# With the First Pick in the NFL Draft, Tversky and Kahneman Select…

## Irrationality in the NFL Draft as a Manifestation of Cognitive Biases, 1967-2013

An honors thesis for the Department of Economics

Michael Roubey

Advisors: David Garman, Department of Economics
Ayanna Thomas, Department of Psychology

Tufts University, May 2014

## Abstract

This thesis analyzes three hypotheses of ways in which the use of cognitive heuristics could lead to suboptimal decision making in the National Football League (NFL) draft. In the first hypothesis, teams draft more players from schools with high-caliber players due to the availability heuristic. This bias is found and is shown to be detrimental to teams. The second hypothesis contends that the availability heuristic will lead teams to overdraft from their local state. Analysis shows that between a fourth and a third of NFL teams draft local players too often, but this does not have any negative effect on performance. The final hypothesis argues that teams will use a mental representation of an NFL player when drafting, even if the mental representation is flawed. This hypothesis is confirmed for quarterbacks, where teams miss a key signal, but is found to be false for offensive skill players, who teams draft based on the correct cues. Prescriptive recommendations are outlined on how teams can avoid biased decision-making processes in the future.

## Acknowledgments

I would like to thank my committee chair, David Garman, and committee member, Ayanna Thomas, for their help and guidance throughout this project. I would not have been able to complete this project without them. I am grateful to the Office of Undergraduate Education at Tufts University and the Undergraduate Research Fund for their financial support. Thank you to Nate Tarrh for his role in data collection. Finally, thank you to my friends and family for their love, support, and interest in my project.

# Table of Contents

## 1. Introduction

In 2011, the National Football League introduced a rookie wage scale as part of their new Collective Bargaining Agreement (Clayton, 2011). This addition systematized the dollar amount and length of a rookie's contract, depending on where they were drafted. The scale dramatically reduced rookie salaries, particularly those of players drafted early. For example, the first overall pick of 2010, Sam Bradford, signed a six-year, $78 million contract with the St. Louis Rams (Spotrac). The following year, the first overall pick, Cam Newton, signed a four-year contract with the Carolina Panthers for just over $22 million (Spotrac).

Though the draft has always been an important event for teams, the new rookie wage scale made it more important than ever. The draft now offers teams the chance to acquire talented players at suppressed rates. Such players are needed because of the hard salary cap, which limits the spending of teams. A team that consistently drafts better than others could gain a competitive advantage over the rest of the league because they would have more talented young players at low costs.

This point is moot if the league is already drafting optimally. It is clear, however, that teams are not consistently successful at drafting. Every year, early-round picks turn into busts, while late-round players turn into steals despite being passed over by other clubs. One strong example that highlights these failings compares the number one overall draft pick in the 2007 NFL draft and the 199$^{th}$ player taken in the 2000 draft. One went on to win three Super Bowls and two league Most Valuable Player (MVP) awards (Pro-Football-Reference). The other played only three seasons in the NFL, completing only 52% of his passes and throwing more interceptions than touchdowns (Pro-Football-Reference). One would naturally assume that the 1$^{st}$ overall pick went on to have the storied career while the late-round pick fizzled and eventually

**Table 1** Descriptive Player Statistics by Round

| *Round* | *Primary Starter* | *Games Played* | *Average Approximate Value Per Year* |
|---|---|---|---|
| *1* | .89 | 99.28 | 5.19 |
| *2* | .72 | 79.33 | 3.71 |
| *3* | .56 | 64.08 | 2.70 |
| *4* | .45 | 52.50 | 2.19 |
| *5* | .34 | 42.11 | 1.65 |
| *6* | .26 | 34.30 | 1.29 |
| *7* | .23 | 30.57 | 1.16 |

left the league. The reverse, however, is true. Jamarcus Russell, the 1st overall pick, is now out of the league while Tom Brady, the 199th pick, continues to excel (Pro-Football-Reference). The important thing to note is this situation is not an isolated incident. The NFL is littered with successful late-round picks while many recent high picks are already out of work.

It would be wrong, however, to say that the draft is entirely random. On the whole, the NFL is successful at drafting better players earlier in the draft. The earlier a player is drafted, the more likely he is to become a primary starter for a team and play more games in the NFL. Earlier draft picks also provide greater value to their respective teams. These data are shown in Table 1. This illustrates that talent evaluation is not random and when teams do evaluate talent correctly, performance benefits. Though there is a steady relationship between draft position and future performance, there remains room for team improvement throughout the draft, particularly in the later rounds.

Given that draft outcomes are not entirely random, suboptimal drafting does exist, and players drafted are now paid at reduced rates, an opening exists for teams to excel through improved drafting. There are two distinct ways that a team could achieve this. One is gaining new knowledge about how players will perform in the NFL. Statisticians are currently using advanced analytics to attempt to find predictors of how a player will perform in the NFL in their college statistics, combine performance, and measurements. This information would reduce the risk of draft picks and improve a team's decision making. The second method also improves team decision making, but without the necessity of gaining new information. Instead, a team can outperform the league by avoiding suboptimal decisions that the rest of the league consistently makes. The field of behavioral economics has found that humans have biased decision-making processes due to the use of heuristics, problem-solving short cuts. By avoiding systematic poor decisions, a team could perform better in the draft than other teams without gaining novel information.

This thesis examines the latter method of improving drafting. Specifically, this thesis analyzes three distinct hypotheses of ways in which cognitive biases could manifest themselves as an irrational drafting strategy. The three hypotheses will be called the Talented Teammate hypothesis, the Geographic Bias hypothesis, and the Signal and Noise hypothesis.

To analyze these hypotheses, a review of previous literature on cognitive biases and heuristics, drafting in professional sports, and the interaction between the two fields is presented. Then, each hypothesis is examined individually. Each hypothesis is described in detail from a psychological perspective. Next, the data used to test the hypothesis are discussed. The models and tests used for the analysis are presented in the Methodology section. The results from these models are presented and discussed. Finally, there is a general discussion of the results spanning across all three hypotheses. This thesis ends with concluding remarks.

## 2. Literature Review

There has been significant previous research done on heuristics and on the NFL draft. Unfortunately, there has been very little done on the interaction of these two fields. Each of these topics is introduced below. This section concludes with a description of what this thesis adds to the literature.

In one of their seminal papers, Tversky and Kahneman (1974) explored three ways in which humans may err in decision making under uncertain conditions. These errors arise via the use of heuristics or mental shortcuts. Two of the heuristics discussed in the article are relevant for this thesis. The first is the availability heuristic, defined as "situations in which people assess the frequency of a class or probability of an event by the ease with which instances or occurrences can be brought to mind." This bias can manifest itself in several ways. Two of note are "biases due to the retrievability of instances" and "illusory correlation". The former is when some aspect of a class, such as size or quality, appears larger due to the ease with which members of the class are retrievable. The latter occurs when individuals incorrectly judge the frequency of two events co-occurring. Both of these manifestations can lead to errors in decision making, both generally and within the domain of sports.

The second potential bias is the representativeness heuristic. Tversky and Kahneman (1974) define this heuristic as the process "in which probabilities are evaluated by the degree to which A is representative of B, that is, by the degree to which A resembles B." As with availability, this can lead to erroneous conclusions in multiple ways. One important example is "insensitivity to predictability". In this case, individuals predict future outcomes based on descriptions. In doing so, they ignore the reliability of the description and the probability of an accurate prediction. Additionally, there is a key potential outcome of this bias, "the illusion of validity". Even when confronted with the limiting factors of representativeness, including base rates, description errors, and faulty predictions, individuals will still feel confident in their predictions. As with availability, this general bias could manifest itself within a team's decision making.

In the years since Tversky and Kahneman first named these heuristics, the field has developed and generated the concept of a dual-processing model (Stanovich, 2012). This model contends that there are two systems for making decisions, System 1 which is based on intuition and heuristics and System 2 which is based on rationality and statistical analysis. System 1 is automatic and always working, where System 2 requires additional effort. Under this model, a person can improve their decision making by using System 2, but at the cost of spending additional mental resources. There are times, however, when this is well worth the cost, and high-profile decisions in sports are likely one of those cases.

One important issue that arises with the topic of heuristics is how it interacts with expertise. The results are decidedly mixed on this topic. Studies have founds that experts use more heuristics than their untrained counterparts (Pachur & Marinello, 2013). The argument for this finding is that experts' skill and knowledge allow them to draw conclusions using smaller amounts of data. The novices, on the other hand, require more information before they are able to decide. Though experts use these shortcuts more often, it is not clear that they are effective in doing so. In fact, research shows that experts perform better when they use an analytical approach (Pretz, 2008). Together, these findings suggest that though experts gain either the ability or confidence to use heuristics in complex environments, they still make stronger decisions when they avoid heuristics and use a System 2 approach. In fact, all people, whether experts or not, make better strategic decisions when they spend more time on a decision (Moxley, Ericsson, Charness, & Krampe, 2012). Thus, expertise does not ensure that a person is protected from biased decision making. Experts are just as capable as novices, and potentially more so, of over-relying on heuristics to a detrimental degree.

Given that heuristics can still bias decisions in specific expert-based fields, the central question of this thesis remains valid. Consequently, it is important to understand the field in question and how decision-makers act.

The NFL draft is an annual event in which the current NFL franchises gain exclusive rights to new league players. The rules governing the draft are outline in the NFL Constitution and Bylaws (2006) and the Collective Bargaining Agreement (2011) between the NFL and the NFL Players Association. The draft takes place during the offseason, after the Super Bowl but before the start of training camp. The length of the draft has varied over the years, ranging from 17 rounds to the current 7 rounds (Pro-Football-Reference). The order of selections is determined by the previous year's standings, whereby the team with the worst record gets the first pick in each round and the Super Bowl winner gets the last pick in each round. A list of procedures is in place in case of a tie between teams. In addition to a team having one pick in each round, there are compensatory picks assigned in each of the rounds 3-7 to teams who lost a player to free agency the previous year. Teams are allowed to trade current and future draft picks. Though it was already generally the practice to pay players less money as the draft progressed, a rookie wage scale is now in place so that players drafted later in the draft necessarily receive less money than earlier draft picks.

There has been a great deal of analytic work done with regards to the NFL draft. There are two basic areas of this research, predicting draft position and predicting future performance. These fields can also be combined to study how successful teams are at drafting.

There are two main data sets that economists use when analyzing the draft. The first is the NFL Combine. The combine is an event held each year before the draft where hundreds of draft-eligible players come to be officially measured, both in terms of physical traits such as height and weight and drills such as the bench press and 40-yard dash (McGee & Burkett, 2003). When accounting for position, the physical results of the combine can be extremely predictive in regards to draft position (McGee & Burkett, 2003). The same cannot be said, however, of the psychological components. At each combine, draft-eligible players are required to take the Wonderlic Personnel Test to measure their intelligence (Mirabile, 2008). This is thought to be most important for quarterbacks due to the cognitive nature of their position. However, Wonderlic scores offer no reflection of how a quarterback performed in college or where they are drafted in the NFL (Mirabile, 2008).

The other dataset of note is a player's college statistics. While it stands to reason that performance at one level should indicate performance at another level, this is not necessarily true in football given the significant differences between the college and professional games. Much of the research done in this area focuses on quarterbacks. Quinn, Geier, and Berkovitz (2007) find that quarterbacks are drafted due more to athleticism than to college success. They suggest that this may be a flawed system, however, because highly drafted quarterbacks do not play systematically better than lower rounds draft picks who get significant playing time. This result, however, must be questioned because the authors do not take into account the different relationship between quality of play and playing time for high-round and low-round players. For the former, they are generally given playing time from the start and lose it only if they fail. Conversely, a member of the latter group plays only if he proves his quality. Berri and Simmons (2011), who use an approach similar to this thesis's Signal and Noise hypothesis, find combine results a stronger predictor of a quarterback's draft position than college statistics. These combine factors, however, are not predictive of future performance, whereas the previously ignored college completion percentage is indicative of future success.

Collectively, the research shows that it is extremely difficult to predict future performance and that teams are drafting in a suboptimal fashion. It is important to note, though, that factors other than statistics and measurements can affect draft decisions. In a paper examining the Major League Baseball (MLB) draft, Whitaker (2013) finds a phenomenon he calls the coattail effect in which teammates of highly drafted players fare better due to the presence of a talented teammate. A school has significantly more players drafted in a year where they have a player taken in the first five rounds of the draft. This trend, which clearly is not grounded in statistical reasoning, suggests teams are influenced by other factors, which could potentially be cognitive heuristics. The Talented Teammate hypothesis of this thesis is derived from Whitaker (2013).

