time to reach the major league by education (Table 13). Boxplots displaying the distribution of the mean time to reach the major league by education level (high school or attended college) were included to visualize the data (Figure 5). The boxplot showed that the distribution for players that have some college education is more spread than the distribution of high school players.

Multiple regression was performed to the test the hypotheses 3. Multiple regression tested the effects of the main independent variables: education and other potential predictors (team drafted, round drafted, position dummy variables) on the dependent variable mean time to reach the major league for a baseball players. The results demonstrated that the overall model was significant (R^2 = .179, F (5,106) = 4.63,

p=.001)), which suggested that the predictors explained 17.9% of the variance in mean time. The estimate of coefficients is shown in Table 14. The results demonstrated that education, round a player was drafted in, and outfielder position significantly predicted the mean time to reach the major league. In particular, the estimated coefficient for education is .77 (p<.001) suggesting that when education level is changed from college (coded as "0") to high school (coded as "1"), it is expected that the mean time to reach the major league is increased by .77 years. The estimated coefficient for round drafted was 0.035 (p=.003) suggesting that when round is increased by 1 unit, the mean time to reach the major league is increased by 0.035 years. In addition, the estimated coefficient for outfielder is -.713 (p=.016) suggesting that the players in outfielder position spent a shorter time to reach major league than the players in catcher position.

Table 13

Descriptive Statistics for Mean Time to Reach the Major League by Education (high school (N=35) vs. some collegiate (N=77))

Variable	Education	Mean	SD	Median	Minimum	Maximum
Mean time	High School	3.66	1.16	4.00	1.00	6.00
	Some collegiate	3.10	1.01	3.00	1.00	6.00

Figure 4. Boxplot for Mean Time to Reach the Major League by Education (high school (N=35) vs. some collegiate (N=77))

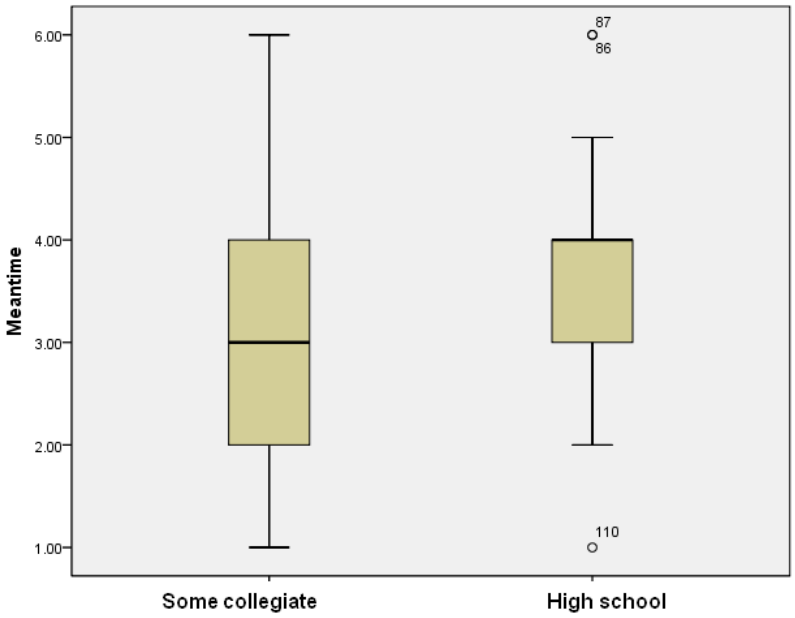


Table 14

Multiple Regression Coefficients for Various Predictors on Mean Time to Reach the Major League (full model)

	B	*SE*	T	*p*-value
(Constant)	3.642	0.329	11.082	<.001
Education	0.772	0.213	3.624	<.001
Team that drafted player	-0.022	0.012	-1.822	0.071
Round Players was drafted in	0.035	0.011	3.073	0.003
Infielder	-0.503	0.283	-1.776	0.079
Outfielders	-0.713	0.290	-2.459	0.016

Note. R^2 = .179, F (5,106) = 4.63, p=.001

Life table and Cox regression (Cox Proportional Hazards Modeling) was used to test the hypotheses 4 and 5 and to estimate the proportion of high school players vs. players with some college education survive, past their first season in the major league, their 2nd season, their 3rd season, their 4th season and finally on to their 5th season. Survival analysis, including life table and Cox regression, is a way of examining the time to an event for one or more groups. In the present study, the event of interest is the end of MLB career for a player. The advantage of survival analysis is it accounts for cases where data about the terminal event is unavailable (censored) due to study cut-off dates. For example, the data in the present study is "right-censored" since many players continued to play beyond the sampled dates. Life table was used to calculate the survival function (i.e., the proportion of the population who were still in MLB in a given time interval). Cox regression model was used to determine how career length of a MLB player was affected by several explanatory variables. The final model from the Cox regression analysis yielded an equation for the hazard as a function of several explanatory variables. The assumptions of survival analysis are that censored and uncensored cases do not differ in terms of the likelihood of the terminal event does not depend upon when the time of the entry event. In Cox's model no assumption is made about the probability distribution of the hazard. However, it is assumed that the hazard ratio does not depend on time. In the present study, these assumptions were fulfilled. The advantage of life table is being able to make a credible analysis without knowing the exact times of censoring. However, life table is not efficient in handling withdrawals, which could be a source of bias. In addition, the choice of the interval is arbitrary. The method assumes that withdrawal occurs at mid-interval which may not be the case. The advantage of Cox

regression is it can incorporate multiple qualitative and quantitative factors, while the limitation is it does not accommodate variables that change over time.

