AN ANALYSIS OF THE NBA DRAFT: ARE TEAMS DRAFTING BETTER AND DOES COLLEGE EXPERIENCE TRULY MATTER

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Kevin Joseph Wolfe

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Statistics

November 2022

Fargo, North Dakota

North Dakota State University
Graduate School

Title

AN ANALYSIS OF THE NBA DRAFT: ARE TEAMS DRAFTING BETTER AND DOES COLLEGE EXPERIENCE TRULY MATTER

By

Kevin Joseph Wolfe

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Rhonda Magel
Chair

Megan Orr

Oudom Hean

Approved:

November 18, 2022 — Rhonda Magel

Date — Department Chair

ABSTRACT

This thesis attempts to answer two questions. Are NBA organizations doing a reasonable job at drafting players and getting better at the process, and does college experience play a significant role in a player's performance during their early NBA career (first 3 seasons).

In regard to these two questions, we determined through our research that NBA organizations are not showing any significant improvements in their ability to draft the best available players, this is surprising given the increase in available scouting data teams currently have access to. We suspected however that this lack of drafting improvements may be related to players entering the NBA with less college experience. However, after we determined that college experience does not appear to play a large role in a player's early career NBA performance, we determined that experience does not appear to be the reason why teams aren't doing a better job of drafting.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

In the National Basketball Association – (NBA) there is an event that takes place before the start of every new season. This event is known as the "NBA Draft" and has long been used as a way for the league to attempt to keep competition as fair as possible, this is accomplished by giving the first draft selections to the teams who had won the fewest number of games during the previous season. The Draft consists of every current NBA team taking turns selecting from the best available, draft eligible players. Throughout most of the draft's history, being draft eligible referred to any player who had completed high school. This changed however starting in 2006 when it was determined that a player now had to turn nineteen within the calendar year of the draft and additionally needed to be one year removed from high school if attending in the United States. (Garnes, 2019)

Since its inception in 1947, the NBA draft has been seen by many as a key factor in determining which NBA franchises will flourish and which will falter in the coming years. As time has passed and we have approached the modern era of the NBA, the level of detail that is involved in the process of scouting and analyzing these young basketball players has increased at an extremely fast rate. In this thesis we would like to test the following: 1. Is a player's draft pick number correlated with the first 3 years of a player's performance in the NBA; and 2. Has this correlation increased significantly through the years. (Berger & Daumann, 2021) (Sailofsky, 2018)

We will measure the first 3 years of a player's performance in the NBA through the usage of the advanced statistic Win Shares. Win shares is calculated through a combination of multiple simple statistics. A few of these key simple statistics are Points Per Game, Assists, and Rebounds. (*NBA win shares | basketball-reference.com 2000*) (*What is a rebound? definition &*

meaning on sportslingo.com 2022) (*What is an assist in basketball?* 2021) (*Scoring system in basketball: how point scoring works*) We will calculate the correlation coefficients between draft pick number and the first 3 years of a player's performance for the years (1981-2018). We will then perform a test to see if there is increasing correlation through the years between a player's draft pick and the players first 3 seasons (Namely, is the accuracy in selecting players through the draft improving?).

Players in the draft are now much younger on average than they were just ten years ago with most top players only participating in one or two years of collegiate basketball. In this research, we would also like to determine if there is any association between the number of years a player has played basketball in college and the first three years of his NBA career performance. Is it better for a player to play college basketball for a while before going into the NBA? If so, for how long?

It should be noted that international players who are selected through the draft lack college basketball experience. Additionally, many international players usually don't join the NBA until multiple years after the time of their original selection. This is often due to previous obligations that they have made with their present teams. As a result of these two facts, international players are not being included in the second part of the research. (Mathewson, 2018)

In Chapter 2 we will go over research related to the topics covered in this thesis. We will then compare our thesis with these other topics and discuss how these other papers can be interpreted in the context of our own research. In Chapter 3 we will be detailing and defining Points, Rebounds, Assists, and Win Shares as well as other statistics, additionally in Chapter 3 we will describe what tests are being performed and what methods are being used for these tests.

In Chapter 4 we will show what the numeric results of these tests are. Then finally, in Chapter 5 we will reach our conclusion and will begin detailing what future research these results could lead us to.

CHAPTER 2 REVIEW OF LITERATURE

In chapter two we will be exploring prior works that have covered topics related to early draft entry and the importance of college experience, as well as studies related to how data for the NBA draft is gathered.

Relevant Research Related to Early Entry Into the NBA Draft.

(Groothuis, Hill, Perri 2007) is an article that explores possible reasons why teams may have begun to draft younger players in the NBA draft. The article makes note of the decrease in the number of college seniors who were drafted in the first round in the year 1997 compared to 2004. The article additionally attempts to delve into the relationship between the structuring of NBA rookie contracts and the trend towards younger players. Overall, this article shows some overlap with our study regarding the acknowledgment of a decrease in age for players being drafted. Differences however, can be seen in the purpose of our studies. While our study is more related to the effects of early entry into the NBA draft on a player's early career performance, their study is more concerned with one of the possible triggers that may have helped usher in the present era of the NBA draft, that trigger being rookie contract structures.

(Arel & Tomas, 2011) looks closely at the decision-making process a player goes through when determining whether they should forgo college experience in order to enter the draft early. This article takes a businesslike approach when viewing the process of leaving college early, the article equates entering the draft early to selling your remaining college experience to an NBA team. This article is very similar to ours in regard to the fact that both of our studies are evaluating the effects of entering the draft early compared to finishing college. Our studies differ however regarding how we go about evaluating success for the players being studied. Our study evaluates whether a player should go pro early or not based on how that effects their early performance in the NBA, while their study evaluates the decision from a monetary perspective. Overall, both factors can be seen as important and which study bears more weight to an individual would likely depend on what they equate as success, performance or income.

(Sugai, 2010) formed a thesis that has many overlaps with ours. The thesis asks the same question that is asked in this thesis however, like the prior article above, the thesis covers the decision-making process mostly from a financial perspective. However, Sugai's thesis also frames the decision to go pro from the familiar life perspective of schooling (staying in college) versus on-the-job learning (entering the pros early). In addition to this perspective the thesis also points out the fact that some college coaches aren't training players to perform in the NBA, but to perform well in college. To explain further, the NBA and college basketball do not share the same style of play. College basketball is usually allowed to be played in a more physical manner and coaches often like running zone based defensive schemes, a zone based defensive scheme is one where each defender guards a portion of the floor instead of a particular offensive player. These schemes are more effective at the college level since players generally aren't as good at shooting as their more skilled NBA counterparts. This results in many players not acquiring the skills needed to effectively guard one on one while playing in college, and this is a skill that is required of players in the NBA. The NBA encourages fast play and limits the amount of contact defenders can make in order to make games as exciting to watch as possible, additionally the high level of shooting that was mentioned prior basically forces teams to defend man to man. Overall, the topics covered by this thesis will be very helpful to consider when evaluating our own data results.

Relevant Research Related to Increases in Draft Accuracy.

(Sailofsky, 2018) formed a thesis related to determining where teams error when drafting players. The thesis analyses players college statistics, and other pre-draft factors to determine why NBA teams aren't seeing an increase in draft accuracy (Defined by our thesis in chapter 1). This thesis also attempts to model a player's career based on different factors using regression. Overall, this thesis is formed on the basis that teams are not drafting more accurately over time despite the presence of increased data. Our thesis seeks to confirm the assumptions that this thesis is performed under, and as a result, this thesis is a logical next step to our thesis if it is indeed confirmed that draft accuracy is not increasing.

(Ichniowski & Preston, 2017) looks at the well televised and overwhelmingly popular march madness national college basketball tournament. The study seeks to determine whether the popularity of this tournament might result in players being overvalued in the draft due to their performances in the tournament. The study concludes that tournament performance is not overvalued and is possibly slightly undervalued. Scouts appear to be trying to ignore the small sample size of play that can be seen during the ncaa tournament, however perhaps more weight could be given to the idea of players playing their best on the national stage under large amounts of pressure since this could possibly allow scouts to observe how clutch a particular player is. Overall, this could perhaps explain a small amount of the large variance that is present in the NBA draft, however more research is likely needed.

(Coates, Dennis, and Oguntimein, 2008) analyzed the college statistics of players drafted between 1987 & 1989 and found that players who scored a large amount of point in college were more likely to be drafted, however their level of college scoring did not consistently translate over to the professional level. However, the study also found that statistics like assists, and rebounds did have a higher level of correlation in regard to predicting NBA performance. This could perhaps be interpreted as scoring being overvalued in the NBA draft due to how hard it is to determine whether a player in college who scores a high volume of points will score at a high volume in the NBA. This could perhaps be another factor that accounts for a small portion of the variance seen in the draft. However, it should be noted that the sample used only contained three NBA drafts and the accuracy of the study could be doubted as a result.