Though both heuristics and the NFL draft have been studied to a great degree, only two studies have investigated their interaction. The first, Hendricks, DeBrock, and Koenker (2003), considers the draft from a labor perspective. They find that statistical discrimination exists in regards to what level school a player attended: either Division 1-A or lower (Division 1-AA, Division II, Division III). Early in the draft, teams use a risk-averse strategy and draft players from well-known schools in an attempt to avoid uncertainty. This suggests that teams know they are supposed to gain value from these picks and thus tread cautiously. In later rounds, however, the reverse is true. Teams now prefer players from smaller schools. Given that these rounds are not expected to produce high quality players, teams choose the riskier options due to the option value they possess. Now that the cost of drafting a player is lower, they are willing to take the chance and evaluate the player to determine their productivity. Collectively, these results show that teams are allowing their perceptions of each round to affect their drafting strategy. Players are viewed more or less favorably depending on factors outside of themselves, indicating the use of heuristics.

The second paper examines overconfidence and its effect on the efficiency of the NFL draft (Massey & Thaler, 2005). The methodology used involves assigning a surplus value to each pick in the draft by calculating his performance value and subtracting his salary. The authors find that surplus peaks in the second round, suggesting that teams overvalue earlier picks. Several psychological factors lead to this result. Teams in general are overconfident about their ability to accurately predict future performance. Therefore, they are willing to spend more money on newly drafted players than is logically advisable. This also derives from non-regressive predictions, whereby teams estimate unrealistic future performance from their drafted players by disregarding base rates. These factors, along with the competitive setting, lead to a winner's curse phenomenon in which a team that is willing to spend the most and draft a player with an early pick is likely to be overvaluing that player. The final psychological factor is the false consensus effect, in which a team thinks other teams agree with their scouting report of a player. While consensus is likely true at the top of the draft, this bias leads teams to draft players too early in the middle and late rounds. All told, this combination of psychological biases lead teams to overvalue high draft picks and draft in inefficient ways. While Massy and Thaler (2005) bring up several points, there is a clear flaw with their surplus definition. Performance per dollar is important, but it is not the only important factor. Given that only eleven players are on the field at a time, concentrating high levels of talent in a few players is more beneficial to a team than have a large number of good, but not great, cheap players.

Clearly, the overlap between cognitive biases and the NFL draft is an area ripe with potential findings. This thesis positions itself in that literature and looks to add to the current body of literature in three key ways. First, it presents three new hypotheses in line with the work done by Hendricks et al. (2003) and Massey and Thaler (2005). By examining how psychological trends affect drafting decisions, this thesis seeks to add to the understanding of how teams make decisions and how heuristics are used in specific fields. Second, the thesis takes analyses formerly done in a purely analytical fashion (Whitaker, 2013; Berri & Simmons, 2011) and provides a new explanation for the results found. This suggests that although only a select few studies highlight psychology's role in drafting biases, many more papers may show a similar effect but fail to recognize it. Third and finally, this thesis provides prescriptive advice to teams, scouts, and general managers on how to avoid using heuristics. General principles found in lab experiments are translated to systems that teams can adopt to help their football operations department make better decisions.

## 3. Hypothesis 1 – Talented Teammate

### 3.1. Hypothesis

The Talented Teammate hypothesis is that schools will have more total players drafted in years when they have a high draft pick than in years where no such player exist. This possibility is based on the Availability Heuristic and occurs through two channels. When a team is preparing for the draft, time is a scarce resource. General managers and scouts must decide how much time to allocate to each draft-eligible player. The most logical allocation correlates time spent scouting a player with the utility that player would offer the team. Therefore, early-round picks are likely to be scouted more in depth than lower-round selections. This allocation system naturally leads to an illusory correlation between time spent scouting a player and said player's talent level.

While this effect will often lead scouts to make rational decisions, it can err because football players are rarely scouted in isolation. Though players can be viewed individually at workouts and at the NFL Combine, the majority of scouting involves watching game tape. This tape, by nature, includes the teammates of the targeted player. So, a scout spending a large amount of time on a high-round prospect will inadvertently spend that same amount of time watching his teammates. When the time comes later for that scout to evaluate the lower-quality players on that team, whether for a scouting report or during the draft itself, the illusory correlation between familiarity and talent will lead the scout to overestimate the talent level. Another effect will also occur due to the additional scouting. Beyond the illusory correlation, the availability heuristic will cause a scout to like a player more simply because he is more familiar with that player. Therefore, players who are potential low-round picks or might go undrafted are more likely to be drafted if they have a very talented teammate because the additional scouting time will make them seem both more talented and more likeable.

There are three steps to analyzing this hypothesis. First, regression analyses will be performed to determine if schools have more players drafted in years with a high-round pick than in years without such a pick. Second, regression analyses will be performed to determine if players drafted in years with a highly drafted teammate underperform relative to other players selected near them. Finally, suggestions will be made to help NFL teams avoid the potential adverse effects of this bias.

## 3.2. Data

The data used in this analysis come from Pro-Football-Reference. The data consist of the top 222 players drafted into the NFL each year in the Super Bowl era, 1967 to 2013. This number of picks was selected because it was the smallest number of players drafted in any one draft. The analysis of this hypothesis is broken down into two parts: School and Player.

### *School Data*

The number of players drafted from a school each year was tallied from the original player data. Overall, 400 colleges and universities are included in the data set. There were also players who could not be tied to a specific university. Over the course of the 47-year period, therefore, there are a total of 18,847 observations. For each school in each year, the total number of players drafted in high rounds and in non-high rounds was tabulated. These calculations were done for three different definitions of high round: round 1, rounds 1 and 2, or rounds 1, 2, and 3. On average, a school had 0.55 players drafted each year. This broke down into 0.07 in the first round and 0.48 in subsequent rounds, 0.15 in the first two rounds and 0.41 in subsequent rounds, and 0.22 in the first three rounds and 0.33 in subsequent rounds. A dummy variable, *Talented Teammate*, was created that had a value of 1 if a school had a high-round draft selection in the given year. 5% of observations had a *Talented Teammate* value of 1 when high round was defined as the first round. This number rose to 10% when the second round was included and 15% when counting rounds 1-3. Table 2 contains the summary of these school data.

School data were also divided between offense and defense. For each year and school combination, two observations were created – one for offense players drafted from the school and one for defense players drafted from the school. All the same metrics were included in these observations. These results are included in Table 3.

**Table 2** Descriptive Statistics of School Data

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Year* | 18847 | 1990 | 13.57 | 1967 | 2013 |
| *Total Drafted* | 18847 | 0.55 | 1.27 | 0 | 13 |
| *High-Round Players (1)* | 18847 | 0.07 | 0.34 | 0 | 6 |
| *Non-High-Round Players (1)* | 18847 | 0.48 | 1.09 | 0 | 11 |
| *Talented Teammate (1)* | 18847 | 0.05 | 0.23 | 0 | 1 |
| *High-Round Players (1-2)* | 18847 | 0.15 | 0.52 | 0 | 7 |
| *Non-High-Round Players (1-2)* | 18847 | 0.41 | 0.93 | 0 | 11 |
| *Talented Teammate (1-2)* | 18847 | 0.10 | 0.30 | 0 | 1 |
| *High-Round Players (1-3)* | 18847 | 0.22 | 0.69 | 0 | 8 |
| *Non-High-Round Players (1-3)* | 18847 | 0.33 | 0.79 | 0 | 9 |
| *Talented Teammate (1-3)* | 18847 | 0.13 | 0.34 | 0 | 1 |

**Table 3** Descriptive Statistics of Offense-Defense School Data

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Year* | 37694 | 1990 | 13.56 | 1967 | 2013 |
| *Total Drafted* | 37694 | 0.27 | 0.73 | 0 | 9 |
| *High-Round Players (1)* | 37694 | 0.04 | 0.21 | 0 | 4 |
| *Non-High-Round Players (1)* | 37694 | 0.24 | 0.64 | 0 | 8 |
| *Talented Teammate (1)* | 37694 | 0.03 | 0.17 | 0 | 1 |
| *High-Round Players (1-2)* | 37694 | 0.07 | 0.32 | 0 | 5 |
| *Non-High-Round Players (1-2)* | 37694 | 0.20 | 0.56 | 0 | 8 |
| *Talented Teammate (1-2)* | 37694 | 0.06 | 0.23 | 0 | 1 |
| *High-Round Players (1-3)* | 37694 | 0.11 | 0.41 | 0 | 6 |
| *Non-High-Round Players (1-3)* | 37694 | 0.16 | 0.48 | 0 | 6 |
| *Talented Teammate (1-3)* | 37694 | 0.08 | 0.27 | 0 | 1 |

## *Player Data*

The dataset includes 10,434 players who were drafted into the NFL. Each observation has data on the year a player was drafted, the round and pick of the selection, the college or university the player attended, and performance statistics.

Eight general statistics were used in this thesis. Due to the difficulty of measuring NFL performance across positions, these eight statistics were chosen to represent four distinct aspects of a player's professional career.

One aspect of a player's career is his talent level within a team. This manifests itself in a team's depth chart. Each observation contains the number of years a player was a primary starter for his team in the NFL. This measure was turned into a dummy variable, *Primary Starter*, which had a value of 1 if a player ever became a primary starter for a team. Overall, 47% of the players drafted in this time period went on to become primary starters.

The second characteristic of a player's career is value. The third measure included was Average Approximate Value (*AV/Year*). This measure came from Approximate Value (AV), a metric created in "an attempt to put a single number on the seasonal value of a player at any position from any year" (Pro-Football-Reference). AV represents an attempt to create one value that can measure any player's on-field performance for a given year, regardless of that player's salary. It is similar in nature to Wins Above Replacement Player (WARP) in baseball analytics and Player Efficiency Rating (PER) in basketball analytics. The exact details of how this measure is created can be found online at http://www.sports-reference.com/blog/approximate-value-methodology/. Each drafted player has a Career Approximate Value, which added together 100% of a player's best AV season, 95% of a player's second-best AV season, etc. *AV/Year* was calculated by dividing a player's Career Approximate Value by number of years in the league. This value ranged from -5 to 18 and had a mean of 2.46. This was also transformed into a logarithmic measure to test functional form. At this transformation, a value of 1 was added to each observation. The two observations that were still negative or had a value of 0 were excluded from this analysis.

Two measures were used to test the star quality of a player: All-Pro selections and Pro Bowl selections. The average player was selected to the All-Pro team 0.09 times and was selected to the Pro Bowl 0.32 times.

The two final measures, games played and years in NFL, addressed longevity of career. The average number of years was 4.07 and the average number of games played was 55.64.

**Table 4** Descriptive Statistics of Player Data

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Year* | 10434 | 1990 | 13.57 | 1967 | 2013 |
| *Round* | 10434 | 4.27 | 2.18 | 1 | 9 |
| *Pick* | 10434 | 111.50 | 64.09 | 1 | 222 |
| *All Pro Selections* | 10434 | 0.09 | 0.52 | 0 | 10 |
| *Pro Bowl Selections* | 10434 | 0.32 | 1.19 | 0 | 14 |
| *Primary Starter (Years)* | 10434 | 2.20 | 3.32 | 0 | 24 |
| *Primary Starter (Y/N)* | 10434 | 0.47 | 0.50 | 0 | 1 |
| *Average Value/Year* | 10434 | 2.46 | 2.58 | -5 | 18 |
| *Log(Average Value/Year)* | 10434 | 0.95 | 0.78 | 0 | 2.94 |
| *Years in NFL* | 10434 | 4.07 | 3.85 | 0 | 25 |
| *Games Played* | 10434 | 55.64 | 56.36 | 0 | 382 |
| *Talented Teammate (1)* | 10434 | 0.37 | 0.48 | 0 | 1 |
| *Talented Teammate (1-2)* | 10434 | 0.57 | 0.49 | 0 | 1 |
| *Talented Teammate (1-3)* | 10434 | 0.70 | 0.46 | 0 | 1 |
| *Offense-Defense Talented Teammate (1)* | 10434 | 0.28 | 0.45 | 0 | 1 |
| *Offense-Defense Talented Teammate (1-2)* | 10434 | 0.46 | 0.50 | 0 | 1 |
| *Offense-Defense Talented Teammate (1-3)* | 10434 | 0.60 | 0.49 | 0 | 1 |

The dummy variable *Talented Teammate* was added that gave a value of 1 to players who attended a school that had a high-round draft selection in the year of their draft. Six versions of this binary were included, one for each definition of high rounds across the two school groupings (all players from a school and players on offense/defense from a school). 37% of players attended a school that had a player selected in the first round during the year they were drafted, while 57% had a teammate drafted in the first two rounds and 70% had one drafted in the first three rounds. When broken down by offense and defense, the percentages are 28%, 46%, and 60%, respectively. Table 4 contains the descriptive statistics of these player data.