The life table analysis results showed 100% of the high school players and players with four year or junior college survived the first three seasons. The proportion of surviving for high school players for the fourth season was still 100%, while it decreased to 99% for players with four year or junior college. The proportions of surviving for the fifth season were 97% for both high school player and players with four year or junior college. The results showed there is no significant difference in the overall survival between high school players and players with four year or junior college, Wilcoxon (Gehan) statistic (1) = .18, p = .68. The Plot of survival for high school players and players with some college education is shown in Figure 6.

Cox regression was performed to determine the effect of age, education, team drafted, round drafted, certain offensive performance statistics (AVG, RBI, OBP, SLG, and OPS), and years in the minors, and player's position on overall survival of MLB. The offensive performance statistics were added in the model one at a time.

The results showed that the overall models were not significant. The parameter estimates and regression statistics are listed in table 15-19. The offensive performance statistics (RBI, SLG, and OPS) had a significant effect on the survival of MLB. The hazard ratio (exp(β)) was less than 1 for the three variables, which suggested that the players with a better offensive performance tend to have longer careers in MLB. None of the other independent variables appeared to significantly predict the survival of MLB (Table 15-19). The findings supported the hypothesis that the overall survival is not different between high school players and players with 4 year or junior college.

Figure 5. Plot of Survival for High School Players (denoted by green line) and Players with Some College Education (denoted by blue line)

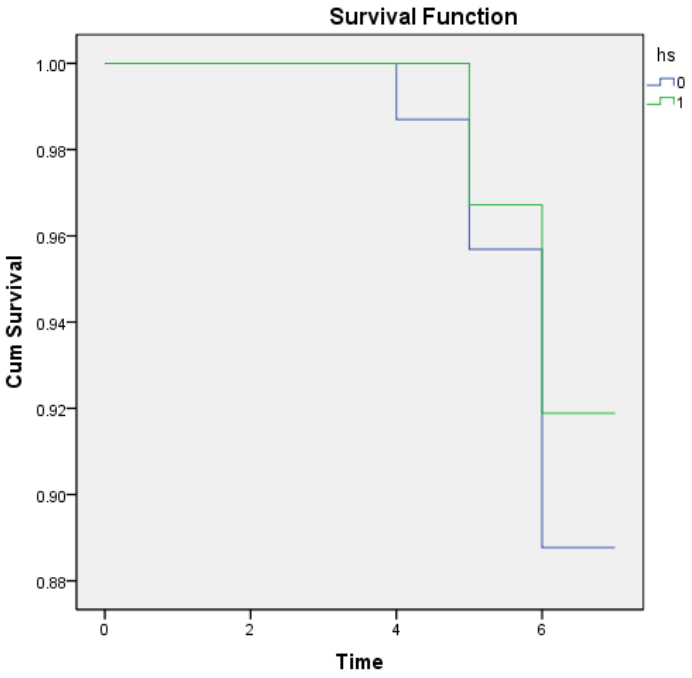


Table 15

Cox Regression Coefficients for various predictors and AVG on overall survival of MLB

	B	SE	Wald	p-value
Age	-0.194	0.616	0.099	0.753
Education	-1.257	1.865	0.455	0.500
Team that drafted player	0.002	0.048	0.002	0.966
Round Players was drafted in	-0.066	0.070	0.898	0.343
Batting average (AVG)	-15.540	10.453	2.210	0.137
Years in Minor Leagues	0.058	0.411	0.020	0.887
Catcher			0.797	0.671
Infielder	10.368	157.129	0.004	0.947
Outfielder	11.062	157.128	0.005	0.944

Note. χ^2 (8) =6.11, p=.635

Table 16

Cox Regression Coefficients for Various Predictors and RBI on Overall Survival of MLB

	B	SE	Wald	p-value
Age	-0.019	0.543	0.001	0.972
Education	-0.442	1.736	0.065	0.799
Team that drafted player	0.010	0.050	0.042	0.838
Round Players was drafted in	-0.068	0.067	1.036	0.309
Years in Minor Leagues	-0.023	0.425	0.003	0.956
Catcher			0.342	0.843
Infielder	10.907	167.854	0.004	0.948
Outfielder	11.370	167.854	0.005	0.946
Runs batted in (RBI)	-0.079	0.040	3.986	0.046

Note. χ^2 (8) =8.90, *p*=.351

Table 17

Cox Regression Coefficients for Various Predictors and OBP on Overall Survival of MLB

	B	SE	Wald	p-value
Age	-0.190	0.620	0.094	0.759
Education	-1.335	1.873	0.508	0.476
Team that drafted player	-0.003	0.049	0.004	0.948
Round Players was drafted in	-0.060	0.070	0.724	0.395
Years in Minor Leagues	0.069	0.396	0.030	0.863
Catcher			0.674	0.714
Infielder	10.410	151.959	0.005	0.945
Outfielder	11.059	151.958	0.005	0.942
On base percentage (OBP)	-14.301	10.071	2.017	0.156