Relevant Research Related to Increases in Draft Accuracy & Early Draft Entry.

(Teramoto, 2018) analyses the effectiveness of the NBA draft combine in predicting a player's performance in the NBA. The 2010-2015 draft combines were analyzed and broken into 3 main categories that were being tested for in general, length-size, power-quickness, and upper-body strength. Principal component analysis was performed on these factors to determine which types of tests correlated most strongly with the on-court performance statistics of Win Shares, Box Plus/Minus, and Value Over Replacement Player. The results of this thesis showed that length and size was the most powerful predictor and upper body strength was also an important predictor. However, the power-quickness factor was not significant. The results of this thesis provides interesting context about the results of our thesis. For example, length and size are both factors that will continue to improve as a player matures during their college years, and the same goes for upper body strength. The fact that these two factors can predict performance implies that at the very least the maturing that takes place during the college years is significant in regard to a player's early career performance. Additionally, the fact that the draft combine has been shown to be predictive of performance in the NBA brings up questions regarding why more detailed draft combines hasn't led to more accurate drafts.

CHAPTER 3 RESEARCH METHODOLOGY

Recall that in the Introduction it was mentioned that the research in this thesis will be divided into two topics. The first research topic relates to how well teams are doing at drafting players for the NBA and whether they are getting better at drafting as time goes on. This research topic is divided into the following two parts: 1. Is a player's draft pick number correlated with the first 3 years of a player's performance in the NBA; and 2. Has this correlation increased significantly through the years. (Berger & Daumann, 2021) (Sailofsky, 2018)

For this first research topic, data had to be gathered. Basketball reference keeps a detailed list of prior NBA drafts as well as each drafted player's career and seasonal statistics. (*Basketball-reference.com - basketball statistics and history* 2000) All relevant data from basketball reference was then recorded on a self-made Excel document and all players who had failed to play either 50 career games or 250 career minutes were filtered out of the dataset. In this thesis, the 1981-2018 drafts will be used. It should be noted that prior to 1985, the NBA draft had ten rounds but ultimately decided to do away with most of the later rounds due to the rarity of players drafted in those rounds having any success. Additionally, these later rounds of the draft were being treated as a joke by many of the NBA's teams. For example, it wasn't uncommon for teams to draft famous celebrities, close friends, or even newborns in the later rounds of the draft. The decrease in the draft's length started in the 1985 draft when the number of rounds was shortened to seven. The seven round format was then used until 1988 when the draft was again shortened to three rounds before being shortened to its present two round format starting with the 1989 draft. Over the years, the number of picks in each round of the draft has varied based on how many teams are in the NBA. During the 1981 draft, only twenty-three picks occurred in each round, then in 1982 the round length's extended to twenty-four due to the addition of a new team in the NBA. The draft then stayed at twenty-four picks per round until 1988, when a twenty-fifth team joined the NBA. Next, another two teams joined the NBA in 1989 increasing the number of picks to twenty-seven. After a few more years once again two more teams joined the NBA in 1995 taking the total to twenty-nine before finally, in the year 2004 one additional team joined the league which brings us to our modern-day total of thirty total picks per round of the draft. For all these drafts mentioned above, players outside the first two rounds were not included in the draft analysis. (*The Long Weird History of the NBA Draft* 2013)

For each player, we calculated multiple different statistics based on their first three years of playing in the NBA. These statistics were then used to create multiple different performance measures with which we could evaluate each players production. One of these statistics was the advanced statistic Win Shares. To quote basketball reference, “Win Shares is a statistic which attempts to divvy up credit for team success to the individuals on the team”. (*NBA win shares | basketball-reference.com 2000*) Win Shares is calculated using a combination of other statistics that are recorded over the course of a basketball game. Win Shares is calculated using points scored, rebounds, assists, turnovers, and a multitude of other statistics for each individual player during a basketball game.

Three additional measures of a player’s performance were calculated for each player, these measurements relate to a player’s average amount of points scored per game, average amount of assists per game, and average amount of rebounds per game.

Points is defined as the number of points an individual player is responsible for scoring on average during one of their team’s games. Points are scored through the process of putting the basketball through the basketball hoop. Note that this is often seen as the most important statistic due to the fact that it in theory should most directly relate to winning a basketball game since the winning team is decided once the clock hits 0 based on which team has scored more points. A rebound is scored when a player grabs the basketball after another player has attempted to score but failed to put the ball into the basketball hoop. This statistic is often regarded as being important due to its correlation to a team's ability to score the ball into the basket, for example, if a player gets a rebound after one of their teammates fails to score, then their team acquires an additional chance to score. An assist occurs when a player passes the ball to one of their teammates who then immediately scores the basket. Assists are seen as important due to how closely they relate to scoring. Note, in basketball it is usually considered very difficult to score by yourself without being set up "having the ball passed to you" by a teammate, as such, assists are considered nearly as important as points in the game of basketball. (*Calculating win shares | college basketball at sports-reference.com* 2000) (*Scoring system in basketball: how point scoring works*) (*What is a rebound? definition & meaning on sportslingo.com* 2022) (Warner, 2021)

Win Shares can be further broken down into Offensive Win Shares and Defensive Win Shares. Offensive Win Shares only uses offensive statistics when calculating its score and likewise Defensive Win shares only uses defensive statistics. Offensive and Defensive Win Shares when added together, will be equivalent to total Win Shares earned during a game. Also, it should be noted that the sum of a team's total Win Shares is usually very close to that team's total wins over the course of a season, as a result of this fact we can conclude that Win Shares is a very good estimator of a player's performance in the NBA. (*NBA win shares | basketball-reference.com 2000*) (*Calculating win shares | college basketball at sports-reference.com* 2000)

For our Offensive Win Shares measure we will be calculating the total amount of Offensive Win Shares that an individual player accounts for during their first three seasons played. We will also be calculating our Defensive Win Shares measure using the same method used for Offensive Win Shares. After including these two performance measures we now have six total. The six performance measures are: Total Win Shares during first three seasons, Offensive Win Shares during first three seasons, Defensive Win Shares during first three seasons, average points per game, average rebounds per game, and average assists per game.

We wanted to see if for each year, a player's draft pick and a player's measure of performance during their first three years in the NBA were correlated. Since we are considering six measures of performance, six correlation coefficients will be calculated for each year. Before we could begin calculating correlation coefficients, we had to first decide how to rank draft picks. We decided that each draft pick should be reorganized into categories, where picks 1-5 in each draft would be considered draft rank 1, picks 6-10 would be considered draft rank 2, picks 11-15 would be considered draft rank 3…. with the final picks 56-60 being considered draft rank 12. Our reason for organizing different pick ranges into groups is due to the fact that players drafted one pick before or after each other are going to have less obvious differences when it comes to expectations. However, by grouping every five draft picks into ranks, we can examine what should be a more obvious difference in expected performance between groups.

Now that picks had been reorganized into ranks, we began calculating our correlation coefficients. Six different sample correlation coefficients were calculated for each year of the draft with 38 draft years considered. For our coefficients we used Spearman correlation because it is the most appropriate method to use with our rank data. To get our different coefficient scores, a correlation coefficient was calculated between each of our six measures of a player's performance during their first three seasons and that player's draft pick rank, this process was done for all 38 draft years. When calculating these coefficients, we would expect negative correlation coefficients between each of the six measures of performance and draft picks. This is due to the fact that a higher measure for performance indicates that a player is performing well and a lower draft pick rank means that the player was selected earlier in the draft. As a result, the optimal correlation coefficient score would be -1 which would indicate that a player's performance would align perfectly to their draft pick rank.

It should be noted that correlation coefficients can range anywhere from -1 to 1, where the value represents the relationship between two variables. If the correlation score is -1 for example, that would imply that a decrease in one variable would directly relate to an increase in another variable, and likewise a correlation score of 1 would imply that an increase in one variable directly relates to an increase of another variable. Finally, a correlation of 0 would imply that the two variables have no effect on one another and can be seen as having no relationship. (Spearman Correlation - an Overview | ScienceDirect Topics)

Six different Spearman correlation coefficients were calculated between draft rank and each of our six performance measures. These coefficients were calculated using our six measures for a player's performance during their first three seasons and that same player's draft pick rank. We did this for each of the 38 drafts observed in this thesis. In each draft, correlation coefficients were calculated based on the rank draft pick and performance measure for all players drafted in the first 2 rounds.

If franchises were doing a reasonable job in drafting players, we would expect a low rank draft pick to be associated with a high rank performance measure since the higher the performance measure, the better the performance of a player. Hence, the closer a correlation coefficient is to -1, the stronger the association between rank draft pick and the performance measure with -1 being the optimal correlation.