## 3.3. Methodology

Fixed Effect models were used in this analysis. This was done to remove the effect of high quality football schools that produce a large number of NFL players, both in the high rounds and subsequent rounds. The analysis was broken down into two sections: determining if the Talented Teammate effect increased the number of players a school had drafted and if players drafted in part due to the Talented Teammate effect performed worse in the NFL than their counterparts.

### *School Level Models*

Regressions were used to determine if teams with a high-round draft selection had systematically more players drafted. All of the models took the basic form:

$$D_{it} = \beta_0 + \beta_1 * T_{it} + a_i + u_{it}$$

where $D$ is a measure of the number of players drafted, $T$ is a measure of players drafted in a high round, $\beta_0$ and $\beta_1$ are parameters, $a$ denotes school-specific fixed effects, $u$ is the error term, $i$ indexes the school while $t$ indexes the year.

Twelve models were estimated. The first six models used the holistic school level data. Three of the six models used the number of players drafted from a school minus the first player selected in a high round as the dependent variable. These models examine the idea that the most talented member of a draft class can pull everyone else up. The other three models had as a dependent variable the total number of players drafted outside the early rounds. This analysis examines whether the presence of a high-caliber teammate helps out only the lower quality players on a team. For each of these dependent variables, the definition of high round was manipulated. Three different definitions were used: Round 1 players, Round 1-2 players, and Round 1-3 players.

Round was chosen as the determining factor instead of pick number because round better reflects the time management of a team's scouting. Even though the first pick of the second round and the last pick of the first round are only one selection apart, they represent significantly different situations to the two teams. For the team drafting last in the first round, it is still their first pick in the draft and therefore deserves the largest time commitment. The first pick in the second round, however, is most often a team's second pick and therefore will receive less scouting time.

The final six models used the same form in terms of independent and dependent variables. The key difference was the data used. While the first six models analyzed the Talented Teammate effect at the overall school level, these models examine the hypothesis at the school offensive/defensive level. Consequently, the offense-defense school data were used. In all models, the coefficient on the Talented Teammate variable is expected to be positive, indicating that the presence of a high-round draft selection significantly increases the number of players drafted from that player's school.

### *Player Level Models*

Models were estimated to see if players drafted in part due to a talented teammate effect performed systematically worse in the NFL than their counterparts drafted near them. The models had the basic form:

$$P_i = \delta_0 + \delta_1 * R_i + \delta_2 * T_i + a_i + u_i$$

where $P$ is a measure of NFL performance, $R$ is a vector of variables related to when the player was selected in the draft, $T$ is a measure of players drafted in Round 1, $\delta_0$, $\delta_1$, and $\delta_2$ are parameters, $a$ denotes school-specific fixed effects, $u$ is the error term, and $i$ indexes the player.

There were eight different performance statistics used in this analysis: years as a primary starter, whether a player ever became a primary starter, Pro Bowl selections, All Pro selections, years in NFL, games played, and approximate value per year in both a linear and log form. For each performance statistics, six models were created that had varying draft location vectors.

Three used round number as the primary variable, while the other three used pick number. In both cases, each variable was included in a model in a linear, squared, and cubed form.

Those 48 models were run 18 times each. They varied on which data set they used (school level or offense-defense school level), definition of high round (Round 1, Round 1-2, or Round 1-3), and which players were analyzed (all players, all players minus first in high round, or all non-high-round players). Collectively, 864 regressions were analyzed. To avoid Type I errors stemming from the large number of models, only consistent trends in the data across similar metrics will be presented.

In all of these models, it is predicted that the talented teammate binary will have a negative value, indicating that players drafted due to this bias underperform relative to their draft position. Similarly, the combined value of the parameters on the draft location variables is expected to be negative. This would suggest that players drafted early on outperform those drafted later in the draft.

## 3.4. Results

Regressions were run using the models discussed above.

### *Schools and the Talented Teammate Effect*

The twelve school models were estimated. School-specific fixed effects were significant in all models. A Hausman test showed that these were not significantly different from random effects. However, fixed effects were possible with the large number of degrees of freedom available in the model and were kept because of the logical approach of measuring the number of players a school has drafted relative to itself.

Of the twelve models, six are presented here. The results were similar between models that regressed the talented teammate binary on all players drafted minus the first player selected

**Table 5** Regression Results for *Non-High-Round Players* Holistic School Models

| VARIABLES | (1)<br>*Non-High-Round Players* | (2)<br>*Non-High-Round Players* | (3)<br>*Non-High-Round Players* |
|---|---|---|---|
| *Talented Teammate* | 0.471*** | 0.394*** | 0.279*** |
| | (0.0700) | (0.0430) | (0.0272) |
| Constant | 0.456*** | 0.370*** | 0.296*** |
| | (0.00384) | (0.00416) | (0.00362) |
| Observations | 18,847 | 18,847 | 18,847 |
| R-squared | 0.014 | 0.019 | 0.015 |
| Number of Schools | 401 | 401 | 401 |
| Fixed Effects? | Yes | Yes | Yes |
| High-Round Definition | Round 1 | Round 1-2 | Round 1-3 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

in a high round and those that regressed on non-high-round selections. In keeping with the psychological understanding of this bias, the dependent variable reported here is number of non-high-round players since these are the players who would benefit the most from a talented teammate.

Table 5 shows the results of the analysis on the holistic school-level data. These models found a significant effect where schools with a player selected in a high round, by any of the definitions used, had significantly more players drafted outside the high rounds. The effect sizes were 0.47 players when high round meant Round 1, 0.39 players when high round meant Round 1 and 2, and 0.28 players when high round meant Rounds 1-3, on average and ceteris paribus. The decreasing effect size is likely due to the decrease in the total number of players drafted in the non-high rounds, since each model includes fewer rounds as the dependent variable. A similar trend can be found in the constants of the models. When all of the models were run using the same data observations of players drafted outside of Round 3, the results maintained their significance throughout.

**Table 6** Regression Results for *Non-High-Round Players* Offense-Defense School Models

| VARIABLES | (1)<br>*Non-High-Round Players* | (2)<br>*Non-High-Round Players* | (3)<br>*Non-High-Round Players* |
|---|---|---|---|
| *Talented Teammate* | 0.298*** | 0.224*** | 0.147*** |
| | (0.0443) | (0.0264) | (0.0176) |
| Constant | 0.226*** | 0.187*** | 0.150*** |
| | (0.00138) | (0.00151) | (0.00143) |
| Observations | 37,694 | 37,694 | 37,694 |
| R-squared | 0.008 | 0.010 | 0.007 |
| Number of Schools | 401 | 401 | 401 |
| Fixed Effects? | Yes | Yes | Yes |
| High-Round Definition | Round 1 | Round 1-2 | Round 1-3 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The same analysis was done using the dataset broken down into offense and defense. In these models, the presence of a talented teammate on a specific side of the ball increased the number of players drafted from the school on that side of the ball by 0.30 after Round 1, 0.22 after Round 2, and 0.15 after Round 3, on average and ceteris paribus. All of these results were significant at the 1% level. As before, the downward trend in size is indicative of a shrinking dependent variable. These results are listed in Table 6.

Overall, these models suggest that schools with high-round draft selections get more players drafted overall, on average and ceteris paribus. Though significant effects were found, these models were not very strong. The within-$R^2$ values range from 0.014 to 0.19 in the holistic school models and 0.007 to 0.01 in the offense-defense models. This suggests that while there is a significant effect of having a highly scouted player, this alone does not explain the variation in the number of players a school has drafted from year to year. Such a result makes sense because the Talented Teammate hypothesis looks to describe how the availability heuristic could bias decisions. It is not meant to explain fully each draft selection. Much of the explanatory power that could be found in these models is removed due to the inclusion of fixed effects. The notable outcome is that the key variable is found to be significant both statistically and practically. Therefore, the low $R^2$ values are less important relative to the other findings.

### *Players and the Talented Teammate Effect*

Of the 864 regressions that were run, only a small subset of them will be presented here. Due to the decision at the school level to use non-high-round players as the dependent variables, the models analyzed here will use that subset of the data. All models discussed here include the variable *Round* in its linear and squared form. Round was chosen over pick number due both to similar results and to mirror the definition of talented teammate. The cubed version of Round was dropped due to insignificance, as were models with only the linear form of Round.

Although eight different performance metrics were used as dependent variables, only three will be presented here: whether a player ever became a primary starter, games played in the NFL, and the linear form of average approximate value per year. These metrics were chosen because they speak to different aspects of a player's NFL career. *Games Played* relates to a player's longevity and is similar to the years in NFL metric. Whether a player became a primary starter provides a binary way of seeing whether a player ever became a significant part of a team and is directly related to the number of years a player was a starter. Average approximate value per year is an attempt to quantify a player's value. The linear form is used both because that is how it was originally intended to be used by its creators and because the results between the linear and log forms were similar.

Finally, there is no star-quality variable, such as pro bowl or all selections, mentioned here because it related to such a small portion of the sample size. Therefore, 18 models will be discussed in depth. Overall, these models show mixed results.

**Table 7** Regression Results for *Non-High-Round Players* Holistic Players Models, Talented Teammate = Round 1

| VARIABLES | (1)<br>*Primary Starter* | (2)<br>*Games Played* | (3)<br>*Average Approximate Value per Year* |
|---|---|---|---|
| *Round* | -0.238*** | -22.60*** | -1.280*** |
| | (0.0139) | (1.499) | (0.0702) |
| *Round*$^2$ | 0.0156*** | 1.423*** | 0.0862*** |
| | (0.00137) | (0.149) | (0.00688) |
| *Talented Teammate* | -0.0274** | -2.641* | -0.0719 |
| | (0.0111) | (1.461) | (0.0544) |
| Constant | 1.141*** | 120.0*** | 5.895*** |
| | (0.0322) | (3.409) | (0.159) |
| Observations | 9,074 | 9,074 | 9,074 |
| R-squared | 0.118 | 0.104 | 0.144 |
| Number of Schools | 400 | 400 | 400 |
| Fixed Effects? | Yes | Yes | Yes |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7 shows the results of models using the holistic school data to define *Talented Teammate* and using Round 1 as the high-round definition. Under these parameters, *Talented Teammate* is found to have a detrimental effect on both whether a player becomes a primary starter and the number of games a player participates in. These effects, while statistically significant, are rather small in size. It suggests that being drafted in part because of a talented college teammate is associated with a lesser frequency of becoming a primary starter by 2.8% and reduces the number of games you will play in by 2.64, or about a sixth of a season, on average and ceteris paribus. Though the coefficient is in the expected negative direction, there is no significant relationship between having a high-round draft selection teammate and average approximate value.

These results are not consistent through different high-round definitions. When high round is defined as the first two rounds, all statistical significance for the *Talented Teammate* variable is lost. These results are shown in Table 8. When high round is defined as the first three

**Table 8** Regression Results for *Non-High-Round Players* Holistic Players Models, Talented Teammate = Round 1-2

| VARIABLES | (1)<br>*Primary Starter* | (2)<br>*Games Played* | (3)<br>*Average Approximate Value per Year* |
|---|---|---|---|
| *Round* | -0.242*** | -22.54*** | -1.103*** |
| | (0.0210) | (2.449) | (0.102) |
| $Round^2$ | 0.0159*** | 1.420*** | 0.0713*** |
| | (0.00188) | (0.217) | (0.00910) |
| *Talented Teammate* | -0.0163 | -0.982 | -0.0281 |
| | (0.0122) | (1.444) | (0.0561) |
| Constant | 1.152*** | 119.5*** | 5.396*** |
| | (0.0554) | (6.480) | (0.268) |
| Observations | 7,697 | 7,697 | 7,697 |
| R-squared | 0.061 | 0.057 | 0.070 |
| Number of Schools | 395 | 395 | 395 |
| Fixed Effects? | Yes | Yes | Yes |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

rounds, however, the coefficients are again significant. Similar to the Round 1 definition, having a talented teammate predicts a decrease in a player's chance of becoming a primary starter by 3% and a reduction in the number of games a player will play in by 2.82, on average and ceteris paribus. Unlike before, *Talented Teammate* now has a significant relationship with average approximate value, decreasing it by roughly 0.1. These numbers are presented in Table 9. Again, though statistical significance exists, the small effect size creates questions about the economic validity of these results.