Note. χ^2 (8) =5.97, *p*=.651

Table 18

Cox Regression Coefficients for Various Predictors and SLG on Overall Survival of MLB

	B	SE	Wald	p-value
Age	-0.171	0.554	0.095	0.757
Education	-1.272	1.695	0.563	0.453
Team that drafted player	0.004	0.048	0.006	0.938
Round Players was drafted in	-0.090	0.075	1.426	0.232
Years in Minor Leagues	0.032	0.420	0.006	0.939
Catcher			1.287	0.525
Infielder	10.634	147.197	0.005	0.942
Outfielder	11.532	147.197	0.006	0.938
Slugging (SLG)	-16.846	6.892	5.974	0.015

Note. χ^2 (8) =9.73, *p*=.285

Table 19

Cox Regression Coefficients for various predictors and OPS on overall survival of MLB

	B	SE	Wald	p-value
Age	-0.218	0.597	0.133	0.715
Education	-1.373	1.779	0.596	0.440
Team that drafted player	-0.008	0.048	0.025	0.875
Round Players was drafted in	-0.070	0.070	0.997	0.318
Years in Minor Leagues	0.046	0.418	0.012	0.913
Catcher			0.712	0.701
Infielder	10.790	141.561	0.006	0.939
Outfielder	11.463	141.560	0.007	0.935
On base + slugging (OPS)	-10.391	4.727	4.832	0.028

Note. χ^2 (8) =.654, *p*=.372

Discussion

Finding prospects is critical for any professional sport. MLB has invested significant resources in player selection process (Olson & Schwab, 2000). Yet, decisions regarding whom to scout, whom to draft, whom to promote, whom to acquire, etc are still difficult. One of the major selection philosophies is focusing on college players is a more sensible strategy as the return on investment of college prospects is higher because they are more predictable and reach the majors faster. However there are counterarguments as evidences were also found that performance of high school players are comparable to college players. Therefore, one of the main aims of this study was to examine the impact of college education on the outcomes of the early stages of a professional baseball player's career.

According to the Human Capital Theory, educational investments and on-the job training will make a player more productive. By attending and playing baseball in college, a player might develop into a more rounded individual who is more cognizant of team goals and committed to sacrificing personal statistics for wins. This was supported by several previous studies (e.g., Spurr, 2000; Newman, 2003; Lee, 2010). The main arguments are college players are more mature, which enable them to better handle the stresses of competition at the major league level. Second, college players play against stronger and more advanced competition more often than high school players and they play a longer schedule and usually practice year round (Lewis, 2003). Finally, college players may go off the board more quickly than high school players. The age, experience and maturity contribute to shorter incubation periods for college prospects: two years from draft to the majors for top players drafted from college, as opposed to three-and-a

half to four years for top players drafted from high school (Law, 2007). Therefore, a college player is considered as a better investment than a high school player. Such drafting strategy is the well known Beane's risk-averse philosophy (Lewis, 2003), and is being shared by a number of general managers (GMs) and top management teams (TMTs), such as Marlins scouting director Jim Fleming (Newman, 2003). Contradicting this philosophy, however, a study conducted among 30 college players and 30 high school players (Wassermann et al., 2005) revealed no significant differences in two offensive statistics (on base percentage and OPS) between these two groups of players. The study argued the results may be yielded due to several reasons. First, high school players may adapt more easily to new changes and college players may have developed a certain approach to hitting from college that contradicts a new approach at the professional level. Second, high school players are usually placed in lower levels of minor league baseball than their counterparts college players, which in turn may even the offensive statistics. Similarly, another study noted that attending college could be viewed as a negative signal because a player is not focusing all of his time on being drafted (Olbrecht, 2007). Furthermore, for some of the best high school prospects, college is not an option as many do not have the academic ability to gain entry into college and players who go to college do not necessarily receive good instruction as they are used in a way that is designed for winning more games, not in a way that maximizes the player's professional development (Law, 2007).

In the present study, out of a sample of 112 players drafted between 1990 and 2000, 35 (31.3%) players were drafted from high school and 77 (68.8%) players were drafted out of college. This indicated that managers shared the Beane's risk-averse

philosophy to favor college players over high school players as college players were drafted more than twice as many as high school players during this period. Similarly, of the 2999 players drafted in the 2004 and 2005 first year player entry drafts, 1978 (66%) were from college (MLN Sports Group, 2005). However, the tendency to choose college players over high school players was not supported by further findings of the present study. Using linear regression, the present study tested the effect of education on five offensive performance measures (AVG, RBI, OBP, SLG, and OPS). The results showed that education did not significantly affect any of these the offensive performance statistics. In fact, high school players even had a slightly higher mean offensive performance than college player although a higher variability was also observed in high school players than college players. Several factors may contribute to the non-significant results. First, the study was based on relatively small sample size (N=112). A more significant result could have been established if much larger data set was used. Second, the overall performance of a player may be difficult to quantify using individual performance data.

