UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE RELATIONSHIP BETWEEN NCAA COLLEGE BASKETBALL RECRUITING STRATEGIES AND TEAM SUCCESS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

Master of Science

By

SAMUEL DAVID LUPTON

Norman, Oklahoma

2021

THE RELATIONSHIP BETWEEN NCAA COLLEGE BASKETBALL RECRUITING STRATEGIES AND TEAM SUCCESS

A THESIS APPROVED FOR THE

DEPARTMENT OF HEATH AND EXERCISE SCIENCE

BY THE COMMITTEE CONSISTING OF

Dr. Daniel J. Larson, Chair

Dr. Christopher Black

Dr. Le Wang

Table of Contents

Abstract

The study explored whether there is a relationship between the different strategies (based on past outcomes) used by college basketball teams while recruiting and how successful teams are after implementing their specific recruiting strategy. The study analyzed data from ten conferences in NCAA Division One Men's College Basketball over six seasons (2013-14 to 2018-19). For analysis, a K-Means Cluster Analysis was used to identify different recruiting groups (clusters) among teams based on the number of five stars signed, the number of first-season (freshman) departures to the NBA draft, and the number of second through fourth year departures to the NBA draft by a team. Based on the results from the cluster analysis, an OLS regression was implemented to test the relationship between recruiting clusters and team success (winning percentage) in the regular-season, along with other university characteristics. The last part of the analysis consisted of an ordered logistic regression to test the relationship between these same clusters/university characteristics and the number of appearances in the NCAA Men's Basketball tournament by a team. The results concluded that there was no significant effect on regular-season performance based on team recruiting clusters, but there was an effect on NCAA tournament appearances with teams who appeared in clusters "OneAndDone" and "HighRecDev".

Chapter One: Introduction

The amateur draft is something used by sports leagues to promote competitive balance and to keep the cost of a player at a minimum by giving a team exclusive rights to negotiate with a new player who is entering a league (Fort and Quirk, 1995). Unlike many other professional sports leagues, the National Basketball Association (or NBA) requires a player to be at least 19-years of age and one year removed from high school before they can enter their name into the NBA draft (Giambalvo, 2018). Each professional sports league has its advantages and disadvantages regarding draft age and eligibility, but if a player decides to go to a college to play a sport, the NBA's requirement on age and college eligibility is the smallest of the "Big Four" North American sports leagues. For college basketball programs, this policy creates player turnover consequences that can affect the way in which a team decides to recruit their players. Should they recruit the best talent that could leave after one season? Or should programs recruit slightly less talented players that will provide stability for a program for what could be up to four seasons?

The history of NBA draft eligibility is one that endured a significant change in the last 20-years. From 1995-2005, notable NBA stars like Kobe Bryant, Kevin Garnett, and LeBron James entered their names into the NBA draft directly out of high school (Burleyson, 2018). While this strategy was profitable for some players, others would either go undrafted or find themselves out of the NBA in a few years with no education to their name. In response to this, the 2005 collective bargaining agreement between the NBA and its player's union agreed that the age of draft eligibility should be at least 19-years old with one year removed from high school (Burleyson, 2018). This rule change did not affect many of the 18-year old college prospects, as

the early entry pathway was only an option for the players who were considered rising stars who did not need college basketball development. Still, some NBA prospects were most definitely ready for the NBA draft immediately out of high school but had to play a year in college instead.

Since the NBA changed its draft eligibility requirements as of the 2006 draft (Burleyson, 2018), NCAA (National Collegiate Athletic Association) college basketball powerhouses like Duke, Kansas, Kentucky, and others have been able to attract the nation's best players for a one-year stay at the collegiate basketball level in preparation for the NBA draft (NBA.com, 2017). These players have been nicknamed "one-and-done" by many. The "powerhouses" mentioned above are considered historic by many and they continue to gather plenty of national attention to this day, much to the benefit of a "one-and-done" player.

Talented college basketball prospects are evaluated by experts who work for companies like ESPN, and 247Sports. These companies grade players by giving the player a numerical grade with an associated number of stars, ranging from one to five (Grading Scale, 2006). However, these evaluators have different standards for assigning numerical grades for prospects, as ratings are subjective to the company they are made by and are not consistent between the two. Because of this, this study will focus only on the numerical ratings from ESPN instead of cross-referencing between the two evaluators (ESPN and 247Sports).

For example, in the 2018 ESPN Top 100 College Basketball Recruiting Database, Zion Williamson (the eventual first overall pick in the 2019 NBA draft) was given a numerical rating of 96 and five stars (ESPN Basketball Recruiting, 2018). A five-star caliber player on ESPN's Top 100 database will have a numerical grade of 90-100, while a four-star player will have a numerical grade ranging from 80-89, and so on (Grading Scale, 2006). In this study, these

players will be referred to as “five-star” or “four-star” athletes, but for the analysis, a player’s numerical score will be used (ex: Zion Williamson: 96).

Table 1

2018-19 ESPN Five Star Recruits

ESPN Rankings 2018	Name	ESPN Numerical Grade	Draft Position
1	R.J. Barret	96	2
2	Zion Williamson	96	1
3	Cam Reddish	96	10
4	Bol Bol	96	44
5	Romeo Langford	95	14
6	Nassir Little	95	25
7	Keldon Johnson	94	29
8	Quentin Grimes	94	Not Drafted in 2019
9	Anfernee Simons	94	Not Drafted in 2019
10	Jalen Smith	93	Not Drafted in 2019
11	Louis King	93	Not Drafted in 2019
12	Nazreon Reid	93	Not Drafted in 2019
13	Darius Bazley	93	23
14	E.J. Montgomery	93	Not Drafted in 2019
15	Moses Brown	93	Not Drafted in 2019
16	Darius Garland	93	5
17	Tre Jones	93	Not Drafted in 2019
18	Charles Bassey	92	Not Drafted in 2019
19	Simisola Shittu	92	Not Drafted in 2019
20	Ashton Hagans	92	Not Drafted in 2019
21	Emmit Williams	91	Not Drafted in 2019
22	Jaylen Hoard	91	Not Drafted in 2019
23	Coby White	91	7
24	Devon Dotson	90	Not Drafted in 2019
25	Immanuel Quickly	90	Not Drafted in 2019
26	Jahvon Quinerly	90	Not Drafted in 2019

Ideally, a player that is given a numerical grade over 90 (or five-stars) should fall into the category of a potential “one-and-done” player, meaning they can leave college to enter the NBA draft after just one season. The 2018 ESPN Top 100 contained 26 five-star players, with the rest of the 100 containing four-star talent (ESPN Basketball Recruiting, 2018). Of the 26 players that were assigned a numerical rating above 90 in the class of 2018, nine of them were selected in the first round of the NBA draft (2019 NBA Draft results, 2019), while another was selected in the second round (see table below for 2018 five-stars). The remaining five-star players who did not declare for the draft (after their freshman year) could be drafted later in their college career, but

this is by no means a guarantee. Three out of the 16 five-star players in the 2018 class not selected in the 2019 NBA draft were drafted after their sophomore season (in the 2020 draft). Another three players in this group are still playing for their respective university, while the remaining ten players entered the NBA as undrafted free agents. So, although ESPN does their best to predict the potential of these players, there is by no means a guarantee that highly-rated players will get drafted in accordance with their predicted potential. If a player declares for the draft and is not selected, they may return to school and continue playing college basketball (Giambalvo, 2018). In the 2018 NBA draft, 37 players were selected (not including players who competed internationally) who weren't evaluated at a level of 90 or above (Five-stars) by ESPN. However, only five of these players (<90) left college after one season (the rest played at least two seasons in college).

Purpose of the Study

A single five-star player could possess the potential to have a monumental impact on a college basketball program's performance. This comes from the fact that there are only five players competing on a basketball court for a team at a time, which is the lowest of any big four sports league (football: 11, baseball: 9, hockey: 6). This gives each player more individual responsibility (within a team game) for the teams overall performance. This individual impact of a five-star player, of course, would not be possible without the current NBA rules that require a player to be at least one year removed from high school (Giambalvo, 2018). Powerhouse schools like Duke, Kentucky, Kansas and others have made a habit of not only attracting these high-caliber prospects but winning with them too. However, at the end of the NCAA Basketball Tournament (otherwise known as March Madness), only one National Champion is crowned, and it is not always the team with the most five-star talent. One could actually argue that the

development of returning college players could match up equally (or better) than that of a team made up of first-year players with high recruiting grades.