For the first part of our first research topic, we wanted to determine if organizations are doing a reasonable job at drafting players. For each performance measure we tested whether the average correlation coefficient between rank draft pick and that performance measurement is less than -.5. In other words, after finding our mean correlation score for each of these performance measures with rank draft pick across the 38 drafts, we then went on to perform t-tests to determine if the true mean correlation was less than -.5.

We next moved on to examine the second part of our first research project. For this, we wanted to determine if there was an increase in the associations between draft pick selection and each of our six measures of NBA performance. In our case, this would mean that correlation coefficients are getting closer to -1 as we move from year to year. As such, it can be said that for this hypothesis we are testing to determine if a decreasing trend is occurring among the correlation coefficients. We performed a Cox-Stuart test for trend for each of the sets of correlations between rank draft pick and one of the performance measures to test for a decreasing trend. This test was then conducted for each performance measure. *(How To: Binomial Test Sign Test McNemars Test Cox & Stuart, 2003)*

For our second research topic, we wanted to examine whether there was a relationship between the number of years a player played college basketball and the player's first three years of performance in the NBA. Note, that this hypothesis was formed due to knowledge regarding the fact that players have been spending less years playing college basketball then they did in the past, this was confirmed through an analysis of the average years of college basketball played by drafted players over time. In this case the Cox Stuart test for trend was once again used to confirm this fact.

We next wanted to determine if the amount of time a player spent playing college basketball directly related to their early career performance in the NBA. Three regression models were created with each of the three models using the same independent variables, but with three different dependent variables, Total Win Shares, Offensive Win Shares, and Defensive Win Shares. The first model created used total Win Shares accumulated during a players first 3 seasons to determine whether playing in college was truly important. The second and third model used total Offensive Win Shares and total Defensive Win Shares during a player's first 3 seasons in the NBA, respectively. For these models, Y is equal to a Player's Win Shares in the NBA for their first 3 seasons and the baseline effect is that a drafted player has 0 years of college basketball experience. Additionally, the indicator variables in this model are the variables 1 year of college basketball experience, 2 years of college basketball experience, 3 years of college basketball experience, and finally 4+ years of college basketball experience. Each of the 4 variables consist of the values 1 and 0 where if the amount of experience playing college basketball is the same as the amount of experience in the variables name, then the variable would equal 1, and otherwise 0. Additionally, the baseline for this model was representative of a player who never attended college and had instead gone straight into the NBA. This left us with the model shown below.

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \varepsilon \quad (1)$$

Y = One of 3 (Total Win Shares, Offensive Win Shares, Defensive Win Shares)

X1 = 1 if one year of college basketball was played and 0 otherwise.

X2 = 1 if two years of college basketball was played and 0 otherwise.

X3 = 1 if three years of college basketball was played and 0 otherwise.

X4 = 1 if four or more years of college basketball was played and 0 otherwise.

For these three models, multiple different versions of our dataset were used. The first version contained the first 2 rounds of each draft only and excluded no outliers. The second version also contained every player drafted in the first 2 rounds but also excluded the top 3 leaders in Win shares for each college basketball experience level (0,1,2,3,4+ years of college basketball played). The third version contained only players selected in the lottery (first 14 picks). Finally, the fourth version contained only lottery players but excluded the top 3 outliers from each experience level. These four datasets were then used on our 3 models.

Overtime, it was eventually determined that a second subset of these 3 models should be created as well. This subset of models would be very similar to our first set of models but would also contain a couple key differences. The first of these differences is that in our second subset of models, we do not count players with 0 years of experience and instead we use 1 year of college basketball played as our baseline. Therefore $\boldsymbol{\beta 1 X1}$ was dropped from the model. The other key difference for this second set of models is that the variable 4+ years of college basketball played is changed to being only 4 years of college basketball played.

Overall, there were two main reasons that this second subset of models was made. The first of these reasons was to create models that would compare 1 year of college experience instead of 0. This was done because of the rarity of players with 0 years of college experience. Generally, you want your baseline to have an adequate amount of datapoints so that you can have more accurate analysis. The second reason was related to wanting to exclude the rare and unproductive players that played more than 4 years of college basketball. Overall players who played more than 4 years of college basketball usually did so because of injury problems, as a result these players performed much worse than 4-year players and were dragging down the 4+ years average in the first model. In order to fix this problem players who played more than 4 years were excluded from the second model in order to improve the accuracy of the 4 years of college basketball group.

Finally, it should be noted as well that international players who are selected through the draft lack college basketball experience. As a result, many international players usually don't join the NBA until multiple years after the time of their original selection. This is often due to previous obligations that they have made with their present teams. As a result of these two facts, international players are not being included in these models. (Mathewson, 2018)

CHAPTER 4 ANALYSIS OF DATA

In this chapter we will be showing the statistical results of the multiple tests that were performed as well as some brief analysis related to each of the outputs we display. During this chapter all the data shown will be in the order of its calculation, however before talking about our individual tests and hypotheses we first need to go over how each draft had its correlation score calculated and what those correlation scores are.

Recall from Chapter 3 that we used Spearman Correlation in order to calculate six different correlation scores pertaining to the relationship between draft pick rank and each of the performance measures (Offensive Win Shares, Defensive Win Shares, Total Win Shares, Points, Assists, and Rebounds) for each of our 38 drafts (1981- 2018). These 6 sets of correlations and their graphs can be seen below in Tables & Figures 1 through 6.

Table 1. Yearly Correlation OWS With Draft Pick

YEAR	SPEARMAN CORRELATION	YEAR	SPEARMAN CORRELATION
1981	-.71	2000	-.24
1982	-.56	2001	-.40
1983	-.49	2002	-.19
1984	-.47	2003	-.34
1985	-.35	2004	-.36
1986	-.26	2005	-.45
1987	-.36	2006	-.23
1988	-.28	2007	-.37
1989	-.24	2008	-.64
1990	-.24	2009	-.35
1991	-.31	2010	-.15
1992	-.59	2011	-.50
1993	-.31	2012	-.42
1994	-.33	2013	-.47
1995	-.50	2014	-.27
1996	-.56	2015	-.33
1997	-.44	2016	-.32
1998	-.42	2017	-.21
1999	-.38	2018	-.44

Figure 1. Absolute Value of OWS Correlation by Year

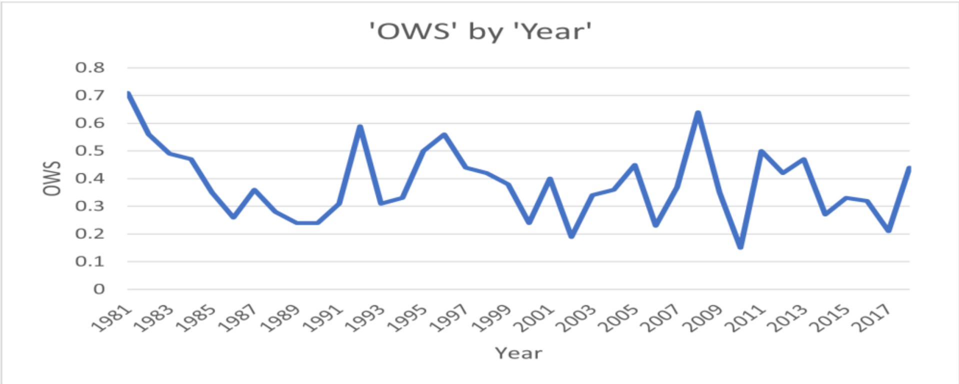


Table 2. Yearly Correlation DWS With Draft Pick

YEAR	SPEARMAN CORRELATION	YEAR	SPEARMAN CORRELATION
1981	-.68	2000	-.76
1982	-.67	2001	-.68
1983	-.63	2002	-.53
1984	-.66	2003	-.61
1985	-.56	2004	-.73
1986	-.44	2005	-.50
1987	-.69	2006	-.64
1988	-.63	2007	-.64
1989	-.60	2008	-.69
1990	-.58	2009	-.61
1991	-.72	2010	-.78
1992	-.78	2011	-.66
1993	-.69	2012	-.66
1994	-.67	2013	-.74
1995	-.64	2014	-.62
1996	-.68	2015	-.76
1997	-.62	2016	-.67
1998	-.71	2017	-.69
1999	-.60	2018	-.62