Nonetheless, the analysis showed that there is a positional difference in terms of the offensive performances, namely, players in the infield position had better batting average (AVG) and runs batted in (RBI) percentages than catchers. In addition, age significantly predicted the RBI. Further investigation is needed to verify whether the positional difference in performance can be generalized to a larger population. Human Capital Theory was also tested on the mean time to reach the major league. It was expected that if Human Capital Theory was taken into account, the mean time to reach the major league would be shorter for a college player than a high school player since

college is expected to experience a tougher competition and have a higher maturity level which gives them a better chance of reaching the majors faster. The results of the present study showed that the mean time to reach MLB for high school players was about half year longer than players with some college education. Furthermore, the regression analysis showed that education significantly predicted the mean time to reach MLB. Specifically, when education level is changed from college to high school, it is expected the mean time to reach the major league is increased by .77 years. The results also suggested that it took from one year to six years for both high school and college players from draft year to reach the major league. On average, it took 3.10 years for college players while 3.66 years for high school players to reach the major league. The variability as measured by standard deviation was higher for high school players than college players, indicating high school players are slightly more unpredictable than college players. These results are consistent with previous findings (Burger et al., 2006; DelGrippo, 2010; Law, 2007). In a previous report, Law (2007) suggested that college players were faster to the majors than high school players. It only took two years from draft to the majors for top players drafted from college, as opposed to three and a half year to four years for top players drafted from high school. Another report (DelGrippo, 2010) also showed similar trend in that top high school players could take up to five or six years to make an impact, whereas highly rated college players (namely pitchers) can make a team better much sooner. Likewise, Burger and Walters (2009) examined first-round picks taken in the 1990-1997 drafts and found that it took high school players who ultimately became ‘stars’ 2.44 years from the time they were drafted to reach the major league; it took college players who became stars 1.97 years to reach the majors. It took

high school draftees who became ‘good’ players 3.56 years, while only 1.84 for college players who became good big leaguers. High school players who became ‘regular’ big leaguers took 4.31 years to reach the major league compared to 2.25 years for college draftees who became regulars to reach the majors. In addition, the results showed that when round is increased by 1 unit, the mean time to reach the major league is increased by 0.035 years and the players in outfielder position spent a relatively shorter time to reach major league than the players in catcher position.

To summarize, the results are mixed. On one hand, the results support the Human Capital Theory and Beane’s drafting philosophy since education showed a positive effect to shorten the time to reach major league. On the other hand, the results contradict these two theories as college players did not show superior offensive performances than high school players. This finding is consistent with an analysis performed by Van Zandt (2006), who compared the rate of success of the players on Beane’s list against that of all 41 first-round picks from 2002 and found that Beane’s choices are lagging behind in terms of performance. The study further compared Beane’s picks against the top high school picks from the first round and found that again, Beane’s choices did not compare favorably and with the A’s draft picks an average of 3.5 years older than the high school draftees they were not likely to catch up either. The present study provided additional evidence with respect to drafting strategies. Success in baseball performance may be in large part determined by ability and training and there may not be a universal draft strategy. Therefore, it is more beneficial to evaluate players comprehensively instead of an all-college approach.

Another important determinant, on-the-job training (OJT) was analyzed in the present study in order to further test whether Baseball performance is determined by OJT. The OJT was defined by two variables in the present study. The first was the previously identified education variable but specifically the four-year college category. When an athlete decided to attend and play baseball in college under the rules of MLB he must stay for at least three seasons, therefore the college variable can be utilized for both measurements of human capital. The second variable (MiLB), accounted for the total number of seasons each player in the sample spent in the minor league system before getting the call up to the major league. Therefore, OJT is a variable measuring the number of years in college and in the minor league and reflect both the player's education and training. The effect of OJT on the offensive performance was tested using regression analysis. The results demonstrated that OJT did not significantly affect any of the five offensive performance measures, which contradict the Human Capital Theory. Based on Human Capital Theory, individuals tend to benefit from on the job training, and thus may have better offensive performance. For example, Singell (2001) showed that the amount of playing experience a major league baseball manager has, establishes a direct correlation to his team's ability to win games and to the individual performance of his players. Results also showed that players drafted in a later round will be faster to reach the majors. Specifically, a player drafted about 29 rounds later will enable him to reach the majors one year sooner, which requires further investigation.

The results demonstrated that more years of on-the-job training did not improve a player's offensive performance. Multiple factors may contribute to such findings. For example, years of training may not be the best variable to represent on-the-job training

since large variations exist in training intensity, training methods, and training types (film specific or general), etc.

In addition to evaluate offensive performance and mean time to reach MLB, the present study explored the career longevity of a player using survival analysis. An understanding of the factors affecting career longevity is useful in modeling the effects of demographic changes on various outcomes. Previous studies demonstrated that age, player performance and era affected baseball career longevities. In particular, age has an inverse effect on career longevity in that players who started at younger ages will enjoy longer careers (Shulz et al., 1994). Moreover, players on the extremes of first-year performance will have vastly different career longevities. Spurr and Barber (1994) suggested that the first year is critical in the evaluation of a player's talent so the worst players will have exceptionally brief careers and the best players will enjoy long careers. Finally, eras with fewer teams will have more players who will witness shorter careers than eras with league expansion. Drawing from these previous researches, the present study examined the effect of age, education, team drafted, round drafted, certain offensive performance statistics (AVG, RBI, OBP, SLG, and OPS), and years in the minors on career longevity of baseball players. The career longevity was defined as the years remaining in majors and the study only focused on early stages of a professional baseball player's career. Specifically, Life table and Cox regression was performed to determine the proportion of higher school players vs. players with some college education who survived the first five seasons and the effect of various factors on the survival of players in major league. The results indicated a similar proportion of players with high school education and college survived their first five seasons and there is no

significant difference in years remaining in majors between high school and college players. The Cox regression further demonstrated that players with a better offensive performance (RBI, SLG, and OPS) tend to have a longer career in MLB. None of the other independent variables appeared to significantly predict the survival of MLB.