March Madness is a 68-team basketball tournament made up of seven rounds, with the goal of crowning one college basketball program the National Champion by the end of the tournament (Wilco, 2020). The 68 teams that compete for the NCAA Basketball Tournament are all a part of Division 1, but they come from 32 different Athletic Conferences (NCAA College Basketball Conferences, n.d.). All told, the 32 conferences in division one contain 353 schools, but five of the conferences in Division 1 (Atlantic Coast Conference, Southeastern Conference, the Big Ten, the Big 12, and the Pac-12) are in what is known as the Power Five. Generally, the Power Five conferences exceed other conferences in terms of revenue and sometimes performance (depending on the sport).

Schools that compete in a Power Five conference dominate in NCAA football, but in the NCAA Basketball Tournament, schools from outside of the Power Five have a legitimate shot of winning the National Championship. 20 years since 2000, a non-Power Five school has won the National Championship seven times (Connecticut (3), Villanova (2), Louisville, Syracuse). In that same timeframe, the Big Ten (Michigan State) and Big 12 (Kansas) have only won the NCAA title once each, and the Pac-12 has none. However, the ACC has crowned eight champions (UNC (3), Duke (3), Virginia, Maryland), while the SEC has a total of three champions (Florida (2), Kentucky) (DI Men's Basketball Championship History, 2019). Yes, some of these teams can bring in five-star recruits consistently, but not all can. Every school would love access to the best recruits every year, but when we consider the fact that there were only 26 five-star players in the 2018 ESPN Top 100 (ESPN Basketball Recruiting, 2018), and

about 1,500+ players start as first year players in Division I basketball, one must assume that some schools have no other choice but to implement a different strategy.

The purpose of this paper is to explore whether there is a relationship between the different strategies used by college basketball teams while recruiting and how successful teams are after implementing their specific recruiting strategy.

Research Questions

Each research question in this study comes with sub-questions that will be critical in answering the overarching questions. The first research question will be whether it is possible to reliably differentiate program recruiting strategies based on historical behavior. The second question is asks if it's clear that some programs rely specifically on multiple "one-and-done" players to come in annually?

The final research question brings recruiting strategy focus back to team success. Can the differential strategies discussed above make a team more or less likely win at the national level when controlling for many program specific conditions; including, but not limited to, available financial resources, local and student populations, local income, and geographic context? If they can, how consistently can they do so? This study will look at the number of NCAA tournament appearances by a team as an indicator of success, along with the regular season winning percentage of a team. In the past, teams competing in the NCAA tournament (a 68-team, single elimination tournament) have been known to drastically under or over perform. The NCAA tournament's outcome variability (created by accomplished teams losing to teams that aren't as successful on paper) creates the need for this study to find another way to measure a team's success, which is why this analysis will include regular season winning percentage of a team.

Hypotheses

H0: Programs do not recruit significantly different across NCAA men's college basketball.

H0: Recruitment clusters are not related to team performance.

H1: Programs recruit differently across NCAA men's college basketball.

H2: Team recruiting clusters will relate to winning percentage at a significant level.

H3: Team recruitment clusters will relate to the number of appearances a team has in the NCAA tournament at a significant level.

Significance

This study's findings could significantly impact the game of college basketball. For example, if it is evident that teams who recruit players that will stay for multiple years tend to perform better conditional on their program characteristics, some coaches may be able to drastically change the way they decide to recruit to optimize their competitive outcomes.

Limitations

One limitation of this study would be its definition of success. A team can be successful without winning the National Championship (getting through multiple rounds of the NCAA Tournament is very difficult and should be applauded). However, many of the historical programs that compete in a Power Five conference would say that winning a National Championship is the end goal of every season. The problem that comes with these expectations is that only one team can be the victor in the NCAA Tournament, and many historic teams from fantastic basketball programs fall short of this goal every year. Another limitation of this study comes from its reliance on the data collected from the ESPN Top 100 database. This is third party data that is collected by ESPN's employees that seems to be mostly accurate, but by no

means is it perfect. A perfect system for evaluating high school players does not exist, as players are over or undervalued at times because of things like human error, evaluator biases, and the differences in talent regarding the skill level of opposing players faced by high school prospects. Information about the qualifications and identity of ESPN's employees is unavailable, as they are identified as "ESPN Analyst" on espn.com.

Delimitations

1. This study considered ESPN's top 100 rankings and NCAA Tournament results from 2013-2019.
2. Ten collegiate athletic conferences were used in a cluster analysis to differentiate recruiting styles. ("*" is given to Power Five Schools)
    1. Atlantic Coast Conference *
    2. Southeastern Conference *
    3. Big-Ten *
    4. Big 12 *
    5. Pac-12 *
    6. Big East
    7. American
    8. Atlantic 10
    9. West Coast Conference
    10. Mountain West Conference

Assumptions

1. The ESPN player rankings are accurate predictors of basketball talent for the most part, but not perfect.

2. Any statistical cluster analysis or regression analysis will be properly conducted.
3. It is assumed that college basketball teams are doing their best to recruit the best players.
4. It can be assumed that college basketball teams are actively trying to win as many basketball games as possible.

Operational Definitions

1. NBA Draft - A draft used by the National Basketball Association that allows team to select the rights to amateur players.
2. NCAA - Governing body for collegiate athletics in the United States.
3. NCAA Tournament (March Madness) - A 68-team, single elimination basketball tournament with seven-rounds.
4. ESPN Top 100 - High school college basketball rankings from college basketball experts who work for ESPN.
5. One-and-done - A high quality collegiate player that only plays one season before entering the NBA draft.

Chapter Two: Literature Review

Introduction

The purpose of this paper is to explore whether there is a relationship between the different strategies (denoted by past recruiting outcomes) used by college basketball teams while recruiting and how successful teams are after implementing their specific recruiting strategy. The first research question revolves around how there are differences in the way that NCAA D1 teams recruit players. Because NCAA Division 1 basketball has 347 teams and 32 conferences, it is reasonable to assume that each team does not have the same recruiting influence or strategy. As in professional sports, some college basketball programs have the ability to go out and get the best players every year, which leaves other teams (who are still expected to win by their fan base and school administration) to turn the recruits that are left behind by more prominent schools into a winning team. Finally, how successful are the results that come from each team's recruiting style? Do teams that have talented recruiting classes win more in the NCAA tournament/regular season, or do the teams that return and develop players have more success?

The literature search process for this study aimed to find similar studies that could be related the topics in this study (NBA Draft, College Basketball recruiting strategies, and team success). To do this, a search was conducted using services provided by the University of Oklahoma's library, including access to Google Scholar. While using these resources, the following keywords and phrases were used in the search process: ("College Basketball", "College Basketball Recruiting", "NCAA Recruiting Determinants", "NBA Draft", "One-and-Done")

Articles about recruiting practices in big-time college football were also included in this literature review, as recruiting information and practices for both sports (for bigger schools in Division One) are highly documented and discussed by sites like Rivals.com, ESPN, and 247Sports. Otherwise, criteria for the literature examined in this section included any peer-reviewed study completed since the year 2000 related to the three variables used in the study.

This first section of this paper will outline the "draft" process in professional team sports and detail the unique policies behind player entry into the NBA draft. The second section (NCAA Basketball Recruiting) in this chapter will review previous research about recruiting and success in college basketball. The articles in this section dive into recruiting practices, along with the value of recruits, and recruiting determinants in NCAA basketball. The third section in this literature review will examine cross-sport recruiting effects, and big-time college football recruiting determinants and success. The final section will discuss the economic theory supporting the proposition that programs will choose to recruit using distinctly different approaches.

NBA Draft Policy

The NBA draft is a reverse order draft (where the worst teams in the prior season get better picks) where the 30 teams that compete in the NBA draft from a pool of prospects who have never played in the NBA before. The draft consists of two rounds (60 picks), where the top 15 picks are determined by a lottery system for the worst 15 teams. NBA draft prospects can come from two different backgrounds: college players and international prospects. Each prospect must declare draft eligibility 60 days prior to the draft date. To be eligible, a (college) player must be at least 19 years old during the calendar year of the upcoming draft, and must wait one season after graduating from high school. If a player does not want to play in college and instead

competes internationally, he is able to declare for the NBA draft at the age of 19 (Smith, 2019). Additionally, if a college player (who has college eligibility left) declares for the draft and is not selected by a team (within the only two-rounds of the draft), he may go back to school as long as he notifies the school's athletic director by the following Monday after the draft, by 5 p.m (Fominykh, 2019).