Figure 2. Absolute Value of DWS Correlation by Year

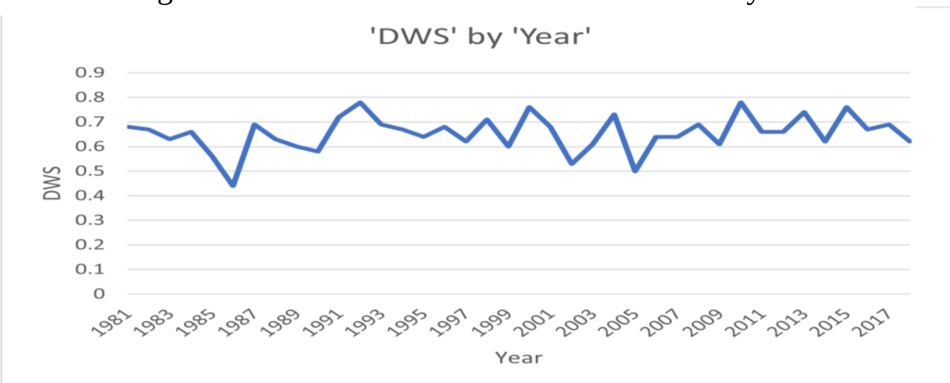


Table 3. Yearly Correlation TWS With Draft Pick

YEAR	SPEARMAN CORRELATION	YEAR	SPEARMAN CORRELATION
1981	-.78	2000	-.59
1982	-.69	2001	-.55
1983	-.69	2002	-.37
1984	-.57	2003	-.51
1985	-.49	2004	-.63
1986	-.32	2005	-.51
1987	-.50	2006	-.44
1988	-.53	2007	-.62
1989	-.46	2008	-.73
1990	-.47	2009	-.45
1991	-.58	2010	-.66
1992	-.74	2011	-.60
1993	-.56	2012	-.58
1994	-.57	2013	-.67
1995	-.61	2014	-.52
1996	-.69	2015	-.58
1997	-.57	2016	-.60
1998	-.55	2017	-.35
1999	-.49	2018	-.56

Figure 3. Absolute Value of TWS Correlation by Year

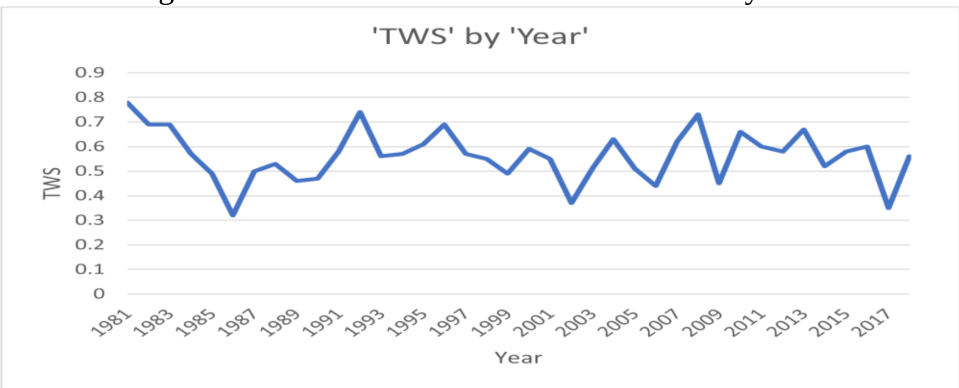


Table 4. Yearly Correlation of Avg. PPG With Draft Pick

YEAR	SPEARMAN CORRELATION	YEAR	SPEARMAN CORRELATION
1981	-.74	2000	-.68
1982	-.64	2001	-.62
1983	-.73	2002	-.51
1984	-.73	2003	-.61
1985	-.56	2004	-.77
1986	-.39	2005	-.47
1987	-.70	2006	-.65
1988	-.60	2007	-.68
1989	-.60	2008	-.77
1990	-.51	2009	-.62
1991	-.57	2010	-.75
1992	-.77	2011	-.65
1993	-.70	2012	-.67
1994	-.66	2013	-.73
1995	-.69	2014	-.60
1996	-.73	2015	-.74
1997	-.53	2016	-.64
1998	-.70	2017	-.61
1999	-.62	2018	-.72

Figure 4. Absolute Value of Points Per Game Correlation by Year

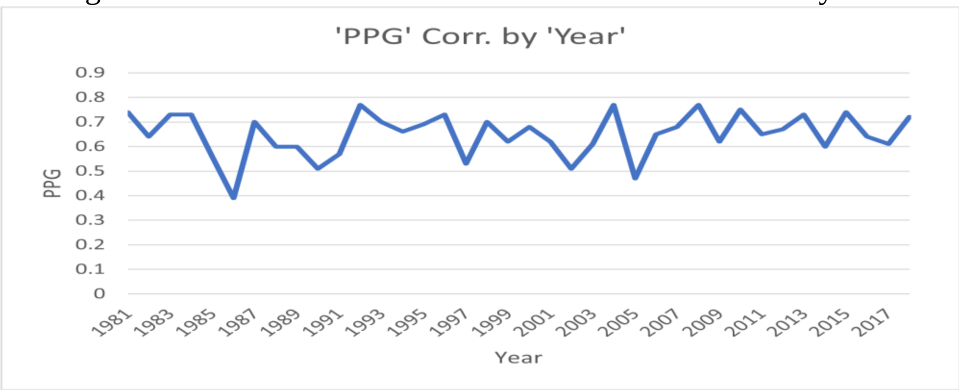


Table 5. Yearly Correlation of Avg. APG With Draft Pick

YEAR	SPEARMAN CORRELATION	YEAR	SPEARMAN CORRELATION
1981	-.63	2000	-.52
1982	-.53	2001	-.45
1983	-.53	2002	-.48
1984	-.55	2003	-.55
1985	-.49	2004	-.64
1986	-.21	2005	-.32
1987	-.57	2006	-.49
1988	-.23	2007	-.56
1989	-.45	2008	-.63
1990	-.34	2009	-.56
1991	-.43	2010	-.68
1992	-.60	2011	-.47
1993	-.62	2012	-.51
1994	-.53	2013	-.49
1995	-.47	2014	-.52
1996	-.48	2015	-.72
1997	-.42	2016	-.54
1998	-.53	2017	-.53
1999	-.65	2018	-.55

Figure 5. Absolute Value of Assists Per Game Correlation by Year

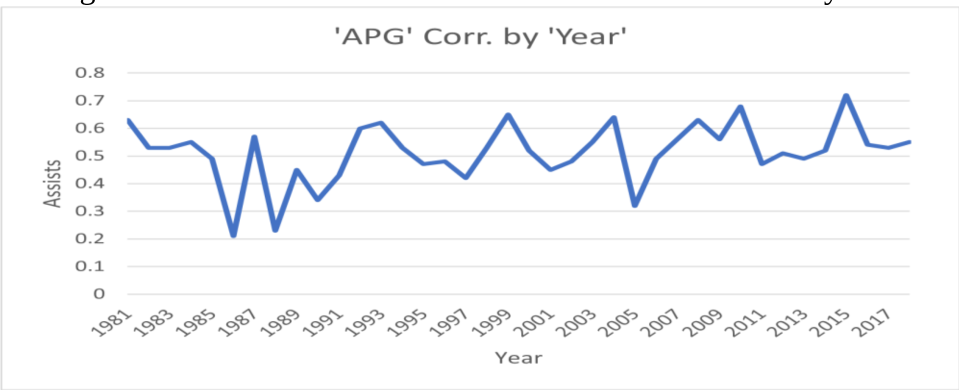
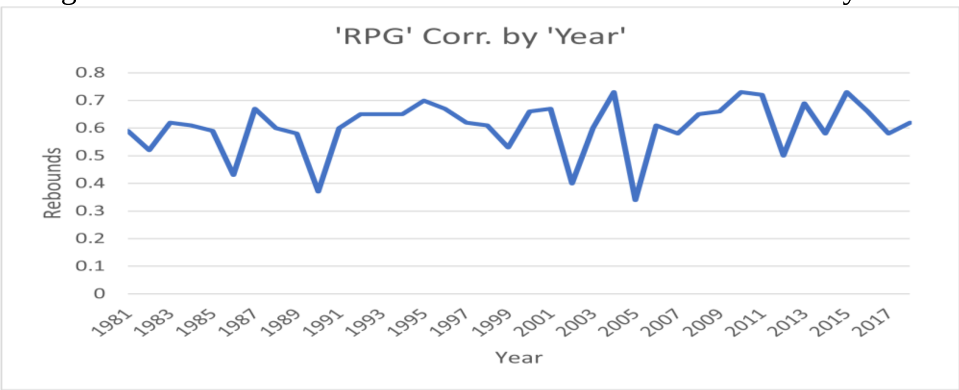


Table 6. Yearly Correlation of Avg. RPG with Draft Pick

YEAR	SPEARMAN CORRELATION	YEAR	SPEARMAN CORRELATION
1981	-.59	2000	-.66
1982	-.52	2001	-.67
1983	-.62	2002	-.40
1984	-.61	2003	-.60
1985	-.59	2004	-.73
1986	-.43	2005	-.34
1987	-.67	2006	-.61
1988	-.60	2007	-.58
1989	-.58	2008	-.65
1990	-.37	2009	-.66
1991	-.60	2010	-.73
1992	-.65	2011	-.72
1993	-.65	2012	-.50
1994	-.65	2013	-.69
1995	-.70	2014	-.58
1996	-.67	2015	-.73
1997	-.62	2016	-.66
1998	-.61	2017	-.58
1999	-.53	2018	-.62

Figure 6. Absolute Value of Rebounds Per Game Correlation by Year



Once all the correlation coefficients had been calculated we could begin answering our first question related to whether or not franchises are doing a good job of drafting. As stated in the previous chapter, a correlation of -.50 will be used. We want to see if the average correlation coefficient between any of the six measures of performance and draft pick ranks is significantly lower than -.50, as this would indicate that franchises are doing a good job of drafting with respect to that measure of performance. We tested the average of the six correlation coefficients using alpha equal to 0.05.