Conclusion

This study addressed the effects of education and on the job training on the performance of professional baseball players using a dataset including 112 players classified as hitters from the 1990 through 2000 seasons. Specifically, the effect of education and on the job training on offensive performance measures, mean time to reach Major League Baseball was tested with linear regression and the effects of multiple factors (age, education, team drafted, round drafted, certain offensive performance statistics (AVG, RBI, OBP, SLG, and OPS), and years in the minors) on career longevity was tested using Cox regression. Overall, the results demonstrated that education significantly affects the mean time to majors, while neither education, nor on-the-job training significantly affects the offensive performance. It was also found that positional differences exist in a few offensive performance measures.

Finally, results demonstrated that the survival of the major league was significantly affected by several offensive performance statistics (RBI, SLG, and OPS), however was not significantly affected by other factors (age, education, team drafted, round drafted, certain offensive performance statistics (AVG, and OBP), and years in the minors).Taken together, these results suggested that due to difficulties and potential pitfalls the MLB managers are facing in choosing among so many highly uncertain prospects, it is more beneficial to fully explore a player's potential by multiple

performance factors instead of using universal strategy that is merely based on the Human Capital Theory.

Chapter Five: Practical Implications

The study is designed to provide empirical evidence of the factors which affect early stages of a professional baseball player's career. A practical implication arising from this study is that there is no universal draft strategy. It is more beneficial to fully explore a player's potential by multiple performance factors instead of merely relying on human capital. The present study demonstrated that education shorten a player's accession into the major league, however neither education, nor on-the-job training significantly affect the offensive performance of 112 players drafted between 1990 and 2000. It was also found that the survival of the major league was significantly affected by several offensive performance statistics (RBI, SLG, and OPS), however was not significantly affected by other factors (age, education, team drafted, round drafted, certain offensive performance statistics (AVG and OBP), and years in the minors). The present study adds to a growing body of literature on understanding the effect of human capital to player performance. From the perspective of the MLB teams, the study is significant in helping managers to develop a more effective drafting strategy. The present study also benefits the amateur baseball players with providing details that help them to make better decisions in their professional careers. The attempt of assessing which attributes have predictive effects on the career duration of the players yields some implication for the drafting process. It is useful for clubs wishing to evaluate trades of established players for draft picks.

Limitations

Certain limitations are present within this study. First, the study used a data set of 112 players drafted between 1990 and 2000. A potential limitation is that it only included

a limited number of players in a certain time period. Therefore, results may not be generalizable to other settings. In order to overcome this limitation, it would be necessary to conduct more comprehensive research which could involve the replication of this study in a number of different settings or with a much larger data set. Second, foreign born players that entered MLB as amateur free agents were not included because the study was limited to those players that graduated high school in the United States or graduated high school in the United States and attended a college or university and were drafted by MLB during the June regular draft. Some of the players that were not included were from countries including Cuba, the Dominican Republic, Venezuela, and other South American countries where players are placed in academies at a young age. The rules for entry into MLB players from countries other than the USA do exists and can include such factors as age, education, and training.

Future Research

The present study only focused on the effect of education and on-the-job training on early stages of a professional baseball player's career. Additional investigation may be warranted to reveal other important factors. For example, the present study showed that the round players drafted significantly affected the mean time reaching the majors. A further study could investigate this in more details. Further investigation with a much larger data set is also recommended in order to reveal the general trend with more statistical power. Finally, a study that examines foreign born players and the training that they receive in the baseball academies as compared to the level of training players in the United States receive prior to entering the minor leagues or the major leagues is an area of interest.

References

Ahlstrom, D., Si, S., & Kennelly, J. (1999). Free-agent performance in major league baseball: Do teams get what they expect? *Journal of Sport Management, 13*(3), 181-196.

Antonietti, R. (2006). Human capital, sport performance and salary determination of professional athletes. Unpublished manuscript.

Becker, G. (1962). Investment in human capital: A theoretical analysis. *The Journal of Political Economy, 70*(5), 9-49.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy, 75*(4), 352-365.

Blass, A. (1992). Does the baseball labor market contradict the human capital model of investment? *The Review of Economics and Statistic,* 261-268.

Blaug, M. (1976). The empirical status of human capital theory: A slightly jaundiced survey. *Journal of Economic Literature, 14*, 827-855.

Cassing, J., & Douglas, R. (1980). Implications of the auction mechanism in baseball's free agent draft. *Southern Economic Journal, 47*(1), 110-121.

Daly, G. (1992). The baseball player's labor market revisited. In Sommers, P.M., (ed), Diamonds *are forever: The business of baseball,(pg 11-28).* Washington D.C.: Brookings Institute.

DeBrock, L., Hendricks, W., & Koenker, R. (1996). The economics of persistence: Graduation rates of athletes as labor market choice. *Journal of Human Resources, 31*(3), 513-539.

Depken, C. A. (2002). Free agency and the concentration of player talent in major league baseball. *Journal of Sports Economics, 3*(4), 335-353.