NCAA Basketball Recruiting

In "The impact of recruiting on NCAA basketball success," Treme and Burrus (2011) set out to observe if highly ranked recruiting classes affected wins, while also testing the effect that certain positions have on success. Instead of testing every team in Division 1, the results from seventy-nine different college basketball programs (from www.ncaa.com) were analyzed from the year 2004 to 2009. The majority of these teams played in conferences that are highly competitive in basketball every year. However, the study also pulled in teams that came from different conferences but consistently made the NCAA tournament. They included the following conferences: ACC, BIG EAST, BIG 12, Big 10, Pac 10 and SEC (All of which are in the Power 5 except for the Big East); and several other single teams: Butler, BYU, Davidson, Gonzaga, Memphis, and Xavier. Some of the team variables that were included in this study's model were previous and current wins by a team (regular season and NCAA tournament), returning players, one-and-done players, coaching change information, the high school player ranking by stars (from rivals.com) and position types of each player.

The results of this study indicated a relationship between both incoming and returning talent and regular-season victories, especially from players labeled as a guard. However, success in the NCAA tournament is harder to come by than regular-season success. The results for

tournament wins indicated that returning players had a more significant impact than one-and-done players.

Five years after their initial study "The impact of recruiting on NCAA basketball success," Treme and Burrus (2016) once again set out to answer questions about college basketball recruiting in "NCAA basketball: when does recruiting talent translate into wins for power conferences?" This study aimed to answer the question of when high-quality high school recruits began to contribute wins for Power 5 schools in NCAA Basketball.

Again, the authors chose to include data from 2004 to 2009 for every team in a power basketball conference (ACC, Big-Ten, Big-12, SEC, Pac 10, and the Big East) but chose to leave out all the other schools in Division 1. Player ranking data came from Rivals.com, and NCAA.com provided the data for wins by a program. Burrus and Treme reused all of the same variables that were mentioned above in their previous study. This time they added a variable denoting the number of Final Fours (the semifinal round of the NCAA Tournament) a team had been to during the timeframe evaluated in the study. Another difference in their regression equation was that they included a variable for three-player specifications (e.g., Class = Junior, Rank = Four-Star, Position = Forward).

The results of this study indicate that a team returning players with a higher than average (above a three-star) recruiting rank (overall rank given to teams based on the quality of recruits attained) will win around two games more per season. The results also say that highly rated freshman guards contribute instantly to team success, but four or five star forwards/centers will need more than one year becoming a driving force that aids a team's win total.

Richard Borghesi (2015) aimed to measure the effect that recruiting had on wins as well in "The Financial and Competitive Value of NCAA Basketball Recruits," but he chose to evaluate success from a different point of view. Instead of looking at regular-season win totals by a team, Borghesi chose to measure team success by using the RPI metric. RPI, or Rating Percentage Index, not only uses win-percentage to evaluate a team, but the winning percentage of their opponents as well. This metric can better determine the success of a team due to its ability to factor in a team's strength of schedule. Because there are many college basketball teams, each team is unable to play every other team in Division 1, making it harder to tell which teams are the best. For example, if Duke University and Wofford College each had 30 wins in a season, are these teams equal in both talent and level of achievement? It is likely that RPI would disagree, as Duke plays higher-quality competition and is a part of a Power Five conference. This study used the numerical star ranking from www.247sports.com to measure recruiting quality.

Borghesi later lists academic ranking, university population, and metro population as some of the determinants of recruiting success. Others would include a team's past success, and the amount of national exposure players get on national television. The results of this article about on-court success indicated that players with higher rankings are indicative of positive on-the-court performance. Notably, the study states that five-star recruits have twice the on-court (in wins) value of a four-star player.

Instead of continuing to look at studies that measure recruiting's effect on wins, this section looks to examine the inverse of that relationship: winning's effect on recruiting. "The

determinants of NCAA basketball recruiting outcomes" (Evans and Pitts, 2017) outlines potential factors that could impact the success of a basketball team's recruiting efforts. Using high school basketball ranking from Rivals.com, Evans and Pitts (2017) employed the 'Rivals 250 Bonus Scale' recruiting system. This system assigns a point total to highly touted recruits, with the #1 ranked recruit possessing 100 points and the #150 ranked recruit receive only 23 points (only the top 150 recruits are assigned points, and 23 is the bottom value used).

This study takes into account many basketball related variables, like the amount of talent in-state, recent and past on-court success (wins and end-of-season rankings), and the team's history of developing NBA talent. Other variables include the school's stadium capacity, the team's conference status, coaching status (new coach or experienced coach), and the school's acceptance rate (Evans and Pitts, 2017).

The authors of the study chose to use a regression analysis to test the impact of each variable, and eight of the eleven variables produced significant results. Notable variables that produced a positive correlation to recruiting success (when a team gathers a group of recruits with a high average rating on the 'Rivals 250 Bonus Scale') were a team's end-of-year ranking in the previous season if they finished in the top 25. A team's number of NCAA championships is also a significant contributor to a team's success in recruiting efforts. Additionally, programs that have sent a player to the NBA show a significant impact on recruiting success, adding eight recruiting points. However, a team's end of year ranking in the previous five seasons did not show significant results and therefore cannot be considered accurate. This study by Evans and Pitts (2017) provides an example of the value that highly touted recruits can bring to a college program by testing recruiting's relationship to different outcome variables.

NCAA Football Recruiting

The assumption that basketball recruiting is driven by the on-field success of a collegiate football program is an intriguing one. "Cross-Sport Recruiting Effects in NCAA D1 Football and Basketball", a study that was also done by Evans and Pitts (2018), examines this assumption by using many of the same variables as their previous research.

Again, the authors pulled their recruiting data from Rivals.com while using "Rivals 250 Bonus Scale" to assign numerical values, or "recruiting points," to recruiting classes. The results of this study concluded that although basketball recruiting success showed no correlation to prestigious football outcomes, football recruiting was impacted by a stellar performance from a school's basketball team.

Although there are differences in basketball and football recruiting, it still might be helpful to observe recruiting determinants in other big-time college sports. A regression analysis was used in “The Relationship between Recruiting and Team Performance in Division 1A College Football” (Langlett, 2003), which was a study that attempted to observe the relationship between a team's performance and recruiting in NCAA Division 1 College Football. The analysis examined the best ten recruiting classes in the previous five years and observed teams that were consistently in the end-of-year top-25 rankings (Coaches Poll and Associated Press). The results from the polls were averaged together into a final ranking for each school. The results of Langeletts' analysis showed evidence that a higher rated recruiting class would provide a better chance of a lower (better) end of year ranking.

In “College Football Success: The Relationship Between Recruiting and Winning”, conducted by Cary A. Caro, the relationship between recruiting and on-field success in Division

One FBS Football was examined. To examine this relationship, Caro's study used a regression analysis and recruiting information from Rivals.com. The conferences included in the study's analysis were the ACC, Big Ten, Big 12, Big East, Pac-10, and the SEC. The recruiting data used in the study was collected from 2004-to-2009, and the outcome data was drawn from the NCAA's website from 2005-to-2010. The study's dependent variable was conference winning percentage, which helped to account for out-of-conference scheduling inconsistencies. The results of this study indicated that for three of the conferences included in the sample (Big Ten, Big-12, and SEC) 63-to-80% of a school's conference wins come as a result of recruiting success.

Non-NBA Professional Sports Draft Policy

The National Football League requires a player to either use up their college eligibility (players are only eligible to play sports for four years in college, in most cases) and be at least three years removed from high school before being deemed draft eligible ("The Rules of the NFL Draft," n.d.). Major League Baseball is more lenient than the NBA and NFL, allowing players to enter the draft directly from High School. However, if a player enrolls in college, the player must complete their Junior or Senior year and must be at least 21-years of age before they acquire draft eligibility ("How the MLB Draft Works," 2008). The National Hockey League requires a player to be 18-years of age on or after September 15th or 20-years of age before December 31st for draft eligibility ("NHL Draft Rules," n.d.).

Game Theory and Nash Equilibrium

Game Theory is a concept that was developed to observe the actions of players (firms, parties, etc.) by using mathematical models in either a cooperative or non-cooperative

competitive environment where players are able to make their own independent decisions based on the firm's goals while using the resources they have available to them (Maschler, 2013). Nash Equilibrium is a concept of Game Theory where firms try to find the optimal solution/direction regarding the game (or competitive scenario) they are currently in, where players won't gain anything from changing their current strategy as long as their competitor's strategies are confidently known and remain the same (Chatterjee, 2014).

Both Game Theory and the concept of a firm's Nash Equilibrium could theoretically be applied to the way in which college basketball programs choose to recruit players. The competitive environment of college basketball recruiting presents a fixed amount of resources (or high school players) that programs have the opportunity to compete for. As mentioned above, some programs might have the ability to attract highly touted five star recruits that could help the program in the short term. On the other hand, smaller programs might rely on a different strategy of annually recruiting lower-ranked players and developing them (in the long-term) so that they may contribute to team success in the future. It could be said that smaller programs (e.g. a non-Power Five basketball program) that recruit outside of their own abilities (attempting to sign five-star players without success or inconsistently doing so) might hurt their programs future success by choosing a recruiting strategy that isn't possible or sustainable.