We first tested to see if Offensive Win Shares had an average correlation with draft pick rank that was significantly less than -0.50. The null and alternative hypothesis can be seen below.

Ho: OWS = -.50

Ha: OWS < -.50

Figure 7. t-Test to Determine if Correlation Mean of OWS <-.5

t-Test: Mean OWS < -.50						
Mean	-0.38105					
Variance	0.016664					
Observations	38					
Hypothesized Mean Difference	0					
df	37					
t Stat	5.68016					
P-Value	0.999999					
t Critical one-tail	-1.68709					

We see from Figure 7 that for the null hypothesis Ho: OWS = -.50 we have a P-Value of Approximately 1. Since this P-Value is greater than .05 we fail to reject our null hypothesis and conclude that we have insufficient evidence to suggest that the true average correlation coefficient of Offensive Win Shares with draft pick is less than -.50. As a result, there is no evidence to indicate that franchises are doing a reasonable job at drafting with respect to the performance measure for Offensive Win Shares.

Next, we tested to see if Defensive Win Shares had an average correlation with draft pick rank that was significantly less than -0.50. The null and alternative hypothesis can be seen below.

Ho: DWS = -.50

Ha: DWS < -.50

Figure 8. t-Test to Determine if Correlation Mean of DWS <-.5

t-Test: Mean DWS<-.50	
	DWS
Mean	-0.653684211
Variance	0.0053266
Observations	38
Hypothesized Mean Difference	0
df	37
t Stat	-12.98063669
P-Value	1.20061E-15
t Critical one-tail	-1.68709362

Observe in Figure 8 that for the null hypothesis Ho: DWS = -.50 we have a P-Value that is approximately equal to 0. Since this P-Value is less than .05 we reject our null hypothesis and conclude that we have significant evidence that the true mean correlation of Defensive Win Shares is less than -.5. This indicates that franchises are doing a reasonable job at drafting with respect to the performance measure for Defensive Win Shares.

For our third t-test we tested to see if Total Win Shares had an average correlation with draft pick rank that was significantly less than -0.50. The null and alternative hypothesis can be seen below.

Ho: TWS = -.50

Ha: TWS < -.50

Figure 9. t-Test to Determine if Correlation Mean of TWS <-.5

t-Test: Mean TWS <-.50	
	TWS
Mean	-0.56263
Variance	0.010874
Observations	38
Hypothesized Mean Difference	0
df	37
t Stat	-3.70247
P-Value	0.000346
t Critical one-tail	-1.68709

Notice in Figure 9 that for the null hypothesis Ho: TWS = -.50 we have a P-Value of .0003. Since this P-Value is less than .05, we reject our null hypothesis and conclude that we have significant evidence that the true mean correlation of Total Win Shares is less than -.5. This indicates that franchises are doing a reasonable job at drafting with respect to the performance measure for Total Win Shares.

We then performed a fourth t-test to see if Points Per Game had an average correlation with draft pick rank that was significantly less than -0.50. The null and alternative hypothesis can be seen below.

Ho: PPG = -.50

Ha: PPG < -.50

Figure 10. t-Test to Determine if Correlation Mean of PPG <-.5

t-Test: Mean PPG < -.50	
	PPG
Mean	-0.64895
Variance	0.007918
Observations	38
Hypothesized Mean Difference	0
df	37
t Stat	-10.3187
P-Value	9.69E-13
t Critical one-tail	-1.68709

Observe in Figure 10 that for the null hypothesis Ho: PPG = -.50 we have a P-Value that is approximately equal to 0. Since this P-Value is less than .05 we reject our null hypothesis and conclude that we have significant evidence that the true mean correlation of Points Per Game is less than -.5. This indicates that franchises are doing a reasonable job at drafting with respect to the performance measure for average Points Per Game.

Next, for our fifth t-test we wanted to see if Assists Per Game had an average correlation with draft pick rank that was significantly less than -0.50. The null and alternative hypothesis can be seen below.

Ho: APG = -.50

Ha: APG < -.50

Figure 11. t-Test to Determine if Correlation Mean of APG <-.5

t-Test: Mean APG < -.50	
	APG
Mean	-0.51237
Variance	0.011786
Observations	38
Hypothesized Mean Differen(	0
df	37
t Stat	-0.7023
P-Value	0.243445
t Critical one-tail	-1.68709

We see from Figure 11 that for the null hypothesis Ho: APG = -.50 we have a P-Value of .243. Since this P-Value is greater than .05 we fail to reject our null hypothesis and conclude that we have insufficient evidence to suggest that the true average correlation coefficient of Assists Per Game with draft rank is less than -.50. As a result, there is no evidence to indicate that franchises are doing a reasonable job at drafting with respect to the performance measure for average Assists Per Game.

Finally, for our sixth and final t-test we wanted to see if Rebounds Per Game had an average correlation with draft pick rank that was significantly less than -0.50. The null and alternative hypothesis can be seen below.

Ho: RPG = -.50

Ha: RPG < -.50

Figure 12. t-Test to Determine if Correlation Mean of RPG <-.5

t-Test: Mean RPG < -.50	
	RPG
Mean	-0.60447
Variance	0.008971
Observations	38
Hypothesized Mean Difference	0
df	37
t Stat	-6.79939
P-Value	2.63E-08
t Critical one-tail	-1.68709

Notice in Figure 12 that for the null hypothesis Ho: RPG = -.50 we have a P-Value of Approximately 0. Since this P-Value is less than .05, we reject our null hypothesis and conclude that we have significant evidence that the true mean correlation of Rebounds Per Game is less than -.5. This indicates that franchises are doing a reasonable job at drafting with respect to the performance measure for average Rebounds Per Game.

Overall, we found evidence indicating that franchises are doing a reasonable job of drafting with respect to the performance measures for Defensive Win Shares, Total Win Shares, average Points Per Game, and average Rebounds Per Game. We however did not find evidence indicating that franchises are doing a reasonable job of drafting players with respect to the performance measures for Offensive Win Shares, and average Assists Per Game. It should be noted however, that Total Win Shares is considered the best overall performance measure of a player, and there was evidence to indicate that franchises are doing a reasonable job of drafting players with respect to this key performance measure.

Next, we will examine the second part of our first research topic, namely, are franchises getting better at drafting players, or as it could also be put, have the correlations of the six performance measures with respect to draft picks decreased over time. (became closer to -1)

To test this, we will perform the Cox-Stuart test for trend over time using the yearly correlating coefficients calculated with respect to each of our six performance measures.

Our first set of hypotheses is

Ho: No trend for OWS coefficient

Ha: There is a decreasing trend (draft accuracy is improving)

The results are shown in Figure 13 below.

Figure 13. Cox-Stuart Test for Trend for OWS

OWS test for trend						
n	38					
m	19					
nneg	8					
npos	11					
P-Value	0.323802948					
P-Value for Improved Drafting Hypothesis	0.676197052					

From Figure 13 we can see that we have a P-Value of .676 for our hypothesis that draft accuracy is improving Offensive Win Shares. This value is greater than .05. As a result, we can conclude that there is insignificant evidence that draft accuracy is improving with respect to Offensive Win Shares.

Our second set of hypotheses is

Ho: No trend for DWS coefficient

Ha: There is a decreasing trend (draft accuracy is improving)

The results are shown in Figure 14 below.

Figure 14. Cox-Stuart Test for Trend for DWS

DWS test for trend	
n	38
m	19
nneg	12
npos	7
P-Value	0.179642
P-Value for Improved Drafting Hypothesis	0.179642

From Figure 14 we can see that the P-Value of .179 for our hypothesis that draft accuracy is improving regarding Defensive Win Shares is greater than .05. As a result, we can conclude that there is insignificant evidence that draft accuracy is improving with respect to Defensive Win Shares.

Our third set of hypotheses is

Ho: No trend for TWS coefficient

Ha: There is a decreasing trend (draft accuracy is improving)

The results are shown in Figure 15 below.