DuMond, J., Lynch, A., & Platania, J. (2008). An economic model of the college football recruiting process. *Journal of Sports Economics, 9*(1), 67-87.

Ehrenberg, R., & Sherman, D. (1987). Employment while in college, academic achievement, and postcollege outcomes. *The Journal of Human Resources,* 23, 1-23.

Eide, E. R., & Ronan, N. (2001). Is participation in high school athletics an investment or consumption good? Evidence from high school and beyond. *Economics of Education Review, 20,* 431-442.

Ferris, E., Finster, M., & McDonald, D. (2004). Academic fit of student-athletes: An analysis of NCAA division I-A graduation rates. *Research in Higher Education, 45*(6), 555-575.

Fort, R. (1992). Pay and performance: Is the field of dreams barren? In Sommers, P.M (ed), *Diamonds are forever: The business of baseball,(pg 134-157).* Washington D.C.: Brookings Institute.

Fort, R. D., & Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature, 33*(3), 1265-1299.

Freeman, R. B. (1980). The facts about the declining economic value of college. *Journal of Human Resources, 15*(1), 124-142.

Frick, B., Prinz, J., & Winkelmann, K. (2003). Pay inequalities and team performance. *International Journal of Manpower, 24*(4), 472-488.

Frost, R. (2003). Recruiting procedures for men's college basketball programs. (Thesis (M.A.), Ball State University). (Available from Kinesiology Publications (formerly Microform Publications), IIHSP, 1243 University of Oregon, Eugene, OR 97403-1243, USA, tel: 541-346-4114, fax: 541-346-0935)

Goff, B. (2000). Effects of university athletics on the university: A review and extension of empirical assessment. *Journal of Sport Management, 14*(2), 85-104.

Groothuis, P. A., Hill, J. R., & Perri, T. J. (2007). Early entry in the NBA draft: The influence of unraveling, human capital, and option value. *Journal of Sports Economics, 8*(3), 223-243.

Hadley, L., Poitras, M., & Knowles, S. (2000). Performance evaluation of national football league teams. *Managerial and Decision Economics, 21*, 63-70.

Haley, W. (1973). Human capital: the choice between investment and income. *American Economic Review, 63*(5), 929-944.

Harder, J. W. (1992). Play for pay: Effects of inequity in a pay-for-performance context. *Administrative Science Quarterly*, *37*(2), 321-337.

Hendricks, W., DeBrock, L., & Koenker, R. (2003). Uncertainty, hiring, and subsequent performance: The NFL draft. *Journal of Labor Economics*, *21*(4), 857-886.

Hotz, V.J., Xu, L. C., Tienda, M., & Ahituv, A. (2002). Are there returns to the wages of young men from working while in school? *Review of Economics and Statistics,* 84, 221-236.

Horowitz, S. A., & Sherman, A. (1980). A direct measure of the relationship between human capital and productivity. *Journal of Human Resources, 15*(1), 67-76.

Howard-Hamilton, M. F., & Sina, J. A. (2001). How college affects student athletes. *New Directions for Student Services*, (93), 35-45.

Imbroscio, D. L. (1998). Reformulating urban regime theory: The division of labor between state and market reconsidered. *Journal of Urban Affairs*, *20*(3), 233-248.

Kahane, L., & Shmanske, S. (1997). Team roster turnover and attendance in major league baseball. *Applied Economics, 29*(4), 425-431.

Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of Economic Perspectives, 14*(3), 75-94.

Kiecolt, K. J. , & Nathan, L. F. (1985). Secondary analysis of survey data. Beverly Hills, CA: Sage.

Kimball, A., & Freysinger, V. J. (2003). Leisure, stress, and coping: The sport participation of collegiate student-athletes. *Leisure Sciences, 25*(2/3), 115-141.

Krautmann, A. (1990). Shirking or stochastic productivity in major league baseball? *Southern Economic Journal, 56*(4), 961-968.

Krautmann, A., & Oppenheimer, M. (1997). Training in major league baseball: Are players exploited?'. In *Baseball Economics*: *Current Research, J.* Fizel, E. Gustafson & L. Hadley (Eds.), Westport, CT: Praeger.

Krautmann, A. C., Gustafson, E., & Hadley, L. (2000). Who pays for minor league training costs? *Contemporary Economic Policy, 18*(1), 37-47.

Krueger, A. B., & Lindahl, M. (2001). Education for growth: Why and for whom? *Journal of Economic Literature, 39*(4), 1101-1136.

Lahey, D. (1993). *Athletic scholarships: Making your sports pay*. Toronto, Canada: Warwick Publishing.

Law, K. (2007). High school or college? Don't lean too hard either way, retrieved from http://sports.espn.go.com/mlb/draft2007/columns/story?id=2885781 June, 23, 2008

Leeds, M., & Von Allmen, P. (2005). The economics of amatereurism and college sports. *The economics of sports* (2nd Ed., pp. 379). Boston: Pearson/Addison Wesley.

Lehn, K. (1982). Property rights, risk sharing, and player disability in major league baseball. *Journal of Law & Economics, 25,* 344-366.

Levhari, D., & Weiss, Y. (1974). The effect of risk on the investment in human capital. *The American Economic Review, 64*(6), 950-963.

Lewis, M. (2004). *Moneyball: The art of winning an unfair game.* NewYork: W.W. Norton and Company.