We can consider the expected outcome using a single period simultaneous move game between two teams competing for resources by deciding on a recruiting strategy. We consider α to be a winning percentage by having more "one and done" players than the other team, where $\alpha > 0.5$. In the case of equivalent teams in terms of recruiting and coaching, and assuming 5 star recruiting is more costly, a simple dominant Nash equilibrium strategy results in a prisoner's dilemma outcome of both teams pursuing 5 star recruits (see Table 2).

Figure 1

Simple Payoff Matrix (Win %)		University A	
		Developing Players	Sign Five Star
University B	Developing Players	.5; .5	$1 - \alpha$; α
	Sign Five Star	α; $1 - \alpha$	.5; .5

However, we consider adding a factor (θ) in the payoff matrix reflecting a general recruiting advantage for one team, in this case provided by structural/market conditions more favorable for 5 star recruiting (See Table 3). If we define $\theta > 0$, the outcome will result in a stable Nash equilibrium where the disadvantaged team will choose to invest in developing players, and the higher resource team will more heavily recruit 5 star players. This result is because the dominant strategy for University A would always be to recruit 5 star players (i.e. $(.5) < (\alpha + \theta)$ and $(1 - \alpha + \theta) < (\alpha + \theta)$), while, coupled with the assumption of increased costs for 5 star recruiting, University B would choose to develop players rather than pursue 5 star recruiting.

Figure 2

Simple Payoff Matrix (Win %)		University A	
		Developing Players	Sign Five Star
University B	Developing Players	.5; .5	$(1– \alpha - \theta)$; $(\alpha + \theta)$
	Sign Five Star	$(\alpha – \theta)$; $(1 - \alpha + \theta)$	$(1 - \alpha – \theta)$; $(\alpha + \theta)$

While this framework begs to be expanded into a repeated interaction, mixed-strategy, dynamic model, this initial formulation offers a rationale for a general expectation that teams may choose different recruiting philosophies, particularly based on their comparative advantages.

The recruiting strategies (clusters) that this study aims to identify could present relevant information to college basketball's current recruiting culture, where a team could be negatively affected if they choose to sway from a strategy outside of their realistic capabilities (while other programs hold to the strategy that consistently provides successful results).

Summary of Review

During the research process for this study, no information has specifically examined whether or not recruiting groups exist and team success in the NCAA tournament. However, some information about the general success of the NCAA teams and their recruiting practices was available. Even though the research focuses in the articles above were different at times, most of these studies used similar methods in gathering data about recruiting classes. Other helpful insights came from studies that did not necessarily measure the outcome results of a recruiting class, but the determinants that caused a program to excel in attracting top talent every year. Additionally, studies regarding cross-sport recruiting effects and secondary information about NBA draft policies were used in this review. Finally, economical concepts such as Game Theory and Nash Equilibrium were observed as well.

While the methods used in the studies above proved fruitful results, the research conducted in this study will attempt to use a cluster analysis to differentiate schools that recruit in distinct ways. After this step, regression and correlation analyses will be conducted to test if a relationship exists between certain recruiting clusters and team wins in the regular and NCAA tournament. The specific approaches are outlined in the following methods section.

Chapter Three: Methodology

Introduction

The online documentation of evaluating high school recruiting class potential (by companies like ESPN, 247Sports, Rivals.com) is something that has been happening in NCAA College Basketball for around twenty years. Prior research has been completed on the impact of a recruiting class, whether financially or on team success, but not one that identifies the NBA draft as one of the main variables that affect the way teams decide to recruit players and its effect on winning. This study aims to place D1 college basketball teams into specific categories, or clusters. Some of these clusters will include schools that target the type of player that is affected by the NBA's draft rules (where a player must be at least 19 years of age and one year removed from high school). (Giambalvo, 2018) After identifying these recruiting clusters, our analysis will determine which group has more success in the regular season and in the NCAA tournament by observing data from sportsreference.com. This chapter will detail the methods used to collect data, the testing methods used to analyze the data, and the programs used for data management.

Research Design

The research design used for this study is correlational and descriptive. Using this design allows this study to be descriptive of the relationship between variables, and a descriptive analysis of this relationship could provide benefits to many NCAA teams for years to come. Because there is only a limited amount of high-quality recruiting talent to go around, only the top NCAA basketball conferences which regularly compete for this talent will be tested in this analysis.

Sample

Recruiting information from ten Division 1 Athletic Conferences will be included in this study:

Atlantic Coast Conference (ACC), Southeastern Conference, (SEC) Big-Ten, Big 12, Pac-12, Big East, American, Atlantic 10 (A-10), West Coast Conference (WCC), Mountain West Conference (MWC)

Five conferences, known as the Power Five (ACC, SEC, Big-Ten, Big 12, and Pac 12), are included in this list. Additionally, the Big East is a conference that is commonly known in basketball circles as a highly competitive conference, as Villanova, Marquette, and Xavier regularly snag recruits inside the top 100 every season. Lastly, the remaining four conferences in our sample (American, A-10, WCC, and MWC) send multiple teams to the NCAA tournament in most seasons. It should be noted that during the time frame that is analyzed in this study, some of the conferences added and lost member teams. With that being said, each season will have a different number of total teams observed. Below is the total number of teams that will be observed, by year (N=718):

(2013-14: n = 116, 2014-15: n = 120, 2015-16: n = 120, 2016-17: n = 120, 2017-18: n = 121, 2018-19: n = 121)

This paper uses recruiting classes from almost every university in each of the ten conferences listed above to determine the recruiting cluster that a school falls into for a given season. The only school that is excluded from this study is the United States Air Force Academy, who competes in the Mountain West Conference. The authors of this study were unable to find complete information regarding school characteristics for the Air Force Academy, therefore they

will not be included in analysis. Each recruit in the sample is given a numerical score of 0-100 by ESPN, who is one of the industry leaders when it comes to evaluating high school recruits. The recruiting and regular season/NCAA tournament results data used in this study were collected from 2013-14 to 2018-19. The reasoning for pulling data from the last six seasons (2013-2019) happens because of a slight grading change that ESPN used before 2013. As an example of this change, this study will look at Marcus Paige, who played four seasons at the University of North Carolina. Before enrolling at UNC, Paige was assigned a numerical grade of 96 by ESPN in 2012 but only received a 4-star rating. It is evident that after the 2012 recruiting class, ESPN changed their rating scale criteria. In 2013, players who were assigned a value of 90 or above automatically received a grade of five stars. For example, Jabari Parker, a one-and-done and an eventual lottery pick in the NBA draft from Duke University, also received a score of 96 by ESPN but was awarded a five-star rating. If this study were to include ESPN's data from before 2013 in its sample, it could lead to inaccurate conclusions and a lack of reliability and validity. Another reasoning for using this time frame is the birth of the American Athletic Conference in 2013. Even though this conference is not in the Power Five, it still provides many teams that compete in the NCAA tournament on a yearly basis, so it would hurt this studies credibility to exclude it before its inception in 2013. This time frame contains 718 teams (over six years) and should be well equipped to show current trends in recruiting. Typical multivariate factor analysis recommendations generally expect a minimum sample size of 300 (Tabachnick & Fidell, 1996), and this study will also gain statistical power through continuous measures and repeated observations of teams. The regular season/NCAA tournament results for the same timeframe and conference selections will be applied to this study as well.

Treme and Burrus (2011) used a five-season (2004-to-2009) sample in "The impact of recruiting on NCAA basketball success." Additionally, they included data from the ACC, BIG EAST, BIG 12, Big 10, Pac 10, and SEC, plus six other teams from smaller conferences. In their second article about basketball recruiting, "NCAA basketball: when does recruiting talent translate into wins for power conferences?", Treme and Burrus (2015) reduced their sample to solely include the ACC, Big-12, Big East, Big Ten, Pac 12, and the SEC.

Instrumentation

The first-ever ESPN Top 100 ranking came out in 2007, with notable players like Kevin Love, O.J. Mayo, and Derrick Rose in its top five. (ESPN Basketball Recruiting, 2007) Ever since then, ESPN has churned out a database that aims to identify the top 100 high school seniors in the country every year. As stated above, ESPN grades each player from 0-100. Not only will this study be using the Top 100 data, but it will also be using ESPN's rating for every other player recruited by a team in one of the ten conferences observed in this study. This study will use R and SPSS to conduct different stages of its analysis. R is a free program used for statistical analysis and graphics building. (“What is R?,” n.d.) Analyzing each player's ratings for this study's 718 teams (the six season summation for the ten conferences) will require the use of a statistical analysis tool that can handle mass amounts of data. RStudio will be used for clustering and an OLS regression, while SPSS will be used for this study’s ordered logistic regression.