The final nine models run the same analyses on the offense-defense definition of *Talented Teammate*. Overall, these models show fairly similar results. In all three definitions of high round, *Talented Teammate* is significant in regards to *Primary Starter* and *Games Played*. For *Primary Starter*, the value ranges from -0.035 in the Round 1 model to -0.048 in the Round 1-3 model. The smallest value for *Games Played* is -2.8 in the Round 1-2 model. That number increases in magnitude to -3.6 in the Round 1 model and -3.8 in the Round 1-3 model. In all

**Table 9** Regression Results for *Non-High-Round Players* Holistic Players Models, Talented Teammate = Round 1-3

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | *Primary Starter* | *Games Played* | *Average Approximate Value per Year* |
| | | | |
| *Round* | -0.271*** | -25.13*** | -1.319*** |
| | (0.0376) | (3.746) | (0.164) |
| *Round*$^2$ | 0.0182*** | 1.632*** | 0.0884*** |
| | (0.00304) | (0.307) | (0.0135) |
| *Talented Teammate* | -0.0299** | -2.820** | -0.0989* |
| | (0.0137) | (1.369) | (0.0569) |
| Constant | 1.247*** | 128.1*** | 6.085*** |
| | (0.112) | (11.10) | (0.482) |
| | | | |
| Observations | 6,276 | 6,276 | 6,276 |
| R-squared | 0.030 | 0.029 | 0.038 |
| Number of Schools | 379 | 379 | 379 |
| Fixed Effects? | Yes | Yes | Yes |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

three models, *Talented Teammate* has no significant relationship with *Average Approximate Value per Year*, though it always has the expected sign. These results are outlined in Tables 10, 11, and 12.

Overall, these models vary greatly in regards to strength. When using a Round 1 definition of high round, the within-$R^2$s range from 0.104 to 0.144. This range drops to 0.057 to 0.070 for Round 1-2 models and 0.029 to 0.038 for Round 1-3 models. This trend suggests that the Talented Teammate bias has more explanatory power when there are stricter definitions of what makes a talented teammate. As with the school level, this hypothesis is not designed to fully explain future performance and thus the low $R^2$ values are less concerning.

## 3.5. Discussion

The present hypothesis argues that a talented teammate effect occurs as a manifestation of the availability heuristic defined by Tversky and Kahneman (1974). The basis for this argument is that scouts spend more time scouting early-round prospects than late-round

**Table 10** Regression Results for *Non-High-Round Players* Offense-Defense Players Models, Talented Teammate = Round 1

| VARIABLES | (1)<br>*Primary Starter* | (2)<br>*Games Played* | (3)<br>*Average Approximate Value per Year* |
|---|---|---|---|
| *Round* | -0.238*** | -22.61*** | -1.280*** |
| | (0.0139) | (1.500) | (0.0701) |
| *Round*$^2$ | 0.0156*** | 1.423*** | 0.0862*** |
| | (0.00137) | (0.149) | (0.00687) |
| *Talented Teammate* | -0.0351** | -3.621** | -0.0202 |
| | (0.0154) | (1.745) | (0.0779) |
| Constant | 1.140*** | 119.9*** | 5.878*** |
| | (0.0316) | (3.400) | (0.158) |
| Observations | 9,074 | 9,074 | 9,074 |
| R-squared | 0.119 | 0.104 | 0.144 |
| Number of Schools | 400 | 400 | 400 |
| Fixed Effects? | Yes | Yes | Yes |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

**Table 11** Regression Results for *Non-High-Round Players* Offense-Defense Players Models, Talented Teammate = Round 1-2

| VARIABLES | (1)<br>*Primary Starter* | (2)<br>*Games Played* | (3)<br>*Average Approximate Value per Year* |
|---|---|---|---|
| *Round* | -0.242*** | -22.57*** | -1.103*** |
| | (0.0211) | (2.454) | (0.102) |
| *Round*$^2$ | 0.0160*** | 1.422*** | 0.0714*** |
| | (0.00189) | (0.217) | (0.00910) |
| *Talented Teammate* | -0.0382*** | -2.754* | -0.0255 |
| | (0.0140) | (1.627) | (0.0636) |
| Constant | 1.157*** | 119.9*** | 5.392*** |
| | (0.0557) | (6.555) | (0.269) |
| Observations | 7,697 | 7,697 | 7,697 |
| R-squared | 0.062 | 0.057 | 0.070 |
| Number of Schools | 395 | 395 | 395 |
| Fixed Effects? | Yes | Yes | Yes |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

**Table 12** Regression Results for *Non-High-Round Players* Offense-Defense Players Models, Talented Teammate = Round 1-3

| VARIABLES | (1) *Primary Starter* | (2) *Games Played* | (3) *Average Approximate Value per Year* |
|---|---|---|---|
| *Round* | -0.273*** | -25.31*** | -1.321*** |
| | (0.0377) | (3.762) | (0.165) |
| *Round*$^2$ | 0.0184*** | 1.645*** | 0.0886*** |
| | (0.00305) | (0.308) | (0.0135) |
| *Talented Teammate* | -0.0477*** | -3.841** | -0.0680 |
| | (0.0140) | (1.677) | (0.0676) |
| Constant | 1.255*** | 128.5*** | 6.064*** |
| | (0.112) | (11.13) | (0.484) |
| Observations | 6,276 | 6,276 | 6,276 |
| R-squared | 0.031 | 0.029 | 0.038 |
| Number of Schools | 379 | 379 | 379 |
| Fixed Effects? | Yes | Yes | Yes |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

prospects. This time allocation is logical given that high quality players, who are drafted early, can impact a team more than lower quality late-round players. When a player is scouted, however, he is not viewed in isolation. In the majority of cases, his teammates are also observed, either in the game tape that a scout watches or during a school's Pro Day where all the draft-eligible players perform drills. Because of this process, players who are late-round or unlikely to be drafted prospects who are teammates of high quality players will be seen more than their talent level deserves.

This scouting process can lead scouts to misjudge players in two distinct ways. The first is based on "biases due to the retrievability of instances" (Tversky & Kahneman, 1974). According to this argument, a group of players that are more retrievable could be thought of as more talented or simply better liked. This would lead scouts to draft players who they are most familiar with or players who are easiest to recall. In the late rounds, when all of the high quality players have already been selected, players who have been seen more often due to their teammates will stick out and therefore are more likely to be drafted. By this process, teammates of early-round talents are more likely to be drafted than they would be without their star teammates.

The second potential explanation is related to “illusory correlation” (Tversky & Kahneman, 1974). Because of the basic time allocation that scouts use, high quality players are seen more than low quality players. This could lead scouts to create a correlation between how often they view a player and the talent of said player. This could happen at a conscious or subconscious level. If this happens, players who are unintentionally scouted beyond their talent level will be considered more talented than they actually are. Thus, teammates of early-round players will be considered talented because they have been seen more often, regardless of their actual talent level. This will make them more likely to be drafted or taken earlier in the draft.

These two mechanisms, “biases due to retrievability of instances” and “illusory correlation”, are potential explanations for the existence of the talented teammate effect stemming from the availability heuristic. Given these arguments, such an effect can be thought of as a manifestation of a human decision-making bias.

The present investigation found that such an effect did exist in the NFL draft whereby teams with players selected in early rounds had more players drafted overall. This was true when early-round definition varied from Round 1 to Round 1-3 and when the schools were broken down into offensive and defensive units. At the player level, the results are mixed regarding how this drafting bias relates to future performance. In most of the models, players who are drafted in part due to a star teammate are less likely to become a starter and play fewer games than the average player in their drafted round. The talented teammate effect, however, is not associated with any difference in a player's average approximate value.

One finding that must be discussed is the disappearance of all significant effects in models that use the first two rounds as the definition for a talented teammate and use the holistic school data. There are several explanations for this surprising finding. First, it could be that teammates of first round picks drafted in the second round perform significantly less well than other second round picks. By removing these from the sample, it improves the overall performance of players who have star teammates. Second, teammates of second round players could perform better than or on par with others in their round. This again would improve the overall play of the group of players who had a talented teammate. To determine which caused this anomaly and why requires further analysis.

Overall, the analysis supports the presence of a talented teammate bias in the NFL draft. Furthermore, this bias is leading teams to select players who will have shorter careers and be less likely to become a starter. Given these detrimental aspects, teams would benefit by removing this bias and drafting players regardless of their college teammates. Possible strategies for accomplishing this are discussed in the general discussion.

## 4. Hypothesis 2 – Geographic Bias

### 4.1. Hypothesis

The Geographic Bias hypothesis contends that teams will be more likely to draft players from their region than the rest of the league. As with the Talented Teammate hypothesis, this theory stems from the Availability Heuristic through both the potential for an illusory correlation between familiarity and player talent and the relationship between familiarity and likeability.

The difference is how players become more familiar in a scout's mind. While at work, scouts are likely to approach the study of players systematically. However, scouts are also likely to see and hear about players outside of the office. This could be through many types of media, including television, radio, newspapers, Facebook, and Twitter. The key aspect of this additional stimulus is that it is biased towards players who attend colleges near the team and employee's home. These players are more likely to be mentioned in local media, appear on a scout's Facebook or Twitter feed in some capacity, and be discussed by the general populace. By being exposed to these additional stimuli, local players have the potential to become available in a scout or general manager's mind. This added familiarity to the local team could then improve a player's chance of getting drafted via both increased fondness towards the player and a perceived increase in the player's talent level.

To analyze this hypothesis, draft tendencies will be studied to see if teams draft from their area at a higher rate than the rest of the NFL. If this bias is confirmed, a regression analysis will be used to see how these local players perform in the NFL given their draft position. Finally, psychological recommendations will be made on how teams can protect against this irrational behavior.

## 4.2. Data

Much of the data used in this analysis are the same as those used in the Talented Teammate hypothesis. Once again, the data looks at players drafted in the top 222 picks during the Super Bowl era. The data are collected at the player level and includes year, round, pick, school, and performance statistics. Using the schools, each player was mapped to a specific state (US News Best Colleges). The same was done for each NFL team. Overall, 5% of players

**Table 13** Descriptive Statistics of Player Data

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Year* | 10434 | 1990 | 13.57 | 1967 | 2013 |
| *Round* | 10434 | 4.27 | 2.18 | 1 | 9 |
| *Pick* | 10434 | 111.50 | 64.09 | 1 | 222 |
| *All Pro Selections* | 10434 | 0.09 | 0.52 | 0 | 10 |
| *Pro Bowl Selections* | 10434 | 0.32 | 1.19 | 0 | 14 |
| *Primary Starter (Years)* | 10434 | 2.20 | 3.32 | 0 | 24 |
| *Primary Starter (Y/N)* | 10434 | 0.47 | 0.50 | 0 | 1 |
| *Average Value/Year* | 10434 | 2.46 | 2.58 | -5 | 18 |
| *Log(Average Value/Year)* | 10434 | 0.95 | 0.78 | 0 | 2.94 |
| *Years in NFL* | 10434 | 4.07 | 3.85 | 0 | 25 |
| *Games Played* | 10434 | 55.64 | 56.36 | 0 | 382 |
| *Same State* | 10434 | 0.05 | 0.22 | 0 | 1 |

drafted were selected by teams in the same state as their college. Table 13 contains the summary statistics at the player level.

These data were then aggregated at the team level. Teams were determined by unique franchise-state dyads. If a team moved to a new state, such as when the Houston Oilers became the Tennessee Titans, they were counted as two different teams. Two franchises that occupied the same city at different times, such as the Baltimore Colts and the Baltimore Ravens, were likewise divided. The only franchise moves that are not reflected in the number of teams is when a team moves within its own state, such as when the Patriots moved from Boston to Foxborough. Using these parameters, there were 37 teams included in the study. The teams are listed in Table 14, along with their home state, years in existence, and total number of picks. While the majority of teams were given one home state, there were three exceptions. The New York Giants and New York Jets actually play in New Jersey, so both states were considered home states for those teams. Similarly, the Washington Redskins have operations in both Virginia and Maryland and thus consider both home states.