Levin, R. C., Mitchell, G., Volcker, P., & Will, G. (2000). The report of the independent members of the Commissioner's blue ribbon panel on baseball economics. *Major League Baseball.*

Light, A. (1998). Estimating returns to schooling: When does the career begin? *Economics of Education Review, 17*(1), 31-45.

Light, A. (2001). In school work experience and the returns to schooling. *Journal of Labor Economics, 19*(1), 65-92.

Long, J., & Caudill, S. (1990). The impact of participation in intercollegiate athletics on income and graduation. *The Review of Economics and Statistics, 73(3), 525-531.*

MacDonald, D. N., & Reynolds, M. O. (1994). Are baseball players paid their marginal products? *Managerial and Decision Economics, 15*(5, Special Issue: The Economics of Sports Enterprises), 443-457.

Mahony, D. F., & Howard, D. A. (2001). Sport business in the next decade: A general overview of expected trends. *Journal of Sport Management, 15*, 275-296.

Maloney, M. T., & McCormick, R. E. (1993). An examination of the role that intercollegiate athletic participation plays in academic achievement. *Journal of Human Resources, 28*(3), 555-570.

Maxcy, J. (1997). Do long term contracts influence performance in major league baseball? In Hendricks, W. (Ed.), *Advances in the economics of sport volume 2 (*157-176). Greenwich, Connecticut: JAI Press.

McCormick, R. E., & Tinsley, M. (1987). Athletics versus academics evidence form SAT scores. *Journal of Political Economy, 95*(5), 1103-1116.

Miceli, T. J., & Scollo, C. (1999). An economic analysis of the reserve system in major league baseball. *Journal of Business & Economic Studies, 5*(1), 55-65.

Michael, R. T., & Tuma, N. B. (1984). Youth employment: Does life begin at 16? *Journal of Labor Economics, 2*(4), 464-476.

Mincer, J. (1974). *Schooling, experience, and earnings.* New York: Columbia University Press.

Mixon Jr., F. G., Treviño, L. J., & Minto, T. C. (2004). Touchdowns and test scores: Exploring the relationship between athletics and academics. *Applied Economics Letters, 11*(7), 421-424.

Mixon Jr., F. G., & Ressler, R. W. (1995). An empirical note on the impact of college athletics on tuition revenues. *Applied Economics Letters, 2*(10), 383-387.

Molitor, C., & Leigh, D. (2005). In school work experience and the returns to two year and four year colleges. *Economics of Education Review, 24*(4), 459-468.

Newman, M. (2003). High School vs. College: Does Either Provide a Better Shot at a "Sure Thing?" November 24, 2003. http://mlb.mlb.com/news/article.jsp?ymd=20030520&content_id=328934&vkey=news_mlb&fext=.jsp&c_id=mlb

Noll, R. G., & Zimbalist, A. (Eds.). (1997). *Sports, jobs and taxes*. Washington, DC: Brookings Institution Press.

Price, S. L. (1997). Is it in the genes? *Sports Illustrated, 87*(23), 52-55.

Quirk, J., & Fort, R. D. (1992, 1997). *Pay dirt: The business of professional team sports*. Princeton, NJ: Princeton University Press.

Robst, J., & Keil, J. (2000). The relationship between athletic participation and academic performance: Evidence from NCAA division III. *Applied Economics, 32*(5), 547-558.

Rosen, S., & Sanderson, A. (2001). Labour markets in professional sports. *Economic Journal, 111*(469), 47-68.

Rosentraub, M. S. (1999). *Major league losers: The real cost of sports and who's paying for it*. New York: Basic Books.

Rottenberg, S. (1956). The baseball players' labor market. *Journal of Political Economy,* 64, 242-258.

Rottenberg, S. (2000). Resource allocation and income distribution in professional team sports. *Journal of Sports Economics, 1*(1), 11-20.

Rumberger, R. W. (1980). An exchange: The economic decline of college graduates: Fact or fallacy? *Journal of Human Resources, 15*(1), 99-112.

Ruhm, C. (1995). The extent and consequences of high school employment. *Journal of Labor Research, 16*(3), 293-303, DOI: 10.1007/BF02685757

Ryan, F. J. (1989). Participation in intercollegiate athletics: Affective outcomes. *Journal of College Student Development, 30*(2), 122-128.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.

Saks, M. (2004). College athletics: Living in a glass bowl. *Athletics Administration, 39*(2), 40-48.

Schwartz, E., & Thornton, R. (1980). Overinvestment in college training?? *Journal of Human Resources, 15*(1), 121-123.

Scully, G. W. (1974). Pay and performance in major league baseball. *American Economic Review, 64*(5), 915-930.

Shughart, W., & Goff, B. (1992). Field of dreams: On the construction of professional baseball talent in colleges and the minor leagues. In Scully G.W. (Ed.), *Advances in the economics of sport* (pp. 91-114). Greenwich, Connecticut: JAI Press.

Sigler, K. J., & Sackley, W. H. (2000). NBA players: Are they paid for performance? *Managerial Finance, 26*(7), 46-51.

Singell Jr., L. D. (1991). Baseball-specific human capital: Why good but not great players are more likely to coach in the major leagues. *Southern Economic Journal, 58*(1), 77-86.