Data Collection Procedures

The regular season/NCAA Tournament success data for men's collegiate basketball will be pulled from sportsreference.com and inputted into Microsoft Excel 2016. Population of a school along with men's athletic recruiting expenditure will be collected from the U.S. Department of Education's EADA site, while population with 400 miles of a school will be pulled from StatsAmerica. This data will contain information regarding our sample team's success in every year's regular season and the years that they made the NCAA Tournament, along with other school characteristics. After this data has been successfully collected and inputted into Excel, it will be imported into a data frame in Rstudio and SPSS for further cleaning and analysis. A similar process will be used for the recruiting data from ESPN. This process will involve scraping each team's recruiting results (the numerical ranking for each incoming player) for the study's specific timeframe and inputting the results into Microsoft Excel. This portion of data collection will be done manually, but will include an accuracy audit to ensure accuracy. After this, this information will be inputted into Rstudio for analysis. Recruiting ratings of players and the number of wins by a team in a given year can be identified as continuous data. Lastly, the author of this study will be responsible for the data collection process. The R code used for data collection will be included in this study's appendix.

Data Management and Analysis

The data (recruiting and wins by a team) used in this analysis was inputted into Microsoft Excel and then read into an Rstudio terminal. Once in Rstudio, the data will be converted into a data frame. After the recruiting data has been cleaned and saved, it will be clustered into $\boldsymbol{N}$ specific groups by year using a K-means cluster analysis. Before running the K-Means cluster

analysis, the discrete measure of a schools number of players recruited at a five star level, first year (Freshman) players drafted into the NBA, and players drafted into the NBA after their Sophomore season (ex: Sophomore, Junior, Senior, etc.) were averaged together to assign one number to a school to account for the entirety of the six-season data. If a coaching change (hence a change in recruiting strategy) occurred during this time for a program, it was accounted for in the dataset as a separate observation for whatever school it happened at, i.e. each unit of observation was a Coach-School pairing. The R code used for data collection will be included in this study's appendix.

Table 2

Variables	Description
RegWinPerc	Dependent variable for equation one. Predicted percentage (continuous) of regular season wins over losses for a team.
TournamentPerformance	Dependent variable for equation two. Predicted number (continuous) of postseason appearances for a team.
PreviousYearWinPerc	Independent variable for both equations. Continuous measure for a team's winning percentage in the past season.
Football	Independent variable for both equations. If the school has a football team, this dummy variable will read "1", and "0" if not.
RecruitingCluster	Independent variable for both equations. Recruiting type (categorical dummy variable coded to "OneAndDone", "HighRecDev" or "PureDev", where "OneAndDone" indicates a 5 star recruiting strategy.
CoachTotalGames	Independent variable for both equations. Continuous measure for a coach's number of games.
CoachWinLossPerc	Independent variable for both equations. Continuous measure for a coach's W/L percentage.
SchoolPop1k	Independent variable for both equations. Continuous measure for a school's population.
Pop400mi1M	Independent variable for both equations. Continuous measure for a school's population within 400 miles.
MensAthleticRecruitingExp100k	Independent variable for both equations. Continuous measure for a school's men's recruiting budget.
PFIVE	Independent variable for both equations. Categorical dummy variable for whether a team is in a power five conference (1 if true, 0 if no).

Models

After each group is clustered together, an analysis will be done in order to see if there is a relationship between recruiting level and regular season (OLS Regression)/NCAA tournament appearances (Ordered Logit Regression) by a team in a given year.

Equation One: OLS Regression

$$RegWinPerc = B_0 + B_1 PreviousYearWinPerc + B_2 \text{CoachWinLossPerc} + B_3 CoachTotalGames + B_4 MensAthleticRecruitingExp100k + B_5 SchoolPop1k + B_6 Pop400mi1M + B_7 (MensAthleticRecruitingExp100k * RecruitingCluster) + B_8 Football + B_9 PFIVE + B_{10} RecruitingCluster + \varepsilon$$

Equation Two: Ordered Logit Regression

$$TournamentPerformance = B_0 + B_1 PreviousYearWinPerc + B_2 \text{CoachWinLossPerc} + B_3 CoachTotalGames + B_4 MensAthleticRecruitingExp100k + B_5 SchoolPop1k + B_6 Pop400mi1M + B_7 Football + B_8 PFIVE + B_9 RecruitingCluster + \varepsilon$$

Chapter Four: Results

The purpose of this paper is to explore whether there is a relationship between the different strategies (derived from past outcomes) used by college basketball teams while recruiting and how successful teams are after implementing their specific recruiting strategy. The analysis conducted in this study consisted of three steps, a K-Means Cluster analysis to identify recruiting clusters, an OLS Regression to predict regular season winning percentage, and an Ordered Logistic Regression to predict the number of NCAA tournament appearances a school accrued over the timeframe of the dataset.

Descriptive statistics were calculated and listed below for each of the continuous variables used in this study, as well as a count of each of the categorical variables used in analysis that contained either a yes or a no outcome (whether a university is in the power five or has a football team). There were a total of 718 teams used in the study's regression analyses and 121 teams (with six season averages for recruiting outcomes for each team) used in the K-means Cluster Analysis.

Table 3

Descriptive Statistics of Continuous Variables

Continuous Variables (Regression)	Mean	St. Dev	Min	Max
PreviousYearWinPerc	0.579663997	0.160459507	0.066666667	0.974358974
CoachTotalGames	405.0208914	291.6769512	0	1438
CoachWinLossPerc	0.576372052	0.159996752	0	0.86407767
MensAthleticRecruitingExp100k	883,932.18	595379.6595	110263	4346403
SchoolPop1k	17497.5376	10519.90256	1716	47253
Pop400mi1M	53135041.29	24897227.38	10881900	103010536
			n = 718	

Table 4

Count Statistics of Categorical Variables

Categorical Variables (Regression)	Count
OneAndDone	2
HighRecDev	23
PureDev	96
Total	**121**
Football	87
No Football	34
Total	**121**
Power Five	65
Non Power Five	56
Total	**121**

Table 5

Descriptive Statistics of Cluster Analysis Variables

Variables (Cluster Analysis)	HighRecDev	OneAndDone	PureDev	Overall
5Stars	0.523	3.917	0.049	0.216
1yeardraft	0.265	2.500	0.033	0.124
2-4yeardraft	0.723	0.667	0.111	0.252

K-Means Cluster Analysis

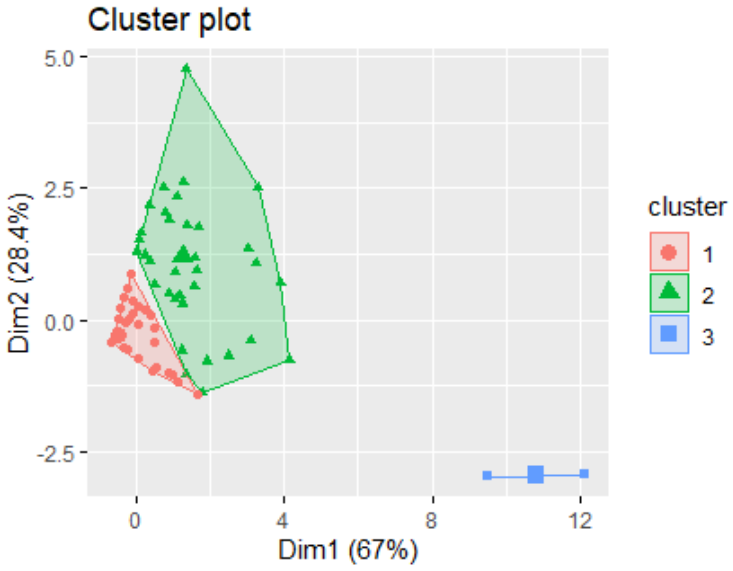
Using recruiting outcomes, the first test in this study set out to define three identifiable recruiting clusters across a six-season dataset (2013-14 – 2018-19) for ten conferences in Division One NCAA Men's College Basketball. The statistical test used to identify these clusters was a K-Means Cluster analysis that was conducted through the statistical software, R. Three categories were identified in this study as "OneAndDone", "HighRecDev", and "PureDev". "OneAndDone" refers to a school that consistently brings in players that are in the five star category (see ESPN's recruiting database). "HighRecDev" indicates a school that might sign players that fall into the five star category, but also relies on players that are rated at three and four stars to develop as a player before reaching their full potential. The final cluster, "PureDev"

denotes a school that almost exclusively relies on players that are ranked at three and four stars to fill out their rosters. This doesn't mean that these school don't want to recruit higher-ranked players, but most of the schools might not be able to do this consistently due to barriers like resources and brand image.