Figure 15. Cox-Stuart Test for Trend for TWS

TWS test for trend	
n	38
m	19
nneg	9
npos	10
P-Value	0.5
P-Value for Improved Drafting Hypothesis	0.5

From Figure 15 we can see that the P-Value of .5 for our hypothesis that draft accuracy is improving regarding Total Win Shares is greater than .05. As a result, we can conclude that there is insignificant evidence that draft accuracy is improving with respect to Total Win Shares.

Our fourth set of hypotheses is

Ho: No trend for Avg. PPG coefficient

Ha: There is a decreasing trend (draft accuracy is improving)

The results are shown in Figure 16 below.

Figure 16. Cox-Stuart Test for Trend for PPG

PPG test for trend	
n	38
m	19
nneg	10
npos	9
P-Value	0.5
P-Value for Improved Drafting Hypothesis	0.5

From Figure 16 we can see that the P-Value of .5 for our hypothesis that draft accuracy is improving regarding Points Per Game is greater than .05. As a result, we can conclude that there is insignificant evidence that draft accuracy is improving with respect to Points Per Game.

Our fifth set of hypotheses is

Ho: No trend for Avg. APG coefficient

Ha: There is a decreasing trend (draft accuracy is improving)

The results are shown in Figure 17 below.

Figure 17. Cox-Stuart Test for Trend for APG

APG test for trend						
n	38					
m	19					
nneg	9					
npos	8					
nzero	2					
P-Value	0.5					
P-Value for Improved Drafting Hypothesis	0.5					

From Figure 17 we can see that the P-Value of .5 for our hypothesis that draft accuracy is improving regarding Assists Per Game is greater than .05. As a result, we can conclude that there is insignificant evidence that draft accuracy is improving with respect to Assists Per Game.

Our sixth set of hypotheses is

Ho: No trend for Avg. RPG coefficient

Ha: There is a decreasing trend (draft accuracy is improving)

The results are shown in Figure 18 below.

Figure 18. Cox-Stuart Test for Trend for RPG

RPG test for trend						
n	38					
m	19					
nneg	11					
npos	8					
P-Value	0.323803					
P-Value for Improved Drafting Hypothesis	0.323803					

From Figure 18 we can see that the P-Value of .323 for our hypothesis that draft accuracy is improving regarding Rebounds Per Game is greater than .05. As a result, we can conclude that there is insignificant evidence that draft accuracy is improving with respect to Rebounds Per Game.

Overall, none of our six Cox-Stuart tests for trend showed significant evidence towards their being an increase in drafting accuracy, but why is this? One theory we formed from these results was the idea that a decrease in the average drafted player's college experience may have negated the effects of improvements made in scouting.

We next turn to our second research topic which was to examine whether or not there was a relationship between the number of years a player played college basketball and the player's first three years of performance in the NBA. We began this topic by testing whether there was a decreasing trend in the average number of years of college basketball played among drafted NBA players. The average years of college basketball played for drafted players during the years 1981-2018 is given in Table 7.

Table 7. Average Years of College Basketball Experience for Drafted Players

YEAR	AVG. YEARS OF COLLEGE	YEAR	AVG. YEARS OF COLLEGE
1981	3.889	2000	3.163
1982	3.739	2001	2.875
1983	3.979	2002	3.302
1984	3.979	2003	3.079
1985	3.891	2004	2.778
1986	3.844	2005	2.596
1987	3.978	2006	3.111
1988	3.880	2007	3.085
1989	4.038	2008	2.857
1990	3.769	2009	3.125
1991	3.906	2010	2.849
1992	4.038	2011	2.978
1993	3.827	2012	2.863
1994	3.745	2013	2.891
1995	3.764	2014	2.826
1996	3.358	2015	2.652
1997	3.549	2016	2.568
1998	3.415	2017	2.400
1999	3.434	2018	2.340

Figure 19 is a graph of the average number of years of college basketball played for each of the years in the draft 1981-2018

Figure 19. Yearly Average Years of College Basketball Played Among Drafted Players

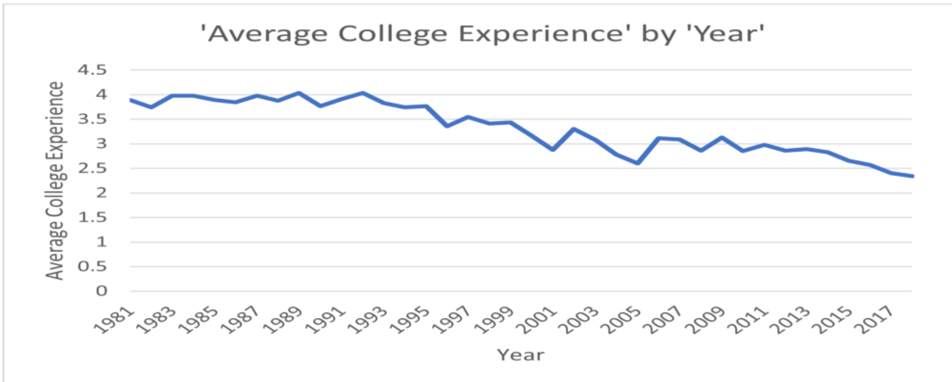


Figure 19 does indicate the average is decreasing but we want to test whether this is a significant decreasing trend. The Cox-Stuart test was performed based on the following set of hypotheses

Ho: No decreasing trend in average number of years of college basketball played

Ha: There is a decreasing trend in average number of years of college basketball played

Results of this test can be seen in Figure 20.

Figure 20. Cox-Stuart Test for Trend for College Experience

Average Years of College Basketball Played test for trend	
n	38
m	19
nneg	0
npos	19
P-Value	1.91E-06
P-Value for Decreasing College Play Hypothesis	1.91E-06

From Figure 20 we can see that the P-Value is approximately 0 for our hypothesis that drafted players are playing less years of college basketball then they used to. Since this value is less than .05, we can conclude that there is significant evidence that the average amount of college basketball played by drafted players has decreased. The results of this test fall in line with our previous thought that players are playing less college basketball.

Based on our prior test we can conclude that it's a possibility that players entering the draft with less college experience may be part of the reason why we did not see any improvement in drafting accuracy. In order to try and help answer this question, we developed regression models to determine whether there was a significant relationship between a player's first three years of performance in the NBA with the number of years they spent playing college basketball. The models in (1) found in Chapter 3 were fit based on the draft data.

Before going into our individual model coefficients however, we first want to exclude models that did not have a F-Statistic that is significant. We are looking to exclude these models because an F-Statistic being non-significant means that the model and all of its coefficients can't be used for analysis. Table 8 Shows all of the models used and highlights the P-Value of models that have a significant F-Statistic.

Table 8. P-Values of F- Statistics for all Regression Models

MODEL BEING USED	WIN SHARES STAT BEING PREDICTED	P-VALUE OF F-STATISTIC	SAMPLE SIZE
All Players, Baseline 0	**Offensive Win Shares**	**0.000**	**1840**
All Players, Baseline 0	**Defensive Win Share**	**0.000**	**1840**
All Players, Baseline 0	**Total Win Shares**	**0.000**	**1840**
Lottery, Baseline 0	**Offensive Win Shares**	**0.290**	**497**
Lottery, Baseline 0	**Defensive Win Shares**	**0.174**	**497**
Lottery, Baseline 0	**Total Win Shares**	**0.151**	**497**
All Players, (Top 3 Removed) Baseline 0	**Offensive Win Shares**	**0.000**	**1808**
All Players, (Top 3 Removed) Baseline 0	**Defensive Win Share**	**0.000**	**1808**
All Players, (Top 3 Removed) Baseline 0	**Total Win Shares**	**0.000**	**1808**
Lottery, (Top 3 Removed) Baseline 0	**Offensive Win Shares**	**0.166**	**469**
Lottery, (Top 3 Removed) Baseline 0	**Defensive Win Shares**	**0.027**	**469**
Lottery, (Top 3 Removed) Baseline 0	**Total Win Shares**	**0.028**	**469**
All Players, Baseline 1	**Offensive Win Shares**	**0.000**	**1650**
All Players, Baseline 1	**Defensive Win Share**	**0.000**	**1650**
All Players, Baseline 1	**Total Win Shares**	**0.000**	**1650**
Lottery, Baseline 1	**Offensive Win Shares**	**0.755**	**464**
Lottery, Baseline 1	**Defensive Win Share**	**0.182**	**464**
Lottery, Baseline 1	**Total Win Shares**	**0.414**	**464**
All Players, (Top 3 Removed) Baseline 1	**Offensive Win Shares**	**0.003**	**1627**
All Players, (Top 3 Removed) Baseline 1	**Defensive Win Share**	**0.000**	**1627**
All Players, (Top 3 Removed) Baseline 1	**Total Win Shares**	**0.000**	**1627**
Lottery, (Top 3 Removed) Baseline 1	**Offensive Win Shares**	**0.456**	**442**
Lottery, (Top 3 Removed) Baseline 1	**Defensive Win Share**	**0.097**	**442**
Lottery, (Top 3 Removed) Baseline 1	**Total Win Shares**	**0.170**	**442**

Based on Table 8 we can see that all of the models that used (All players) as a dataset had significant F-Statistics. This however was not the case for models that used lottery players only.