Singell, Jr., L.D. (1993). Managers, specific human capital, and firm productivity in major league baseball. *American Economic Journal*, *21*(3), 47-59.

Smart, D. L., & Wolfe, R. A. (2000). Examining sustainable competitive advantage in intercollegiate athletics: A resource-based view. *Journal of Sport Management, 14*(2), 133-153.

Spurr, S. J. (2000). The baseball draft: A study of the ability to find talent. *Journal of Sports Economics, 1*(1), 66-85.

Spurr, S. J., & Barber, W. (1994). The effect of performance on a worker's career evidence from minor league baseball. *Industrial & Labor Relations Review, 47*(4), 692-708.

Staw, B.M., & Hoang, H. (1995). Sunk Costs in the NBA: Why Draft Order Affects Playing Time and Survival in Professional Basketball. *Administrative Science Quarterly, 40*(3), 474-494.

Stewart, D. W. (1984). *Secondary research: Information sources and methods*. Beverly Hills: Sage Publications.

Stone, R. (1987). The graham plan of 1935: An aborted crusade to de-emphasize college athletics. *North Carolina Historical Review, 64*(3), 274-293.

Sweetland, S. R. (1996). Human capital theory: Foundations of a field of inquiry. *Review of Educational Research, 66*(3), 341-359.

Thorn, J., & Palmer, P. (1984). *The hidden game of baseball.* New York: Doubleday and Company, Inc., Garden City.

Toma, J. D., & Cross, M. E. (1998). Intercollegiate athletics and student college choice: Exploring the impact of championship seasons on undergraduate applications. *Research in Higher Education, 39*(6), 633-661.

Tomasini, N., Frye, C., & Stotlar, D. (2004). National collegiate athletic association corporate sponsor objectives: Are there differences between divisions I-A, I-AA, and I-AAA? *Sport Marketing Quarterly, 13*(4), 216-226.

Trail, G., & Chelladurai, P. (2000). Perceptions of goals and processes of intercollegiate athletics: A case study. *Journal of Sport Management, 14*(2), 154-178.

Wassermann, E., Czech, D. R., Wilson, M. J., & Joyner, A. B. (2005). An examination of the Moneyball Theory: A baseball statistical analysis. *Sport Journal*, 8(1), unpaginated.

Watt, S. K., & Moore III, J. L. (2001). Who are student athletes? *New Directions for Student Services,* (93), 7-18.

Weiss, A. (1995). Human capital vs. signaling explanations of wages. *Journal of Economic Perspectives, 9(*4), 133-154.

Winfree, J. A., & Molitor, C. J. (2007). The value of college: Drafted high school baseball players. *Journal of Sports Economics, 8*(4), 378-393.

Witmer, D. R. (1980). Has the golden age of American higher education come to an abrupt end? *Journal of Human Resources, 15*(1), 113-120.

Wolfe, R. A. (2000). Understanding university athletics: Cultural, strategic, and economic perspectives. *Journal of Sport Management*, *14*(2), 79-84.

Zimbalist, A. (1999). *Unpaid professionals: Commercialism and conflict in big time college sports*. Princeton, NJ: Princeton University Press.

Zimbalist, A. (1992). Salaries and performance: Beyond the Scully model in Sommers, P.M., (ed), *Diamonds are forever: The business of baseball,(pg 109-133).* Washington D.C.: Brookings Institute.

Appendix

yob	**Year of birth**
age	**Age at the time player was drafted**
observ_yr	**Year in which the observation was recorded**
strike_yr	**1994 was a strike year. Dummy variable coded as 0 for no and 1 for yes**
tm_code	**Team player was on during observation year. Coded 1-30**
drft_yr	**Year player was drafted**
drft_tm	**Team that drafted player during the June Regular Draft**
drft_rd	**Round in which player was drafted**
collegefy	**Drafted from a four year college variable. Dummy code 0 no/ 1 for yes**
hs	**Drafted from high school. Dummy coded 0 for no and 1 for yes**
collegety	**Drafted from 2 yr college. Dummy coded 0 for no and 1 for yes**
milb	**Minor league baseball seasons. Total years spent in the minors**
frst_yrmj	**First year in the major leagues**
lst_yrmj	**Last year in the major leagues**
active	**Active. Is the player still playing major league baseball as of 2000**
tot_seas	**Total seasons spent in major league baseball**
pos	**Position played on the field**
g	**Games. Total number of games played in observation year**
ab	**At bats. Total number of at bats during the observation year**
r	**Runs. Total runs scored during the observation year**
h	**Hits. Total number of hits during the observation year**

doubles	**Total number of doubles hit during the observation year**
triples	**Total number of triples hit during the observation year**
hr	**Home runs. Total number of home runs hit during the observation year**
rbi	**Runs batted in for each observation year**
tb	**Total bases**
bb	**Base on balls. Walks**
so	**Strike outs**
sb	**Sacrifice bunts**
cs	**Caught stealing.**
obp	**On base percentage**
slg	**Slugging**
avg	**Batting average**
sf	**Sacrifice flies**
sh	**Sacrifice hits**
hbp	**Hit by pitch**
ibb	**Intentional base on balls. Intentional walks**
gdp	**Grounded into a double play**
tpa	**Total plate appearances**
xbh	**Extra base hits**
sb_perc	**Stolen base percentage**
ops	**On base + slugging**