At the top recruiting outcome cluster, "OneAndDone" (or group 3), two schools (Duke and Kentucky) appeared due to their extreme ability to sign five star players. 42 schools appeared in the next category, "HighRecDev" (group 2). This group involves teams that might pull in some high level recruits but also tend to rely on retaining and developing players rather than getting five star talent and losing them after a year like the first group. Finally, the final cluster, "PureDev" (group 1), consisted of 163 observations that came from schools that did not recruit players at a high level, but rather relied on developing prospects that played more than one season at the college level before being drafted into the NBA. A graphed output is shown in Figure 3 to illustrate where each team fell in their respective cluster.

Figure 3

K-Means Cluster Analysis Chart



The results of this cluster analysis indicate that H0 (H0: Programs do not recruit significantly different across NCAA men's college basketball) should be rejected, and the alternate hypothesis that programs recruit differently across NCAA men's college basketball, should be viable based on the observation of differences in recruiting outcomes between schools.

OLS Regression

An OLS regression (multi-variate) analysis was conducted on the entire data set to test the relationship between recruiting cluster and team performance in the regular season, controlling for several other known/hypothesized performance predictors discussed above. The overall regression results and coefficient estimates are shown in table 6. The R-Squared associated with this multivariate OLS regression was 0.374. The rest of the variable coefficients can be seen below.

Table 6

	Dependent Variable:	RegWinPerc
Variable	β estimate	P Value
PreviousYearWinPerc	-.581	**< .001**
CoachWinLossPerc	.235	**< .001**
CoachTotalGames	.421	.523
MensAthleticRecruitingExp100k	<.001	.367
SchoolPop1k	<.001	.705
Pop400mi1M	<.001	.878
Football	<.001	.872
PFIVE	-.001	.249
OneAndDone	.798	.301
HighRecDev	.842	.265
PureDev	.766	.310
MensAthleticRecruitingExp100k:OneAndDone	<-.001	.400
MensAthleticRecruitingExp100k:HighRecDev	<-.001	.356
MensAthleticRecruitingExp100k:PureDev	<-.001	.371
R^2 = .374	Observations: 718	
F-statistic = 31.7 on 14 and 703 DF		

The dependent variable for this equation was identified as "RegWinPerc." Multiple independent variables were used in this step of analysis, including "PreviousYearWinPerc",

“CoachTotalGames”, “CoachWinLossPerc”, “SchoolPop1k”, “Pop400mi1M”, “MensAthleticRecruitingExp100k”, “MensAthleticRecruitingExp100k*RecruitingCluster”, “Football”, “PowereFive”, and “RecruitingCluster”.

Only two variables came back with statistically significant p values ($\alpha = .05$): PreviousYearWinPerc ($p < .001$) and CoachWinLossPerc ($p < .001$). “PreviousYearWinPerc” carried a coefficient of 0.4214 and “CoachWinLossPerc” had a coefficient of 0.235, which indicated that they both could be used as predictors of a team’s winning percentage in a given year and were both positively related to regular season winning performance. The three recruiting cluster variables generated by this study’s cluster analysis produced statistically insignificant p values of 0.301 (OneAndDone), 0.265 (HighRecDev), and 0.310 (PureDev). Therefore, there was not support for the second hypothesis (H2) that team recruiting clusters would be related to overall winning percentage, i.e, the null hypothesis was not rejected. An interaction variable was also created to interact a schools recruiting expenditure with their recruiting cluster (MensAthleticRecruitingExp * RecruitingCluster) as a separate variable for this linear regression model. This variable carried a p value that was greater than 0.05 for each recruiting cluster that recruiting expenses shared an interaction with, and is therefore not significant.

Ordered Logistic Regression

For the final test used in this study, an Ordered Logistic Regression was implemented to predict the number of tournament appearances a college basketball team would have in the NCAA Men’s Basketball Tournament. The statistical tool, SPSS, was used to conduct this step of analysis. The tournament outcome values ranged from zero to six. A value of zero would indicate a result of “NoBid.” If a team appeared in the tournament one or two times in a given

year, the result would read “FirstRound” or “SecondRound”, respectively. If a team made it to the third round (three appearances) in the NCAA tournament, the result would be “Sweet16”, indicating that the team was one of the final 16 teams remaining in the tournament. The fourth and fifth round of the tournament were coded as “Elite8” and “Final4”, and the sixth appearance of the tournament indicated that a team appeared in the National Title game (“ChampGame”).

The maximum likelihood estimation procedure produced a statistically significant model fit (Chi-Square = 265.4, $p < .001$). SPSS also produced three pseudo R-Squared values that assess the model’s ability to accurately predict the dependent variable (tournament appearances). The first was the Cox and Snell, which produced an R-Squared value of .309. The second was the Nagelkerke, which produced the highest pseudo R-Squared value of the three at .337. The last pseudo R-Squared value (the McFadden) was assessed at .149. The overall regression results are shown in Table 7.

Table 7

	Dependent Variable:	Tournament Performance
Variable	β estimate	P Value
0-1 Appearances	6.255	< .001
1-2 Appearances	7.195	< .001
2-3 Appearances	8.145	< .001
3-4 Appearances	9.046	< .001
4-5 Appearances	9.916	< .001
5-6 Appearances	10.685	< .001
PreviousYearWinPerc	5.736	**< .001**
CoachWinLossPerc	3.000	**< .001**
CoachTotalGames	0.000	.317
MensAthleticRecruitingExp100k	-0.002	.921
SchoolPop1k	0.004	.685
Pop400mi1M	0.008	**.012**
Football = 0	.137	.615
Football = 1	0^{α}	
PFIVE = 0	-0.555	.051
PFIVE = 1	0^{α}	
OneAndDone	1.505	**.012**

HighRecDev	.602	**.004**
PureDev	0^{α}	
$\boldsymbol{Cox\ and\ Snell\ R^2}$ = .309	McFadden $\boldsymbol{R^2}$ = .149	
Nagelkerke $\boldsymbol{R^2}$ = .337	Observations: 718	
Chi-Square = 265.356, p < .001		

Five variables came back with significant p values and coefficients that show that they are related to the likelihood of advancing through the NCAA Men's Basketball Tournament. Like our first model shown above, this model produced a significant response (p value <0.001) for a school previous year's winning percentage. This variable carried a positive coefficient of 5.736, which indicates a positive relationship shared with the dependent variable (TournamentPerformance). The variable for a coach's career winning percentage also provided a significant p value (< .001) and a positive coefficient (3.000) in relation to this model's dependent variable. The next independent variable that showed a significant response (p < .012) was a team's population within 400 miles of the school's location. This variable's coefficient was smaller, but still positive at .008. Two of the three recruiting clusters generated by the K-Means Cluster analysis also showed a significant response to the dependent variable. The cluster level "OneAndDone" provided a P value at < .012 with a coefficient estimate of 1.505, and the group "HighRecDev" accrued a P value of (p < .004) and a coefficient of .602. The final cluster was used as a reference condition for this model, and therefore didn't produce a P value or coefficient. The results of the cluster variables indicate support for the third hypothesis of this study (H3) that team recruitment clusters will relate to the number of appearances a team has in the NCAA tournament at a significant level, and the null hypothesis can be rejected.

While all of the significant variables are listed above, this step in the analysis did produce a variable (PowerFive) that was just above the study's significance threshold of .05.

“PowerFive” produced a P value of .050807 (> .05) which was very close to being seen as a significant predictor of appearances in the tournament and is worth noting, but can’t officially be documented as sharing a relationship with the dependent variable. The specific coefficient estimates and significance levels for each variable can be found above in Table 5.