In fact, the only lottery models that were significant had no significant coefficients present. As a result, we will only be examining the models that use all players as their dataset.

Table 9 below shows the outputs of our regression models that contain a baseline of 0 years of college basketball played when taking into account all players drafted. In this Table's first column we are shown which specific Win Share statistic is being predicted while our second column contains what we would predict the value of our statistic to be if we were to play 0 years of college basketball. Our third column shows us which level of college experience we are comparing our baseline to, and our fourth column shows us what level of change in Win Shares we should expect of see if the level of college experience being compared to our baseline occurs instead of 0. Finally, our fifth column calculates our P-Value which signifies whether the difference between our baseline of 0 years played in college and the level of college experience we are comparing to is significant or not.

We can see from Table 9 below that only 4 years and 5+ years of college basketball played are significantly different than our baseline of 0 years played. However, for both of these two significant differences, the added years of experience actually appear to be decreasing the projected number of Win Shares which is the opposite of what we expected.

Table 9. Regression Results for Models Using Baseline 0 Years of College Basketball Played With Dataset of all Players

STATISTIC BEING PREDICTED (DEPENDENT VARIABLE)	BASELINE VALUE (INTERCEPT)	INDICATOR VARIABLE BEING COMPARED (IN YEARS)	ESTIMATE OF INDICATOR VARIABLE EFFECT	P-VALUE OF INDICATOR VARIABLES EFFECT
Offensive Win Shares	**3.2707**	**1 Year**	**-0.3781**	**0.587**
Offensive Win Shares	**3.2707**	**2 Years**	**-0.4872**	**0.475**
Offensive Win Shares	**3.2707**	**3 Years**	**-0.4362**	**0.510**
Offensive Win Shares	**3.2707**	**4 Years**	**-1.3465**	**0.034**
Offensive Win Shares	**3.2707**	**5+ Years**	**-2.4815**	**0.000**
Defensive Win Shares	**3.7756**	**1 Year**	**-0.4155**	**0.397**
Defensive Win Shares	**3.7756**	**2 Years**	**-0.5833**	**0.224**
Defensive Win Shares	**3.7756**	**3 Years**	**-0.5934**	**0.204**
Defensive Win Shares	**3.7756**	**4 Years**	**-1.4956**	**0.001**
Defensive Win Shares	**3.7756**	**5+ Years**	**-2.4112**	**0.000**
Total Win Shares	**7.0463**	**1 Year**	**-0.7936**	**0.465**
Total Win Shares	**7.0463**	**2 Years**	**-1.0705**	**0.314**
Total Win Shares	**7.0463**	**3 Years**	**-1.0296**	**0.319**
Total Win Shares	**7.0463**	**4 Years**	**-2.8421**	**0.004**
Total Win Shares	**7.0463**	**5+ Years**	**-4.8927**	**0.000**

Table 10 below shows the outputs of our regression models that contain a baseline of 0 years of college basketball played (with the top 3 performers in each experience level removed) taking into account all players drafted. In this Table's first column we are shown which specific Win Share statistic is being predicted while our second column contains what we would predict the value of our statistic to be if we were to play 0 years of college basketball. Our third column shows us which level of college experience we are comparing our baseline to, and our fourth column shows us what level of change in Win Shares we should expect of see if the level of college experience being compared to our baseline occurs instead of 0. Finally, our fifth column calculates our P-Value which signifies whether the difference between our baseline of 0 years played in college and the level of college experience we are comparing to is significant or not.

We can see from Table 10 below that only 5+ years of college basketball played is significantly different than our baseline of 0 years played. However, the added years of experience actually appears to be decreasing the projected number of Win Shares which is the opposite of what we expected.

Table 10. Regression Results for Models Using Baseline 0 Years of College Basketball Played With Top 3 Players of Each Experience Level Removed

STATISTIC BEING PREDICTED (DEPENDENT VARIABLE)	BASELINE VALUE (INTERCEPT)	INDICATOR VARIABLE BEING COMPARED (IN YEARS)	ESTIMATE OF INDICATOR VARIABLE EFFECT	P-VALUE OF INDICATOR VARIABLES EFFECT
Offensive Win Shares	**1.9270**	**1 Year**	**0.4914**	**0.447**
Offensive Win Shares	**1.9270**	**2 Years**	**0.5467**	**0.387**
Offensive Win Shares	**1.9270**	**3 Years**	**0.6516**	**0.290**
Offensive Win Shares	**1.9270**	**4 Years**	**-0.1064**	**0.858**
Offensive Win Shares	**1.9270**	**5+ Years**	**-1.3409**	**0.040**
Defensive Win Shares	**2.9730**	**1 Year**	**0.1226**	**0.793**
Defensive Win Shares	**2.9730**	**2 Years**	**0.0280**	**0.951**
Defensive Win Shares	**2.9730**	**3 Years**	**0.0066**	**0.988**
Defensive Win Shares	**2.9730**	**4 Years**	**-0.7604**	**0.076**
Defensive Win Shares	**2.9730**	**5+ Years**	**-1.8126**	**0.000**
Total Win Shares	**4.9000**	**1 Year**	**0.6140**	**0.542**
Total Win Shares	**4.9000**	**2 Years**	**0.5748**	**0.560**
Total Win Shares	**4.9000**	**3 Years**	**0.6582**	**0.493**
Total Win Shares	**4.9000**	**4 Years**	**-0.8668**	**0.348**
Total Win Shares	**4.9000**	**5+ Years**	**-3.1535**	**0.002**

Table 11 below shows the outputs of our regression models that contain a baseline of 1 year of college basketball played when taking into account all players drafted. In this Table's first column we are shown which specific Win Share statistic is being predicted while our second column contains what we would predict the value of our statistic to be if we were to play 1 year of college basketball. Our third column shows us which level of college experience we are comparing our baseline to, and our fourth column shows us what level of change in Win Shares we should expect of see if the level of college experience being compared to our baseline occurs instead of 1. Finally, our fifth column calculates our P-Value which signifies whether the difference between our baseline of 1 year played in college and the level of college experience we are comparing to is significant or not.

We can see from Table 11 below that only 4 years of college basketball played is significantly different than our baseline of 1 year played. However, the added years of experience actually appears to be decreasing the projected number of Win Shares which is the opposite of what we expected.

Table 11. Regression Results for Models Using Baseline 1 Year of College Basketball Played With Dataset of all Players

STATISTIC BEING PREDICTED (DEPENDENT VARIABLE)	BASELINE VALUE (INTERCEPT)	INDICATOR VARIABLE BEING COMPARED (IN YEARS)	ESTIMATE OF INDICATOR VARIABLE EFFECT	P-VALUE OF INDICATOR VARIABLES EFFECT
Offensive Win Shares	**2.8926**	**2 Years**	**-0.1091**	**0.799**
Offensive Win Shares	**2.8926**	**3 Years**	**-0.0581**	**0.883**
Offensive Win Shares	**2.8926**	**4 Years**	**-0.9684**	**0.005**
Defensive Win Shares	**3.3601**	**2 Years**	**-0.1679**	**0.576**
Defensive Win Shares	**3.3601**	**3 Years**	**-0.1779**	**0.521**
Defensive Win Shares	**3.3601**	**4 Years**	**-1.0801**	**0.000**
Total Win Shares	**6.2528**	**2 Years**	**-0.2769**	**0.677**
Total Win Shares	**6.2528**	**3 Years**	**-0.2360**	**0.701**
Total Win Shares	**6.2528**	**4 Years**	**-2.0485**	**0.000**

Table 12 below shows the outputs of our regression models that contain a baseline of 1 year of college basketball played (with the top 3 performers in each experience level removed) taking into account all players drafted. In this Table's first column we are shown which specific Win Share statistic is being predicted while our second column contains what we would predict the value of our statistic to be if we were to play 1 year of college basketball. Our third column shows us which level of college experience we are comparing our baseline to, and our fourth column shows us what level of change in Win Shares we should expect of see if the level of college experience being compared to our baseline occurs instead of 1. Finally, our fifth column calculates our P-Value which signifies whether the difference between our baseline of 1 year played in college and the level of college experience we are comparing to is significant or not.