**Table 14** List of NFL Teams

| Team | State | First Year | Last Year | Number of Picks |
|---|---|---|---|---|
| Arizona Cardinals | AZ | 1988 | 2013 | 199 |
| Atlanta Falcons | GA | 1967 | 2013 | 368 |
| Baltimore Colts | MD | 1967 | 1983 | 150 |
| Baltimore Ravens | MD | 1996 | 2013 | 126 |
| Buffalo Bills | NY | 1967 | 2013 | 391 |
| Carolina Panthers | NC | 1995 | 2013 | 133 |
| Chicago Bears | IL | 1967 | 2013 | 363 |
| Cincinnati Bengals | OH | 1968 | 2013 | 423 |
| Cleveland Browns (1) | OH | 1967 | 1995 | 214 |
| Cleveland Browns (2) | OH | 1999 | 2013 | 114 |
| Dallas Cowboys | TX | 1967 | 2013 | 393 |
| Denver Broncos | CO | 1967 | 2013 | 331 |
| Detroit Lions | MI | 1967 | 2013 | 348 |
| Green Bay Packers | WI | 1967 | 2013 | 382 |
| Houston Oilers | TX | 1967 | 1996 | 250 |
| Houston Texans | TX | 2002 | 2013 | 90 |
| Indianapolis Colts | IN | 1984 | 2013 | 217 |
| Jacksonville Jaguars | FL | 1995 | 2013 | 136 |
| Kansas City Chiefs | MO | 1967 | 2013 | 355 |
| Los Angeles Rams | CA | 1967 | 1994 | 230 |
| Miami Dolphins | FL | 1967 | 2013 | 376 |
| Minnesota Vikings | MN | 1967 | 2013 | 339 |
| New Orleans Saints | LA | 1967 | 2013 | 363 |
| New England Patriots | MA | 1967 | 2013 | 373 |
| New York Giants | NJ, NY | 1967 | 2013 | 336 |
| New York Jets | NJ, NY | 1967 | 2013 | 377 |
| Oakland Raiders | CA | 1967 | 2013 | 319 |
| Philadelphia Eagles | PA | 1967 | 2013 | 356 |
| Pittsburgh Steelers | PA | 1967 | 2013 | 411 |
| San Diego Chargers | CA | 1967 | 2013 | 354 |
| Seattle Seahawks | WA | 1976 | 2013 | 272 |
| San Francisco 49ers | CA | 1967 | 2013 | 188 |
| St. Louis Cardinals | MO | 1967 | 1987 | 141 |
| St. Louis Rams | MO | 1995 | 2013 | 329 |
| Tampa Bay Buccaneers | FL | 1976 | 2013 | 272 |
| Tennessee Titans | TN | 1997 | 2013 | 136 |
| Washington Redskins | MD, VA | 1967 | 2013 | 265 |

## 4.3. Methodology

Fixed Effect models were used in this analysis. Fixed effects were done at the college/university level in an attempt to control for schools that both produce a large number of NFL-caliber players and are located in states with a large number of NFL teams. The analysis broke down into three levels: NFL team, round, and player.

### *Team Level Analysis*

2x2 contingency tables were calculated for each of the 37 NFL teams. The two dimensions of the table were whether or not a player attended college in the same state as the team and whether or not the team drafted the player. Chi-square analysis was used to determine if there was a significant interaction between the two dimensions. The purpose of these tests is to determine which teams suffer from a geographic bias by over-drafting from their state relative to the rest of the league.

### *Round Level Analysis*

This level of analysis addressed the question of how drafting behavior in regards to a potential geographic bias changed throughout the draft. A preliminary analysis was done to see the percentages of same-state picks in each round. Given that this is a potentially biased test because not all players are eligible to be drafted by an in-state team, the same analysis is done on the subset of players who played college in a state with an NFL team.

### *Player Level Analysis*

The analysis here mirrors the player level analysis of the Talented Teammate hypothesis. As was done there, eight different performance metrics were used as dependent variables regressed on a vector of either round or pick variables ranging from linear to cubed form. The independent variable of note in these models is *Same*, which is given a value of 1 if the player was drafted by a team in the same state as their college.

These models were run on two different sets of data. The first dataset includes all players drafted into the NFL during the Super Bowl era. The second set of data includes all players drafted from a state that contains a NFL team. These models were used because these players are the only ones who had the potential to be drafted due to the geographic bias.

Despite the differences between them, the predictions for all these models are the same. The variable of interest, *Same*, is expected to have a negative coefficient, suggesting that teams are drafting suboptimally due to the geographic bias. As before, the draft location parameters are expected to be negative, suggesting that players become less talented the later they are drafted.

## 4.4. Results

The 2x2 contingency tables and regressions discussed above were run and analyzed.

### *Team Level Analysis*

Table 15 presents the contingency tables for all 37 NFL teams. In all, there were ten teams that showed a geographic bias: the Atlanta Falcons, the original Cleveland Browns, and the Jacksonville Jaguars at a 10% significance level, the Pittsburgh Steelers at the 5% significance level, and the Baltimore Colts, the Los Angeles Rams, the Minnesota Vikings, the New England Patriots, the Oakland Raiders, and the Tampa Bay Buccaneers at the 1% significance level. The New York Jets, on the other hand, were the only team in the league that drafted significantly less from their state than the rest of the league. The specific percentage breakdowns of these eleven teams are presented in Table 16.

**Table 15** Team State/College Contingency Tables

| Team | Not From Not To | From Not To | Not From To | From To | $\chi^2$ | P-Value | Significant? |
|---|---|---|---|---|---|---|---|
| Arizona Cardinals | 10012 | 223 | 195 | 4 | 0.03 | 0.872 | No |
| Atlanta Falcons | 9849 | 217 | 355 | 13 | 3.12 | 0.077 | 10% |
| Baltimore Colts | 10178 | 106 | 145 | 5 | 7.45 | 0.006 | 1% |
| Baltimore Ravens | 10199 | 109 | 124 | 2 | 0.33 | 0.564 | No |
| Buffalo Bills | 9931 | 112 | 387 | 4 | 0.03 | 0.865 | No |
| Carolina Panthers | 9894 | 407 | 129 | 4 | 0.31 | 0.578 | No |
| Chicago Bears | 9876 | 195 | 352 | 11 | 2.17 | 0.141 | No |
| Cincinnati Bengals | 9681 | 330 | 405 | 18 | 1.16 | 0.282 | No |
| Cleveland Browns (1) | 9884 | 336 | 202 | 12 | 3.50 | 0.061 | 10% |
| Cleveland Browns (2) | 9976 | 344 | 110 | 4 | 0.01 | 0.917 | No |
| Dallas Cowboys | 9152 | 889 | 366 | 27 | 1.86 | 0.173 | No |
| Denver Broncos | 9901 | 202 | 324 | 7 | 0.02 | 0.883 | No |
| Detroit Lions | 9642 | 444 | 328 | 20 | 1.43 | 0.231 | No |
| Green Bay Packers | 9935 | 117 | 376 | 6 | 0.52 | 0.470 | No |
| Houston Oilers | 9293 | 891 | 225 | 25 | 0.48 | 0.490 | No |
| Houston Texans | 9433 | 911 | 85 | 5 | 1.18 | 0.278 | No |
| Indianapolis Colts | 9835 | 382 | 207 | 10 | 0.44 | 0.505 | No |
| Jacksonville Jaguars | 9672 | 626 | 123 | 13 | 2.83 | 0.093 | 10% |
| Kansas City Chiefs | 9971 | 108 | 349 | 6 | 1.21 | 0.270 | No |
| Los Angeles Rams | 9320 | 884 | 194 | 36 | 13.67 | 0.000 | 1% |
| Miami Dolphins | 9449 | 609 | 346 | 30 | 2.33 | 0.127 | No |
| Minnesota Vikings | 10025 | 70 | 331 | 8 | 12.28 | 0.000 | 1% |
| New Orleans Saints | 9721 | 350 | 349 | 14 | 0.15 | 0.697 | No |
| New England Patriots | 9948 | 113 | 363 | 10 | 7.49 | 0.006 | 1% |
| New York Giants | 9950 | 148 | 331 | 5 | 0.00 | 0.973 | No |
| New York Jets | 9905 | 152 | 376 | 1 | 3.91 | 0.048 | 5% |
| Oakland Raiders | 9239 | 876 | 275 | 44 | 10.13 | 0.001 | 1% |
| Philadelphia Eagles | 9701 | 377 | 342 | 14 | 0.04 | 0.851 | No |
| Pittsburgh Steelers | 9655 | 368 | 388 | 23 | 4.05 | 0.044 | 5% |
| San Diego Chargers | 9199 | 881 | 315 | 39 | 2.21 | 0.138 | No |
| Seattle Seahawks | 9945 | 217 | 265 | 7 | 0.24 | 0.623 | No |
| San Francisco 49ers | 9206 | 885 | 308 | 35 | 0.84 | 0.357 | No |
| St. Louis Cardinals | 10136 | 110 | 184 | 4 | 1.90 | 0.168 | No |
| St. Louis Rams | 10179 | 114 | 141 | 0 | 1.58 | 0.209 | No |
| Tampa Bay Buccaneers | 9554 | 608 | 241 | 31 | 13.51 | 0.000 | 1% |
| Tennessee Titans | 9893 | 405 | 130 | 6 | 0.08 | 0.775 | No |
| Washington Redskins | 9741 | 428 | 250 | 15 | 1.34 | 0.247 | No |

## *Round Level Analysis*

Given that a geographic bias was found in close to a third of the NFL states and teams, the next step is to understand where in the draft those picks are occurring. There is a large difference between picking hometown players in the early rounds and drafting such players later

**Table 16** Significantly Biased Teams

| State | Rest of League Percentage from State | Team Percentage from State |
|---|---|---|
| Atlanta Falcons | 2.16% | 3.53% |
| Baltimore Colts | 1.03% | 3.33% |
| Cleveland Browns (1) | 3.29% | 5.61% |
| Jacksonville Jaguars | 6.08% | 9.56% |
| Los Angeles Rams | 8.66% | 15.65% |
| Minnesota Vikings | 0.69% | 2.36% |
| New England Patriots | 1.12% | 2.68% |
| New York Jets | 1.51% | 0.27% |
| Oakland Raiders | 8.66% | 13.79% |
| Pittsburgh Steelers | 3.67% | 5.60% |
| Tampa Bay Buccaneers | 5.98% | 11.40% |

**Table 17** Round Breakdown of Same-State Rates using All Players

| Round | Percentage of Same-State Picks |
|---|---|
| 1 | 4.93% |
| 2 | 5.37% |
| 3 | 4.79% |
| 4 | 4.89% |
| 5 | 4.87% |
| 6 | 4.88% |
| 7 | 4.51% |

on in the draft. Any discernable pattern in the round could provide insight into why this phenomenon occurs. However, as Table 17 shows, the results are inconclusive. These numbers suggest that teams pick players from their home state consistently throughout the draft, with a floor of 4.51% in the seventh round and maxing out at 5.37% in the second round.

These numbers, however, are potentially misleading. There is an inherent bias because not all players can be drafted by a team in their state. Players who play in a state without an NFL team can never be drafted due to the geographic bias. It might be that the states that have NFL teams also produce higher quality college players, thereby skewing the sample. To better analyze this question, the percentages should be taken out of players drafted from the 22 NFL states. These results are presented in Table 18. As before, the percentages are fairly uniform, suggesting

**Table 18** Round Breakdown of Same-State Rates using Players from NFL States

| Round | Percentage of Same-State Picks |
|---|---|
| 1 | 6.46% |
| 2 | 7.56% |
| 3 | 6.69% |
| 4 | 6.79% |
| 5 | 7.08% |
| 6 | 7.04% |
| 7 | 6.64% |

that any geographic bias present in the draft is present at each level of scouting and player analysis.

## *Player Level Analysis*

The same reasoning used in the Talented Teammate hypothesis was used here. Therefore, this section presents the results of models in which the dependent variables were whether or not a player became a primary starter, games played in the NFL, and the linear form of average approximate value per year. *Round* was used to the second power to control for draft location. Given that this bias could apply only to players in NFL states, the analysis is performed on only this subset of the data. Fixed effects were used at the college level to control for any potential bias towards a state that has a large number of high quality college football programs. Table 19 shows the results of these models.