Chapter Five: Discussion and Conclusion

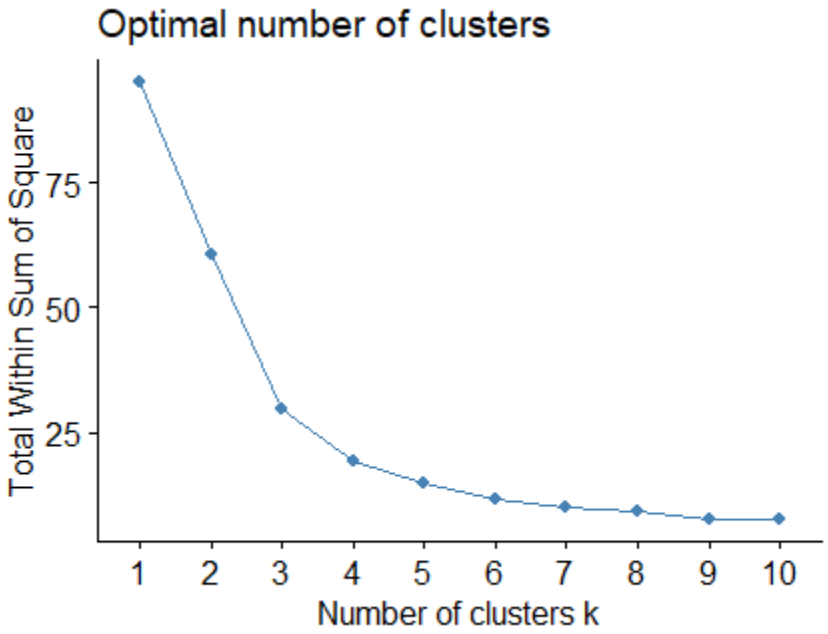
Chapter five discusses the results of whether there is a relationship between the different strategies (denoted by past recruiting outcomes) used by college basketball teams while recruiting and how successful teams are after implementing their specific recruiting strategy in terms of regular season winning percentage and NCAA Men's Basketball Tournament appearances. This chapter also contains a list of limitations for the study, recommendations for future research, and the study's conclusion.

Research Questions and Results

The first research question of this study asked whether it would be possible to reliably differentiate program recruiting strategies based on historical behavior. To answer this question, the analysis of this study first consisted of a K-Means Cluster analysis to identify whether or not schools fell into specific groups (or clusters) based on their past recruiting behavior. The schools were clustered into three groups (OneAndDone, HighRecDev, and PureDev) based on the number of five star recruits they signed, the number of players that were drafted into the NBA after their first season, and the number or players that were drafted after their second, third or fourth season in college. Based on the data being put into the cluster analysis, the decision to cluster on three groups (instead of two or four) was made based on the results of an elbow method test that gives a recommendation on the optimal number clusters. The y-axis measures the total within the sum of square of the data, and the x-axis represents the number of clusters. The elbow of the curve falls right above 3 (x-axis), which is why a three cluster strategy was chosen for this study (see below). After completing the K-Means Cluster analysis, three groups were assigned schools that fit the requirements of each cluster, confirming research question one.

Figure 4

Elbow Method Test



The second research question asks if it's clear that some programs rely specifically on multiple "one-and-done" players to come in annually. The results from the K-Mean's cluster analysis also confirms this, as two teams (Duke and Kentucky) fall directly into the "OneAndDone" category (see Chapter Four). Over the six year data set, Duke averaged 3.33 five star recruits signed per season, while Kentucky averaged 4.5 per season. It should also be said that while these teams were able to consistently sign high-level recruits every season, they won consistently while doing so (see below).

Table 8

Duke and Kentucky Win Totals from 2013-14 to 2018-19

Duke	**Wins**	**Losses**	**Kentucky**	**Wins**	**Losses**
2013-14	26	9	2013-14	29	11
2014-15	35	4	2014-15	38	1
2015-16	25	11	2015-16	27	9
2016-17	28	9	2016-17	32	6
2017-18	29	8	2017-18	26	11
2018-19	32	6	2018-19	30	7
Total	**175**	**47**	**Total**	**182**	**45**

The final research question asks if team recruitment strategies (past outcomes) will make a team more or less likely to win at the national level (regular season and NCAA Men's Basketball tournament) with respect to other institutional characteristics such as university population, population within 400 miles, coaching characteristics (winning % and total games for a coach), and a schools men's athletic recruiting budget. The other institutional characteristics that were controlled for were if the school was in the Power Five (ACC, SEC, BIG Ten, Big-12, or Pac-12) and if the school had a men's football program.

Two sets of analysis were done in ordered to attempt to answer this research question. The first was an OLS regression analysis that aimed to predict a men's basketball program's winning percentage based on the variables listed above along with interacting which recruiting cluster the school fell into with the men's athletic recruiting expenditure for a school. After completing this analysis, only two variables came back as significant predictors of a team's regular season winning percentage. Of the two variables that came back as significant predictors of a team's winning percentage, only one was aligned with the third research question listed above (CoachWinLossPerc). This variable carried a significant P value (< .01) and a coefficient of 0.235, indicating that it affects a team's winning percentage in a positive way.

The model as a whole did not provide enough statistically significant variables or a reasonable enough R-square value to predict the dependent variable with precision. The three recruiting cluster variables did not return with statistically significant results, but it should be noted that recruiting rankings are a measurement that has significant noise due to how it is assessed to each player by an evaluator based on a prediction of a players potential. Therefore, this tool should not be used to answer regular season portion of this research question.

For the second portion of the third research question (pertaining to success in the NCAA tournament), this study used an Ordered Logistic Regression to predict the number of NCAA Basketball Tournament appearances (TournamentPerformance) by a school. The variables that were significant in this model were “CoachWinLossPerc”, “Pop400Mi1M”, “OneAndDone”, and “HighRecDev”. All four of these variables produced positive coefficients that effect tournament appearances in a positive way. “CoachWinLossPerc” (coefficient: 3.00) and “Pop400Mi1M” (coefficient: .008) both fall under the category of institutional characteristic that align with this research question. The final two significant variables were generated by the K-Means Cluster analysis and both fall under the umbrella of recruiting strategies that impact national success (appearances) in the NCAA Men’s Basketball Tournament. The group “OneAndDone” had the greatest positive effect on the dependent variable at 1.505, and “HighRecDev” carried a positive coefficient of .602. The previous statement aligns with what "The impact of recruiting on NCAA basketball success" (Treme and Burrus, 2011) says about the impact of varying player experience with relation to the NCAA tournament. The analysis conducted by the two authors’ state that having more experienced players in the NCAA can lead to more success, which aligns with the “HighRecDev” cluster in this study’s analysis.

While our model produced significant results that impact national success in the tournament in a positive way, it should not be used as a reliable tool to predict appearances. However, the significant variables do align with this research question and are positively related to adding NCAA Tournament appearances for teams.

Limitations

After completing the analysis, one of the biggest limitations of this study was its sample size. While six years and 718 team observations across ten conferences is typically enough for multivariate and repeated measure (panel regression) techniques, there were limitations here in how many observations occurred within the identified clusters within the sample, specifically the "One and done" cluster. This smallest group only included two schools and twelve observations (six seasons), which limited relevant model identification for both panel regression and many independent factor interactions, e.g. cluster X institutional characteristics. However, this timeframe was chosen because of the fact that "one-and-done" players departing after one season wasn't as common of a trend until the early 2010's. Additionally, the American Athletic conference didn't exist until 2013, which was another reason for starting the data collection process during the 2013-14 season.

Another limitation of this study comes from the sample size, but in a different manner than the limitation listed above. Because resource levels fluctuate so much between schools and conferences, having smaller conferences in the data that didn't match up as well with larger ones negatively affected each models fit. With that being said, it would have been difficult to justify excluding the West Coast Conference from the dataset when one of their best schools, Gonzaga, is consistently nationally ranked and appears in the NCAA Tournament every year (Gonzaga placed as National Runner-Up during the 2016-17 season).

The final limitations listed in this section relates to the population data and cluster analysis results that appear in each of this study's regression analyses. The issue with this data was that it did not match year-to-year population levels with each season but instead had to use one population number (for university population and population within 400 miles of a school) for the entire dataset. Regarding the data used in the K-Means Cluster analysis, this study used the average number of five star recruits, first year draftees, and second through fourth year draftees for a school instead of a single-season number of recruits or draftees used in a cluster based on each year. While this can offer a picture of general recruiting outcomes across the time period, it could be diluting the impact of recruiting efforts in any single season (or smaller time frame).

Recommendations for Future Research

In terms of recommendations for future research relating to college basketball recruiting and team success, one thing that could widen the scope of this study's topic would be incorporating transfer portal data. This could be done to measure whether schools benefit more from recruiting high school players or former high quality recruits who decided to transfer elsewhere at some point during their college basketball career. Another area of improvement that could be made would be to expand this dataset by another two to three years. The one-and-done player and his effect on college teams is more important now more than ever, and having more recent data added to this study's dataset would only help a future studies credibility.

Conclusion

In conclusion, Division One NCAA Men's College basketball programs do recruit differently. While this isn't as much of a known strategy but more of an analysis on past recruiting outcomes, there are significant differences between schools that consistently bring in

one-and-done talent and those that do not. The analysis done in the study doesn't demonstrate a strong relationship with recruiting clustering and success in the regular season, but recruiting clustering is positively related with appearances in the NCAA Tournament on the levels of "one-and-done" and "high-recruiting/development" along with other institutional characteristics like coaching quality (winning percentage) and a school's population within 400 miles. Therefore, if a program's objective is to advance in the NCAA tournament more regularly, a "professional development" and "one and done" generally produces better results than a developmental approach. If a programs objective in regular season wins, there doesn't seem to be a difference across recruitment patterns based on this sample.