We can see from Table 12 below that only 4 years of college basketball played is significantly different than our baseline of 1 year played for both Defensive and Total Win Shares. However, the added years of experience once again appear to actually be decreasing the projected number of Win Shares which is the opposite of what we expected.

Table 12. Regression Results for Models Using Baseline 1 Year of College Basketball Played With Top 3 Players of Each Experience Level Removed

STATISTIC BEING PREDICTED (DEPENDENT VARIABLE)	BASELINE VALUE (INTERCEPT)	INDICATOR VARIABLE BEING COMPARED (IN YEARS)	ESTIMATE OF INDICATOR VARIABLE EFFECT	P-VALUE OF INDICATOR VARIABLES EFFECT
Offensive Win Shares	2.4185	2 Years	0.0553	0.887
Offensive Win Shares	2.4185	3 Years	0.1602	0.657
Offensive Win Shares	2.4185	4 Years	-0.5979	0.059
Defensive Win Shares	3.0955	2 Years	-0.0946	0.736
Defensive Win Shares	3.0955	3 Years	-0.1159	0.654
Defensive Win Shares	3.0955	4 Years	-0.8830	0.000
Total Win Shares	5.5140	2 Years	-0.0393	0.948
Total Win Shares	5.5140	3 Years	0.0442	0.937
Total Win Shares	5.5140	4 Years	-1.4809	0.003

After excluding the results of models that weren’t significant, we found that their only appears to be significant evidence regarding the idea that players with 0 or 1 years of college basketball experience perform better in their early NBA careers than players who play 4 or more years of college basketball. This decrease in performance however is likely due to the fact that in the modern college basketball landscape, players who play 4 or more years of college basketball usually only do so because they either suffer from injuries or were not good enough to enter the draft early.

In conclusion, there does not appear to be a significant relationship between years of college basketball played and early career performance in the NBA.

CHAPTER 5 CONCLUSION & FUTURE RESEARCH

We began this thesis with one main question, we wanted to know if the increase in available scouting data provided to NBA organizations was helping said organizations to draft with better accuracy. To test this theory, we began by collecting every player's statistics during their first three seasons for Points, Rebounds, and Assists as well as the advanced statistics Offensive, Defensive, and Total Win Shares. (*Basketball-reference.com - basketball statistics and history* 2000)

We then compared these six statistics to where each player was drafted in their draft class. By doing this we were able to create a correlation score for every draft between the years (1981-2018) for each of our six statistics. Using these correlation scores, we then performed testing to determine which if any, of our six statistics appeared to have at least decent correlation (less than -.50) with where a player was selected in the draft. The results of this testing resulted in Defensive Win Shares, Total Win Shares, Points, and Rebounds showing at least moderate correlation. However, why did Offensive Win Shares and Assists have less significant correlations? We theorize that perhaps Offensive Win Shares are harder to predict for a player than Defensive Win Shares, that is to say, perhaps it is easier to determine which players will be good defenders in the NBA then it is to determine which players will be good offensively. As for Assists, we hypothesize that NBA organizations may simply be more interested in players who can score points and rebound then those who can pass. Both are just theories however and would need to have their own analysis done in order to know with certainty whether this is the case or not.

After testing for significant correlation of less than -0.50 for each of our six statistics, we then performed testing on each of them to determine which, if any of them were showing a decreasing trend which implied that drafting was getting more accurate. However, we were shocked to find that none of our six statistics produced significant evidence in favor of drafting becoming more accurate.

Based off these results we determined that there was insignificant evidence to conclude that drafting was becoming more accurate, but why was this the case? We theorized that the main reason that the increase in scouting is not producing better results may be due to the simultaneous decrease in college basketball experience that has come as a trickle-down effect of players entering the draft at younger ages than they would have in the past.

To determine whether our theory held weight or not we decided to create regression models to test the effects that additional years in college played, in a player's early career performance (first 3 seasons). Before creating these regression models however, we first confirmed through another test for trend that college experience had decreased significantly over time.

We created two different models, one containing a comparative baseline of 0 years played in college and the other containing a baseline of 1 year played in college. However, regardless of which of these two regression models we used or which of our four data subsets we used (first 2 rounds, lottery only, first 2 rounds top 3 of each age group removed, lottery only top 3 of each age group removed) we always failed to generate any significant evidence that playing longer in college had a positive effect on performance. These results disprove our theory that the change in college basketball experience, especially for players in the draft lottery who are now almost always 1 year college players, was the reason for the lack of improvement in draft accuracy.

In terms of future research there are many different aspects of this thesis that we could explore further. We could look into why it is that Assists per Game and Offensive Win Shares don't contain reasonable correlation with draft pick ranks, and in doing so could hopefully gain a better understanding as to why it is that these measures of performance didn't perform as well as our other four measures.

Another area where we could dig deeper relates to our regression models. For each regression model, only variables related to college experience were used, however if we were to set up the model to account for what year the data was coming from and made note of how the coefficients changed in response to the change in year, we could possibly get a much clearer idea of how much college experience matters on a yearly basis instead of just overall.

An additional thing we could have tested and may do in future research is adding in player positions to our draft accuracy analysis. It would likely be extremely interesting to see how position effected what statistics had the highest correlation with draft pick rank. Also, assuming this experiment was conducted, many additional simple statistics would likely be added to the analysis as well.

Finally, it would likely be very interesting to perform this analysis on other major professional sports that have a draft. As analysis of drafts in other sports would allow us to determine if the idea that teams aren't doing a better job at drafting is true for all major sports or is just isolated to the NBA.

WORKS CITED

Arel, B., & Tomas, M. J. (2011). The NBA draft. *Journal of Sports Economics, 13*(3), 223-249. doi:10.1177/1527002511406128

Basketball-reference.com - basketball statistics and history. (2000). Retrieved 2022, from https://www.basketball-reference.com/

Berger, T., & Daumann, F. (2021, July 21). Jumping to conclusions – an analysis of the NBA draft combine athleticism data and its influence on managerial decision-making. Retrieved October 19, 2022, from https://www.emerald.com/insight/content/doi/10.1108/SBM-11-2020-0117/full/html

Calculating win shares | college basketball at sports-reference.com. (2000). Retrieved 2022, from https://www.sports-reference.com/cbb/about/ws.html

Coates, Dennis, and Babatunde Oguntimein. *The Length and Success of NBA Careers: Does College Production Predict Professional Outcomes?* 2008.

Garnes, K. (2019). The monetary and social opportunity cost of entering the NBA draft directly from high school versus completing one year of college before entering the draft. Retrieved 2022, from https://triceratops.brynmawr.edu/handle/10066/21731

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. https://doi.org/10.1177/1527002505281228

How to: Binomial test Sign test McNemars test Cox & Stuart. (2003). Retrieved 2022, from https://influentialpoints.com/Training/binomial_and_related_tests.htm

Ichniowski, C., & Preston, A. (2017, July 23). Does March madness lead to irrational exuberance in the NBA draft? high-value employee selection decisions and decision-making bias.

Mathewson, T. (2018, December 13). How do players get to the NBA from different countries? Retrieved 2022, from https://globalsportmatters.com/business/2018/12/13/how-do-basketball-players-go-pro-in-different-countries/

NBA win shares | basketball-reference.com. (2000). Retrieved 2022, from https://www.basketball-reference.com/about/ws.html

Sailofsky, D. (2018, April). Drafting errors and decision making theory in the NBA draft. Retrieved 2022, from https://dr.library.brocku.ca/handle/10464/13452

Scoring system in basketball: how point scoring works. (n.d.). Retrieved 2022, from https://www.chaseyoursport.com/Basketball/Scoring-System-in-Basketball/2560

Spearman Correlation - an Overview | ScienceDirect Topics. *Www.sciencedirect.com*, www.sciencedirect.com/topics/mathematics/spearman-correlation.

Sugai, Nick. (2010). The Effect of Early Entry to the NBA. Retrieved 2022, from Docslib. https://docslib.org/doc/854365/the-effect-of-early-entry-to-the-nba

Teramoto, Masaru, et al. "Predictive Validity of National Basketball Association Draft Combine on Future Performance." *Journal of Strength and Conditioning Research*, vol. 32, no. 2, Feb. 2018, pp. 396–408, 10.1519/jsc.0000000000001798.

The Long Weird History of the NBA Draft. (2013). Retrieved 2022, from https://nbahoopsonline.com/Articles/History/Drafthistory.html

Warner, A. (2021, January 19). What is an assist in basketball? Retrieved 2022, from https://thatsportlife.com/what-is-an-assist-in-basketball/

What is an assist in basketball? (2021, June 1). Retrieved 2022, from https://www.hoopsaddict.com/what-is-an-assist-in-basketball/

What is a rebound? definition & meaning on sportslingo.com. (2022, January 05). Retrieved 2022, from https://www.sportslingo.com/sports-glossary/r/rebound/