In all three cases, whether or not a local team drafted a player did not significantly impact his future performance. In fact, in two of the three models, the estimated coefficient was positive, further disagreeing with predictions. The results stay the same when the fixed effects are removed and when the entire set of players drafted is used. Despite this insignificant variable, all the models are decent in predictive power. The $R^2$ values range from 0.17 to 0.26. This predictive power comes exclusively from the draft position markers, all of which were highly significant.

**Table 19** Regression Results for Performance Metrics of NFL State Players

| VARIABLES | (1) *Primary Starter* | (2) *Games Played* | (3) *Average Approximate Value per Year* |
|---|---|---|---|
| *Round* | -0.219*** | -23.54*** | -1.501*** |
| | (0.00786) | (1.498) | (0.0543) |
| *Round*$^2$ | 0.0136*** | 1.492*** | 0.106*** |
| | (0.000934) | (0.162) | (0.00585) |
| *Same State* | 0.0100 | 0.147 | -0.0576 |
| | (0.0239) | (2.594) | (0.119) |
| Constant | 1.100*** | 122.9*** | 6.440*** |
| | (0.0134) | (2.804) | (0.103) |
| Observations | 7,386 | 7,386 | 7,386 |
| R-squared | 0.203 | 0.170 | 0.264 |
| Number of Schools | 270 | 270 | 270 |
| Fixed Effects? | Yes | Yes | Yes |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

## 4.5. Discussion

The Geographic Bias hypothesis argues that teams are more likely to draft players from within their state than the rest of the league. This develops from the availability heuristic and a potential illusory correlation between talent and familiarity. Unlike players affected by the Talented Teammate hypothesis, who are more available because of activities within the realm of scouting, local players are more familiar to a team's scouts because of factors outside of the workplace. Such factors include local media, social interactions, and fandom. As with all of the hypotheses presented in this thesis, any drafting effect stemming from these causes suggests a bias in scouting practices that derives from general human decision-making processes.

There were mixed results in regards to this hypothesis. There were several cases where a geographic bias was found. Ten of the 37 teams included in the sample showed a geographic bias. In all of these cases, the team drafted players from their state at a significantly higher rate than others did. Only one team, the New York Jets, drafted significantly less from their area.

These results, though significant for the teams showing the bias, represent less than a third of the NFL. Also, the bias is fairly consistent throughout the draft, so it does not always affect the most valuable players. Therefore, though an effect does exist, it is not widespread.

Even the teams that displayed the bias might have little to worry about. Analysis at the player level showed that players drafted by nearby teams did no worse than others drafted in their rounds. Given this lack of detrimental effects, a team may benefit from drafting locally if scouting such players required less resources.

An important observation is that this bias is generally small where it is significant at all. The greatest bias existed for the Los Angeles Rams, who drafted players from California 6.99 percentage points more often than the rest of the league. For teams still in existence, the Tampa Bay Buccaneers display the greatest bias, drafting from Florida 5.42 percentage points more often than the rest of the league. These small numbers suggest that this bias occurs only in a few instances, likely when the difference between a local player and a non-local player is sufficiently small. This also would explain why over-drafting locally does not harm teams.

In sum, the analysis suggests that a handful of teams do draft from their local states considerably more than other teams, indicating support for the hypothesis. This bias, however, is small and is not hurting teams when it comes to player performance. Therefore, though teams should always be aware of their biases, there is no concrete evidence suggesting teams should commit resources to avoiding the geographic bias.

## 5. Hypothesis 3 – Signal and Noise

### 5.1. Hypothesis

The Signal and Noise hypothesis argues that teams may exhibit irrational drafting habits because of a systematic misunderstanding of what traits predict future performance. This process occurs through the Representativeness Heuristic, whereby people judge an item's inclusion in a class by judging how similar the item is to the members of that class. In the case of football, teams may have specific beliefs about what defines a successful professional football player. These beliefs could be based on previous experience, physical skill, or other factors.

When these beliefs are applied to the draft, teams will look for players that match their concept of a successful NFL player and will draft players based on this comparison. While this process can work well if the scouts truly understand what makes a great player, it can also backfire if a team consistently draft players based on qualities that are unrelated to future performance. This will lead to suboptimal draft decisions and a less successful team as a whole.

The first step in analyzing this hypothesis is using regression models to determine what qualities are actually strong signals and predictors of future performance. Once these underlying connections are determined, the same qualities will be analyzed in relation to a player's draft position. Comparison between these models will show whether or not teams are making draft decisions based on the appropriate evidence. Due to the vast differences between positions, all of these models will be position-specific. As with the other hypotheses, the players drafted due to this bias will be studied to see if their performance is detrimental to the team. Finally, best practices will be explored on ways teams can avoid drafting players based on extraneous noise.

## 5.2. Data

The data used in this analysis come from two sources. On one side, statistics concerning a player's professional career comes from the same data used in the previous two hypotheses. As before, the data are restricted to include only those players drafted in the top 222 picks of a draft during the Super Bowl era. This includes a player's name, round, pick, college, position, and

**Table 20** Position Group Breakdown

| Position Group | Number of Players | Included Positions |
|---|---|---|
| Quarterbacks | 395 | QB |
| Running Backs | 1008 | FB, HB, RB |
| Pass Catchers | 1403 | WR, TE, FL, E |
| Offensive Line | 136 | OL, T, G, C |
| Defensive Line | 444 | DL, DE, DT, DE-DT, NT |
| Linebackers | 668 | LB, OLB, ILB |
| Defensive Backs | 1227 | DB, CB, S, FS, SS |
| Special Teams | 154 | K, P, LS, KR |

performance metrics. The eight general metrics discussed in the previous two hypotheses are included as well as position-specific statistics.

The second data source provides a player's college career statistics. This data set, collected from Sports-Reference – College Football, includes a player's name, college, year of graduation, and position-specific career statistics.

The players were subdivided into eight position groups so that regressions could be run using position-specific statistics. The eight classes are: quarterbacks, running backs, pass catchers, offensive line, defense line, linebackers, defensive backs, and special teams. The breakdown of what positions went into each category, as well as the size of each group, is listed in Table 20. Due to data limitations, analyses were conducted on only three of the eight subgroups. Offensive line was removed since there were no college blocking statistics included in the data. Special Teams were removed since each position within the category had a different set of statistics and to subdivide the group would create samples too small. Finally, all three defensive groups were removed because the collection of defensive statistics, most notably tackles, was inconsistent.

This leaves three primary groups of interest: quarterbacks, running backs, and pass catchers. Overall, college statistics were found for 395 of the 477 quarterbacks drafted into the NFL during the Super Bowl era. 1008 of the 1251 running backs were similarly matched.

**Table 21** Summary Statistics for Quarterbacks

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Year* | 395 | 1991.10 | 13.68 | 1967 | 2013 |
| *Round* | 395 | 3.83 | 2.19 | 1 | 8 |
| *Pick* | 395 | 98.27 | 68.52 | 1 | 222 |
| *All Pro Selections* | 395 | 0.07 | 0.43 | 0 | 6 |
| *Pro Bowl Selections* | 395 | 0.46 | 1.48 | 0 | 12 |
| *Primary Starter (Years)* | 395 | 2.15 | 3.54 | 0 | 19 |
| *Primary Starter (Y/N)* | 395 | 0.46 | 0.50 | 0 | 1 |
| *Years in NFL* | 395 | 5.21 | 4.82 | 0 | 20 |
| *Games Played* | 395 | 45.78 | 56.07 | 0 | 302 |
| *Approximate Value per Year* | 395 | 2.37 | 3.15 | -5 | 18 |
| *Log(Approximate Value/Year)* | 394 | 0.86 | 0.82 | 0 | 2.94 |
| *Pro Passing Completions* | 395 | 569.23 | 918.40 | 0 | 6300 |
| *Pro Passing Attempts* | 395 | 992.78 | 1543.72 | 0 | 10169 |
| *Pro Passing YDs* | 395 | 6799.41 | 10998.11 | -2 | 71838 |
| *Pro Passing TDs* | 395 | 40.03 | 70.52 | 0 | 508 |
| *Pro Passing INTs* | 395 | 36.06 | 53.38 | 0 | 336 |
| *Pro Passing Percentage* | 317 | 0.53 | 0.11 | 0 | 1 |
| *Pro Passing YDs/Attempt* | 317 | 6.19 | 1.61 | -1 | 15 |
| *Pro Passing TDs/Attempt* | 317 | 0.03 | 0.02 | 0 | 0.2 |
| *Pro Passing INTs/Attempt* | 317 | 0.05 | 0.05 | 0 | 0.5 |
| *College Passing Attempts* | 395 | 817.72 | 380.88 | 64 | 2183 |
| *College Passing Percentage* | 395 | 56.82 | 5.52 | 40.9 | 70.4 |
| *College Passing YDs/Attempt* | 395 | 7.52 | 0.81 | 5.1 | 10 |
| *College Passing TDs* | 395 | 43.87 | 24.12 | 1 | 131 |
| *College Passing INTs* | 395 | 28.89 | 12.63 | 1 | 75 |

Finally, of the 1793 pass catchers included in the NFL dataset, 1403 were identified in the collegiate dataset.

Each of the three positions has its own set of college and professional statistics in addition to the general information. For quarterbacks, the professional statistics include passing completions, attempts, yards, touchdowns, and interceptions. These are then used to create completion percentage (completions/attempts), yards per attempt, touchdowns per attempt, and interception per attempt. For college, the statistics include passing attempts, completions, completion percentage, touchdowns, interceptions, yards, and yards per attempt. The summary statistics for this position group are presented in Table 21.

**Table 22** Summary Statistics for Running Backs

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Year* | 1008 | 1988.49 | 13.22 | 1967 | 2013 |
| *Round* | 1008 | 4.09 | 2.26 | 1 | 9 |
| *Pick* | 1008 | 105.41 | 65.73 | 1 | 222 |
| *All Pro Selections* | 1008 | 0.08 | 0.43 | 0 | 6 |
| *Pro Bowl Selections* | 1008 | 0.34 | 1.06 | 0 | 10 |
| *Primary Starter (Years)* | 1008 | 1.61 | 2.51 | 0 | 14 |
| *Primary Starter (Y/N)* | 1008 | 0.44 | 0.50 | 0 | 1 |
| *Years in NFL* | 1008 | 3.72 | 3.30 | 0 | 15 |
| *Games Played* | 1008 | 51.71 | 48.37 | 0 | 239 |
| *Approximate Value per Year* | 1008 | 2.63 | 2.88 | 0 | 14.5 |
| *Log(Approximate Value/Year)* | 1008 | 0.97 | 0.82 | 0 | 2.74 |
| *Pro Rushing Attempts* | 1008 | 378.50 | 595.87 | 0 | 4409 |
| *Pro Rushing YDs* | 1008 | 1530.51 | 2491.06 | 0 | 18355 |
| *Pro Rushing TDs* | 1008 | 10.58 | 18.42 | 0 | 164 |
| *Pro Rushing YDs/Attempt* | 793 | 3.77 | 0.96 | 0 | 12.33 |
| *Pro Rushing TDs/Attempt* | 793 | 0.02 | 0.02 | 0 | 0.2 |
| *College Rushing Attempts* | 1008 | 424.46 | 216.31 | 0 | 1220 |
| *College Rushing TDs* | 1008 | 19.99 | 12.60 | 0 | 77 |
| *College Rushing YDs/Attempt* | 1008 | 4.98 | 0.89 | -5 | 10.3 |

Similarly, Table 22 outlines the summary statistics for the running back group. As with the quarterbacks, all the general player information and performance metrics are included. Position-specific statistics include rushing attempts, yards, and touchdowns at the professional level. The final two are further analyzed to create yards per attempt and touchdowns per attempt metrics. At the college level, rushing attempts, yards per attempt, and touchdowns are included. Table 23 presents the same statistics for pass catchers, but looks at receptions instead of rushing attempts.