Appendix

R Code from K Means Cluster Analysis

```
#Reading Data
library(readxl)
avgsforclustering <- read_excel("C:/Users/Samuel/Desktop/Thesis/avgsforclustering.xlsx")
View(avgsforclustering)

#Calling Packages
library(tidyverse)
library(cluster)
library(factoextra)

#DF Conversion
complete_df <- avgsforclustering[complete.cases(avgsforclustering),]
sum(sapply(complete_df, is.infinite))

avgsforclustering[apply(sapply(avgsforclustering, is.finite), 1, all),]

#viewdata
View(sapply(avgsforclustering, class))

#deleting the name columnn
clustering <- avgsforclustering[,-1]

#3 clusters
kcluster <- kmeans(clustering, centers = 3, nstart = 25)
str(kcluster)

#Cluster Plot
fviz_cluster(kcluster, data = clustering, geom="point")
```

R Code from OLS Regression

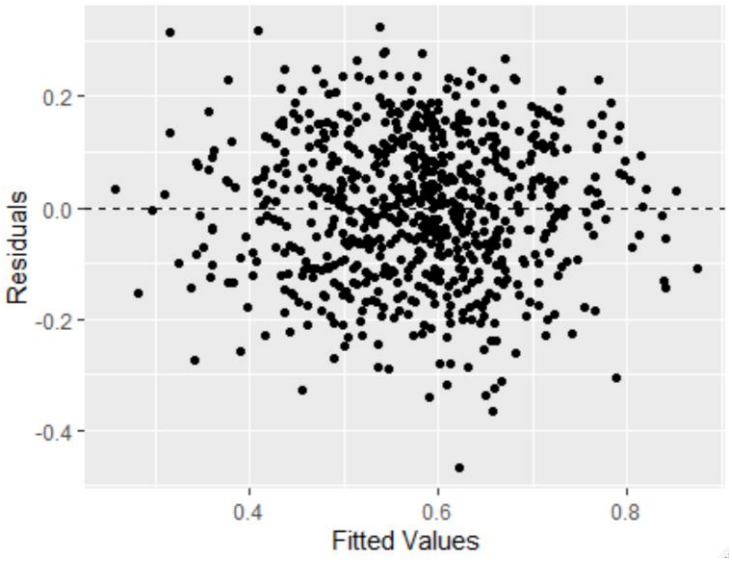
```
#Call package
library(tidyverse)

#OLS with LM
modelOLS1 <- lm(RegWinPerc ~ PreviousYearWinPerc + CoachTotalGames +
CoachWinLossPerc +
                  (MensAthleticRecruitingExp100k*RecruitingCluster) +
                MensAthleticRecruitingExp100k + SchoolPop1k + Pop400mi1M +
                  Football + PowerFive + RecruitingCluster, data = Thesis_Data_1_)

#model summary
summary(modelOLS1)
```

```
#plot
ggplot(data = modelOLS1, aes(x = .fitted, y = .resid))+
  geom_jitter()+
  geom_hline(yintercept = 0, linetype = "dashed")+
  xlab("Fitted Values")+
  ylab("Residuals")
```

Scatter Plot from OLS Regression



References

2019 NBA Draft results: Picks 1-60. (2019, June 21). Retrieved February 13, 2020, from https://www.nba.com/article/2019/06/21/2019-nba-draft-results-picks-1-60

Borghesi, R. (2015). The Financial and Competitive Value of NCAA Basketball Recruits. Retrieved February 28, 2020

Burleyson, N. (2018, July 11). The History of NBA Draft Eligibility and the Elimination of the One-and-Done Rule. Retrieved February 11, 2020, from https://medium.com/@NBurleyson/the-history-of-nba-draft-eligibility-and-the-elimination-of-the-one-and-done-rule-239ba0c7e21c

Caro, C. A. (2012). College Football Success: The Relationship between Recruiting and Winning. International Journal of Sports Science & Coaching, 7(1), 139–152. https://doi.org/10.1260/1747-9541.7.1.139

Chatterjee, K., & Samuelson, William. (2014). Game theory and business applications / Kalyan Chatterjee, William Samuelson, editors. (2nd ed., International series in operations research & management science ; 194). New York: Springer.

ESPN Basketball Recruiting - Player Rankings. (2007). Retrieved March 23, 2020, from http://www.espn.com/collegesports/basketball/recruiting/playerrankings/_/class/2007/order/true

ESPN Basketball Recruiting - Player Rankings. (2018) Retrieved February 13, 2020, from http://www.espn.com/college-sports/basketball/recruiting/playerrankings/_/class/2018/order/true

ESPN. (2021). College Basketball Top 100 [Player Recruiting Rankings from 2013 to 2019] . Retrieved from http://www.espn.com/college-sports/basketball/recruiting/playerrankings

Evans, B., & Pitts, J. (2017). The determinants of NCAA basketball recruiting outcomes. Applied Economics Letters, 24(5), 351-354.

Evans, B. A., & Pitts, J. D. (2018). Cross-Sport Recruiting Effects in NCAA D1 Football and Basketball. Retrieved February 28, 2020

Fominykh, K. (2019, August 20). NCAA will allow undrafted basketball players to return to their schools, 'elite' players to have agents. Retrieved December 11, 2020, from https://www.baltimoresun.com/sports/college/bs-sp-ncaa-new-rules-20180808-story.html

Fort, R., Quirk, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. Journal of Economic Literature, 33(3), business premium collection, 1265-1299. DI Men's Basketball Championship History. (2019, April 8). Retrieved February 15, 2020, from https://www.ncaa.com/history/basketball-men/d1

Giambalvo, E. (2018, August 8). NCAA to allow more flexibility for college basketball players considering the NBA. Chicago Tribune. Retrieved from https://www.chicagotribune.com/sports/college/ct-spt-ncaa-nba-flexibility-20180808-story.html

Grading Scale. (2006, October 26). Retrieved February 13, 2020, from https://www.espn.com/college-sports/recruiting/basketball/mens/news/story?id=2639514

How The MLB Draft Works. (2008, May 23). Retrieved February 11, 2020, from https://www.baseballamerica.com/stories/how-the-mlb-draft-works-2008/

Langelett, G. (2003). The Relationship between Recruiting and Team Performance in Division 1A College Football. Journal of Sports Economics, 4(3), 240–245.

Maschler, M., Solan, E., & Zamir, S. (2013). Game Theory. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511794216

NBA.com. (2017, August 24). All-Time NBA Draft History. Retrieved April 29, 2020, from https://www.nba.com/history/draft

NCAA College Basketball Conferences. (n.d.). Retrieved February 15, 2020, from http://www.espn.com/mens-college-basketball/conferences

NHL Draft Rules - NHL. (n.d.). Retrieved February 11, 2020, from https://www.draftsite.com/nhl/rules/

Smith, J. (2019, April 23). NBA Draft Eligibility Rules: Deadline to Withdraw & Upcoming Changes. Retrieved December 11, 2020, from https://heavy.com/sports/2019/04/nba-draft-eligibility-rules-deadline-to-withdraw-enter/

Sports Reference. (2021). College Basketball Reference [College Basketball Wins and Losses for Coaches and Teams]. Retrieved from https://www.sports-reference.com/cbb/

Stats America. (2021). Big Radius Tool [Population within 400 miles of a University]. Retrieved from https://www.statsamerica.org/radius/big.aspx

The Rules of the NFL Draft. (n.d.). Retrieved February 11, 2020, from https://operations.nfl.com/the-players/the-nfl-draft/the-rules-of-the-draft/

Tabachnick, B. G., & Fidell, L. S. (1996). Using multivariate statistics (3rd ed.). New York: HarperCollins.

Treme, J., R. Burrus, and B. Sherrick. 2011. "The Impact of Recruiting on NCAA Basketball Success." Applied Economics Letters 18: 795–798. doi:10.1080/13504851.2010.507171.

Treme, J., Burrus, 2015. NCAA basketball: when does recruiting talent translate into wins for power conferences?. J Econ Finan 40, 735–753 (2016).

US Department of Education. (2021). EADA [Men's College Recruiting Expenditures from 2013-2019 & School Populations]. Retrieved from https://ope.ed.gov/athletics/#/

Wilco, D. (2020, January 26). What is March Madness: The NCAA tournament explained. Retrieved February 15, 2020, from https://www.ncaa.com/news/basketball-men/bracketiq/2020-01-26/what-march-madness-ncaa-tournament-explained

“What is R?” (n.d.). Retrieved March 23, 2020, from https://www.r-project.org/about.html