## 5.3. Methodology

For each of the three position classes, three sets of models were estimated. In all models for a given position group, the independent variables were the same set of position-specific college statistics. For quarterbacks, the college statistics used are passing attempts, touchdowns per attempt, interceptions per attempt, yards per attempt, and completion percentage. For running

**Table 23** Summary Statistics for Pass Catchers

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| *Year* | 1403 | 1991.89 | 13.26 | 1967 | 2013 |
| *Round* | 1403 | 4.11 | 2.13 | 1 | 9 |
| *Pick* | 1403 | 108.00 | 62.77 | 1 | 222 |
| *All Pro Selections* | 1403 | 0.07 | 0.40 | 0 | 6 |
| *Pro Bowl Selections* | 1403 | 0.28 | 1.07 | 0 | 13 |
| *Primary Starter (Years)* | 1403 | 1.90 | 2.99 | 0 | 16 |
| *Primary Starter (Y/N)* | 1403 | 0.46 | 0.50 | 0 | 1 |
| *Years in NFL* | 1403 | 3.83 | 3.56 | 0 | 17 |
| *Games Played* | 1403 | 53.48 | 52.96 | 0 | 258 |
| *Approximate Value per Year* | 1403 | 2.16 | 2.34 | 0 | 11.33 |
| *Log(Approximate Value/Year)* | 1403 | 0.88 | 0.75 | 0 | 2.51 |
| *Pro Receptions* | 1403 | 109.99 | 175.39 | 0 | 1265 |
| *Pro Receiving YDs* | 1403 | 1489.33 | 2448.04 | -2 | 15292 |
| *Pro Receiving TDs* | 1403 | 9.50 | 16.35 | 0 | 156 |
| *Pro Receiving YDs/Reception* | 1077 | 13.11 | 4.03 | -2 | 46 |
| *Pro Receiving TDs/Reception* | 1077 | 0.08 | 0.09 | 0 | 1 |
| *College Receptions* | 1403 | 95.25 | 58.28 | 0 | 349 |
| *College Receiving TDs* | 1403 | 11.38 | 8.44 | 0 | 60 |
| *College Receiving YDs/Reception* | 1403 | 15.19 | 3.39 | 0 | 36.3 |

backs, the independent variables are rushing attempts, yards per attempt, and touchdowns per attempt. For pass catchers, the models use receptions, touchdowns per reception, and yards per reception. These metrics assess both how often a player was used and how successful a player was when given an opportunity.

The first set of models used draft position as the dependent variable and had two forms, *Round* and *Pick*. These models establish how teams are using college performance to influence drafting decisions.

The second set of models looks at the eight generic, all-purpose statistics used in the previous hypotheses: whether or not a player became a primary starter, years as a primary starter, number of all-pro selections, number of pro bowl selections, years in NFL, games played, and the linear and log forms of approximate value per year. These models allow comparisons between the groups, both in terms of how college performance affects future performance and how similar or different the coefficients of the draft position models are in comparison to the coefficients of the general performance models.

Finally, the final set of models uses position-specific professional statistics as the independent variable. While these models cannot be directly compared between groups, they provide a better picture of what a player provides on the field and how much of that can be known prior to draft day. For quarterbacks, the professional statistics of interest are completion percentage, passing attempts, yards per attempt, touchdowns per attempt, and interceptions per attempt. For running backs, the models use rushing attempts, yards per attempt, and touchdowns per attempt. For pass catchers, the three dependent variables used are receptions, yards per reception, and touchdowns per reception. As with the selection of independent variables, these metrics get at two aspects of on-field performance. Passing attempts, rushing attempts, and receptions indicate how often a player is used. The rest of the statistics are per attempt ratios and observe how successful a player is when used.

If teams were drafting rationally, then each college statistic that was found to improve draft position would also be found to improve some facet of future performance. The reverse, in which a statistic worsens a player's draft position and their future performance, should also hold. Finally, anything that is insignificant in either the draft position or performance models should be insignificant in the other as well, suggesting an understanding of the relationship between draft position and expected future performance. However, these are not the predicted results. Instead, it is predicted that there will be divergences away from these three rules suggesting that the factors that determine where an NFL team drafts a player are not identical to the factors that determine how that player will fare in the NFL.

Finally, should specific statistics be found to be significant predictors in both draft and performance models, there is a question of whether the signals are being properly weighted. Even if teams are picking up on the right cues, it is possible they are not using this information to its fullest extent. Though this possibility is discussed, analysis of such a phenomenon is beyond the reach of this thesis.

## 5.4. Results

Overall, 41 models were tested. For all the position groups, there were strong similarities between the two draft position models and thus only the *Round* models are discussed. Similarly, the general performance metric models were reduced to the three variables used in the previous hypotheses: whether or not a player became a primary starter, games played, and average approximate value per year. Once again, these statistics get at the three larger concepts of a player's professional career – impact, longevity, and value to team.

### *Quarterback Analysis*

The draft position model for quarterbacks finds that college statistics alone cannot account for where a quarterback is taken in the draft. With an $R^2$ value of only 0.082, the model has little explanatory power. The only significant result of note is that more college passing attempts lead a player to be taken earlier in the draft. Even though this result shows statistical significance, its minute size creates questions about economic significance. This result is listed in Table 24.

Table 25 shows a similar story for the general performance metric models. In these cases, none of the independent variables are significant. Logically following from that, the models are all very weak, with no $R^2$ values being higher than 0.052.

**Table 24** Regression Results for Draft Position of Quarterbacks

| VARIABLES | (1)<br>*Round* |
|---|---|
| *College Passing Attempts* | -0.000689** |
| | (0.000344) |
| *College INTs/Attempt* | 5.340 |
| | (10.14) |
| *College Completion Percentage* | -0.0460 |
| | (0.0295) |
| *College TDs/Attempt* | -8.088 |
| | (9.977) |
| *College YDs/Attempt* | -0.206 |
| | (0.203) |
| Constant | 8.784*** |
| | (1.801) |
| | |
| Observations | 395 |
| R-squared | 0.082 |

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

**Table 25** Regression Results for General Performance Metrics of Quarterbacks

| VARIABLES | (1)<br>*Primary Starter* | (2)<br>*Games Played* | (3)<br>*Average Approximate Value per Year* |
|---|---|---|---|
| *College Passing Attempts* | 3.34e-05 | 0.00311 | 0.000534 |
| | (8.15e-05) | (0.00840) | (0.000541) |
| *College INTs/Attempt* | -2.628 | 112.5 | -10.39 |
| | (2.289) | (261.1) | (11.29) |
| *College Completion Percentage* | 0.00314 | -0.337 | 0.0539 |
| | (0.00671) | (0.712) | (0.0386) |
| *College TDs/Attempt* | 1.968 | -267.2 | 6.075 |
| | (2.163) | (208.7) | (11.99) |
| *College YDs/Attempt* | 0.00984 | 8.190 | 0.306 |
| | (0.0462) | (5.144) | (0.256) |
| Constant | 0.172 | 10.76 | -3.355 |
| | (0.426) | (49.43) | (2.335) |
| | | | |
| Observations | 395 | 395 | 395 |
| R-squared | 0.025 | 0.007 | 0.052 |

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

**Table 26** Regression Results for Position-Specific Performance Metrics of Quarterbacks

| VARIABLES | (1)<br>*Pro Completion Percentage* | (2)<br>*Pro Passing Attempts* | (3)<br>*Pro YDs/ Attempt* | (4)<br>*Pro TDs/ Attempt* | (5)<br>*Pro INTs/ Attempt* |
|---|---|---|---|---|---|
| *College Passing Attempts* | 2.06e-05 | 0.400* | 0.000196 | 3.72e-06 | -3.13e-06 |
| | (1.57e-05) | (0.232) | (0.000247) | (3.69e-06) | (8.46e-06) |
| *College INTs/Attempt* | -0.719 | -4,236 | 12.89 | -0.0224 | 0.239 |
| | (0.581) | (6,329) | (9.101) | (0.137) | (0.321) |
| *College Completion Percentage* | 0.00666*** | -15.61 | 0.0797*** | 6.08e-06 | -0.000824 |
| | (0.00216) | (18.75) | (0.0273) | (0.000312) | (0.000612) |
| *College TDs/Attempt* | -0.815 | -4,654 | -3.617 | 0.0680 | -0.0421 |
| | (0.597) | (5,037) | (8.362) | (0.0894) | (0.196) |
| *College YDs/Attempt* | -0.00198 | 186.6 | -0.0458 | -0.000936 | -0.00640 |
| | (0.0143) | (128.4) | (0.185) | (0.00215) | (0.00442) |
| Constant | 0.220* | 561.0 | 1.517 | 0.0318 | 0.141*** |
| | (0.117) | (1,294) | (1.554) | (0.0250) | (0.0459) |
| Observations | 317 | 395 | 317 | 317 | 317 |
| R-squared | 0.131 | 0.013 | 0.045 | 0.007 | 0.039 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In the position-specific models, however, there are some significant results. It was found that an increase of one point in a player's college completion percentage leads to an increase of 0.0067 points in a player's professional completion percentage and an increase of 0.08 yards in the player's professional yards per attempt, on average and ceteris paribus. In addition, a player increases his number of professional attempts by an average of 0.4 with each additional collegiate attempt. These results are displayed in Table 26.

Overall, the quarterback models find positives and negatives in the current drafting decisions of NFL teams. On the positive side, teams are accounting for collegiate number of attempts. Players who throw more in college are both being selected higher in the draft and throwing more in the NFL. The scouts, however, are missing the importance of a player's

**Table 27** Regression Results for Draft Position of Running Backs

| | (1) |
|---|---|
| VARIABLES | *Round* |
| *College Rushing Attempts* | -0.00311*** |
| | (0.000298) |
| *College YDs/Attempt* | -0.434*** |
| | (0.114) |
| *College TDs/Attempt* | -8.973*** |
| | (2.776) |
| Constant | 8.004*** |
| | (0.576) |
| Observations | 1,006 |
| R-squared | 0.145 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

completion percentage in college. Currently, this is not factored into most drafting decisions, yet it is a significant predictor of future success.

### *Running Back Analysis*

In contrast to the quarterback model, the draft position model for running backs finds each of the three key independent variables to be extremely significant. An increase in college rushing attempts, yards per attempt, or touchdowns per attempt improves a player's draft position, on average and ceteris paribus. Table 27 outlines this result.

Table 28 shows that these metrics are also significant predictors of the general performance metrics. For *Games Played* and *Approximate Value per Year*, all three college statistics predict future performance in a positive way. While rushing attempts and touchdowns per attempt are also significant in the primary starter model, yards per attempt is not.

Finally, the position-specific models provide a mixture of results. An increase in any of the three college statistics leads to an increase in a player's professional attempts, on average and ceteris paribus. Collegiate success with attempts and yards per attempt also improves a player's

**Table 28** Regression Results for General Performance Metrics of Running Backs

| VARIABLES | (1)<br>*Primary Starter* | (2)<br>*Games Played* | (3)<br>*Average Approximate Value per Year* |
|---|---|---|---|
| *College Rushing Attempts* | 0.000309*** | 0.0403*** | 0.00365*** |
| | (7.22e-05) | (0.00737) | (0.000400) |
| *College YDs/Attempt* | 0.0353 | 4.653*** | 0.599*** |
| | (0.0230) | (1.694) | (0.126) |
| *College TDs/Attempt* | 1.510** | 111.1* | 8.318** |
| | (0.693) | (66.46) | (3.543) |
| Constant | 0.0619 | 6.131 | -2.302*** |
| | (0.115) | (8.604) | (0.648) |
| Observations | 1,006 | 1,006 | 1,006 |
| R-squared | 0.030 | 0.048 | 0.132 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

**Table 29** Regression Results for Position-Specific Performance Metrics of Running Backs

| VARIABLES | (1)<br>*Pro Rushing Attempts* | (2)<br>*Pro YDs/Attempt* | (3)<br>*Pro TDs/Attempt* |
|---|---|---|---|
| *College Rushing Attempts* | 0.754*** | 0.000506*** | 4.89e-06 |
| | (0.0958) | (0.000170) | (4.10e-06) |
| *College YDs/Attempt* | 105.6*** | 0.198*** | 0.000180 |
| | (22.66) | (0.0513) | (0.00112) |
| *College TDs/Attempt* | 1,738*** | -1.729 | 0.0139 |
| | (654.0) | (1.847) | (0.0472) |
| Constant | -550.9*** | 2.619*** | 0.0204*** |
| | (118.9) | (0.278) | (0.00558) |
| Observations | 1,006 | 792 | 792 |
| R-squared | 0.120 | 0.042 | 0.003 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

yards per attempt in the NFL. Finally, none of the college statistics are significant predictors of a player's professional touchdown per attempt rate. These results are provided in Table 29.

In general, it appears that scouts and GMs are using successful representations of running backs when they draft. Each of the three college statistics used in this regression both significantly affected draft position and future performance. Further analysis would be needed to

**Table 30** Regression Results for Draft Position of Pass Catchers

| VARIABLES | (1)<br>*Round* |
|---|---|
| *College YDs/Reception* | -0.0844*** |
| | (0.0203) |
| *College Receptions* | -0.0100*** |
| | (0.000888) |
| *College TDs/Reception* | -2.100* |
| | (1.163) |
| Constant | 6.608*** |
| | (0.277) |
| Observations | 1,396 |
| R-squared | 0.099 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

see if the teams are weighting each of the three statistics correctly in regards to draft position and future performance. There is still more to the story, however, as the $R^2$ values in these models were generally low. The draft position model set the highest mark with a value of 0.145, while the performance metric models ranged from 0.003 to 0.132. Thus, even though the teams are keying into to valuable signals, there is still more to the story of predicting future success for a running back.

## *Pass Catcher Analysis*

The results for pass catchers mirror the results for running backs. Starting with the draft position model, all three of the college statistics (receptions, yards per reception, touchdowns per reception) significantly improve draft position. For precise numbers, see Table 30.

In the general performance metric models, each independent variable has mixed significance. In determining whether or not a player becomes a primary starter, an increase in either college receptions or touchdowns per reception improves a player's odds while yards per reception has no meaningful effect. In contrast, yards per reception is the only predictor of games

**Table 31** Regression Results for General Performance Metrics of Pass Catchers

| VARIABLES | (1)<br>*Primary Starter* | (2)<br>*Games Played* | (3)<br>*Average Approximate Value per Year* |
|---|---|---|---|
| *College YDs/Reception* | 0.00477 | 0.977** | 0.0967*** |
| | (0.00449) | (0.473) | (0.0187) |
| *College Receptions* | 0.000641*** | 0.0361 | 0.00873*** |
| | (0.000230) | (0.0229) | (0.00109) |
| *College TDs/Reception* | 0.764*** | 14.88 | 2.328** |
| | (0.224) | (21.57) | (0.923) |
| Constant | 0.231*** | 33.35*** | -0.427 |
| | (0.0686) | (7.306) | (0.281) |
| Observations | 1,396 | 1,396 | 1,396 |
| R-squared | 0.019 | 0.006 | 0.073 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

**Table 32** Regression Results for Position-Specific Performance Metrics of Pass Catchers

| VARIABLES | (1)<br>*Pro Receptions* | (2)<br>*Pro YDs/Reception* | (3)<br>*Pro TDs/Reception* |
|---|---|---|---|
| *College YDs/Reception* | 6.997*** | 0.473*** | 0.000863 |
| | (1.427) | (0.0456) | (0.000948) |
| *College Receptions* | 0.490*** | 0.00351* | -9.30e-05** |
| | (0.0732) | (0.00198) | (4.15e-05) |
| *College TDs/Reception* | 106.5* | 0.171 | 0.0277 |
| | (61.92) | (2.526) | (0.0526) |
| Constant | -56.39*** | 5.500*** | 0.0740*** |
| | (21.33) | (0.670) | (0.0170) |
| Observations | 1,396 | 1,071 | 1,071 |
| R-squared | 0.047 | 0.149 | 0.006 |

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

played in the NFL. Finally, an increase in any of the three college metrics leads to an improvement in a pass catcher's *Average Approximate Value per Year*, on average and ceteris paribus. Table 31 summarizes these findings.

The results are equally inconsistent in the position-specific performance metric models, as shown in Table 32. A player's college receptions significantly affect his professional receptions, yards per receptions, and touchdowns per receptions, on average and ceteris paribus. While the first two of these results are in the expected direction, college receptions have an inverse relationship with touchdowns per reception at the NFL level. A player's yards per reception in college also significantly improves the player's number of receptions and yards per reception in the pros. It does not, however, affect a player's touchdowns per reception. Finally, the collegiate touchdowns per reception of a player significantly increases the number of receptions a player will have professionally, but does not affect the yards or touchdowns per reception.

In mirroring the results of the running backs, this analysis suggest that NFL teams are drafting pass catchers based on signals that are significantly correlated with future performance. As before, further analysis is required to see if the weighting of these signals is optimal. Further research is also needed to determine the other predictors of future success, since the $R^2$ values found in these models are relatively low.

## 5.5. Discussion

The Signal and Noise hypothesis contends that teams will draft players based on internal models of what it takes to make it and succeed in the NFL. Via the representativeness heuristics, teams will draft players who fit their vision of an NFL player, despite the fact this might be a faulty predictor of future performance. Since the possibility also exists that teams do not fully understand what does and does not lead to success, teams may further be hurting themselves by drafting players based on an inaccurate representation. If this occurs, teams are using a suboptimal drafting strategy due to internal decision-making processes.

The present analysis examined this hypothesis in regards to three categories of players: quarterbacks, running backs, and pass catchers. The most striking results occurred in the quarterbacks' analysis. The results show that the main predictor of draft position is a player's number of passing attempts in college. Teams are correct in drafting based off this metric, as it significantly predicts how much a player will throw professionally. There is, however, another college statistic that is a true signal of future performance, namely completion percentage. Teams are not factoring this into their representations and therefore are leaving a signal unrecognized.

Drafting performance is better for both running backs and pass catchers. In these two classes, scouts are picking up key signals in all of a player's college statistics. These statistics, in turn, each predict improved future performance. The key question remaining for these position groups are whether or not general managers are accurately weighting each signal.

Overall, it appears as though there are internal models that NFL teams used to draft players. In terms of the factors measured in this thesis, however, these representations include the significant signals. The only signal currently being missed in a quarterback's college completion percentage. This thesis finds no evidence that teams are drafting based on a false signal. These results suggest that while the representativeness heuristic might be influencing decision, it is doing so in a predominantly positive way.

## 6. General Discussion

Overall, this thesis finds that teams are using a variety of heuristics when making draft decisions. All three models found some form of heuristic use in the selection of players. The results are much more split when it comes to the effect this has on player performance. The Talented Teammate findings suggest that teams are suffering due to irrational drafting, while the Geographic Bias and Signal and Noise results suggest that these trends do not adversely affect teams.

Given the presence of cognitive biases and the potential for detrimental effects, teams would be wise to implement strategies that would help them avoid such errors in the future. Psychological research of overcoming heuristics suggests that there are strategies individuals and groups can use to make more rational decisions.

Ferreira, Garcia-Marques, Sherman and Sherman (2006) found that there were four distinct ways to affect what processes an individual uses when making a decision, three of which are applicable here. The first is to change the processing goal of the individual. They were able to do this by defining their study as either a study of intuition or a study of rationality. Thomas and Millar (2012) find similar results when instructing participants to either "think like a gambler" or "think like a scientist". Second, decision makers use more rational processes when they have increased cognitive resources. People whose minds are busy working on other problems will use heuristics more often to solve the task at hand. Third and finally, decision-makers can be trained to respond more logically. This can be done either by priming the participant with logic or statistics-based problems (Ferreira et al., 2006; Thomas & Millar, 2012) or by directly teaching individuals about statistical laws (Kosonen & Winnie, 1995). A key finding is that such training can be successful both when it is field-specific and when it is general.

Hirt and Markman (1995) found a fourth applicable strategy. They instruct decision-makers to use a "consider-an-alternative" strategy when they are potentially biased in favor of one answer. By considering an alternative, even if that alternative is not directly opposite the original answer, individuals make available competing arguments. They then are forced to use more rational methods for deciding on a final outcome.

Each of these four strategies can be applied to the world of the NFL. First, an NFL team can change the goals of their football operations department by creating an analytics department. Outside of the benefit that the analytic department's research will provide, merely creating the department will change the mindset of the scouts and general manager. It will move the team from an intuition, "gut"-based process to one that is focused on making sound, rational decisions.

Second, teams should hire more scouts and analysts. This would reduce the cognitive workload of each individual. Instead of rushing through as many players as possible, scouts would be able to properly focus on a potentially specialized subset of players and devote their whole focus to justly analyzing those players.

Third, teams should mandate training sessions for their employees. These training sessions should be both about rational decision making, including logic and statistics, as well as biased decision making, such as heuristics. The former will teach individuals how to go about making strong decisions while the later will inform them about where errors are possible. This last effect will put them on their guard and help protect them from unknowingly using heuristics.

Fourth and finally, teams should design their scouting reports so that they require a consider-an-alternative approach. Instead of boiling down a player to one number or grade, teams should mandate that scouts provide either a range for a player's score or a worst-case, average-case, and best-case score. These systems will force scouts to consider multiple outcomes and will trigger the use of natural analytic processes in judging players.

These many recommendations would be unnecessary, however, if biases were not present. Though this thesis contends that the results found here indicate irrationality, there are two significant alternative explanations.

First, it is possible that there is heuristic use but that it is rational, given the environment. Gigerenzer and Gaissmaier (2011) argue that situations exist in which heuristics lead to better, or at least equivalent, quality of decisions than more formal statistical models. This can happen for two reasons. One, statistical models are designed to optimize fit on previous data. However, decisions have to deal with predictions. It may be that statistical models overvalue aspects of the previous data that improves the overall fit of the model but reduces its predictive power. Heuristics, on the other hand, use only the most noteworthy or important information to make a decision and therefore avoid being biased by past outcomes. The second reason has to deal with the cost of making a decision. Even if using analytical thinking leads to better outcomes, it costs the decision maker something. If the benefit of using such thinking does not outweigh the cost, an individual is better off using a heuristic to come to a satisfactory decision. This understanding of heuristics may explain why heuristic use was consistently found in the present study, but detrimental effects were inconsistent. It is possible that teams are using heuristics, but it is actually rational for them to do so.

The second possible explanation concerns the goals of NFL teams. This thesis has assumed that the goal of a franchise is to maximize wins and on-field performance. Anything that took away from these fields was defined as irrational. It can be argued, however, the teams are not win-maximizing entities, but profit-maximizing businesses. Using this understanding, actions that have been labeled illogical may in fact make sense. For the Talented Teammate hypothesis, it may be that having players who are connected to more famous early picks improves revenue via ticket or merchandise sales. The same case can be made for the Geographic Bias. Teams may be trying to entice fans of local teams to spend their money on NFL entertainment. Finally, the statistics that were defined as signals in the final analysis could also be construed as measures of college fame. Intentionally bringing in well-known college players could help the bottom line and be seen as rational for a profit-maximizing entity. Without knowing whether teams aspire to increase profits or wins, it is impossible to definitively argue what makes for rational behavior in the NFL draft.

Even without these alternative explanations, there were limitations to this thesis that must be considered. First, the performance metrics used throughout the thesis were fairly weak. They were used to capture large aspects of a player's career but failed to precisely measure what someone brought to a team. There were also hypothesis-specific issues. For the Geographic Bias, the use of states as the definition of local failed to account for teams near borders. For the Signal and Noise hypothesis, the results are limited by the sample size, the quality of college statistics, and the inability to measure how teams should weight signals. A general concern present throughout this thesis is the relatively small effect size and explanatory power. Though many significant results were found, it is unclear how significantly these biases are actually affecting team outcomes. Finally, the recommendations provided above for reducing heuristic use are grounded in previous literature, but are not tested within the realm of professional football. Without such tests, it cannot be confirmed that they would benefit teams. In sum, though this thesis presents many interesting findings, there are still limitations that require understanding before any prescriptive action is taken.

Given these limitations and the questions still remaining, there are many areas for future research. First, there are several ways to strengthen the results found here. Similar analyses could be run using position-specific statistics as the dependent variable. This would help confirm whether or not these biases are causing suboptimal performance at various positions. The Geographic Bias hypothesis could be reworked using a college's mileage from a team's home to determine if that school is considered local. Further processing could be done on the college statistics data set to increase the sample size for the Signal and Noise hypothesis. Moving beyond the questions asked in this thesis, there is a great deal of work still to do relating cognitive heuristics to strategies in the NFL and sports in general. In relation to the draft, further hypotheses could be created directly related to team-specific scouting processes. Understanding how a football operations department functions is critical to the creation of legitimate questions. Outside of the draft, there are many ways that cognitive biases could affect football decisions. One could apply prospect theory and the idea of risk-seeking and risk-averse behavior to a coach's in-game decision about going for it on fourth down or punting. A team's training regimen could be analyzed to determine if the status quo bias is leading to suboptimal practices. A study of trades in the NFL could examine the potential for the endowment effect as an impediment to maximizing team quality. These are but a few examples of the wide range of topics still to be explored.

## 7. Conclusion

The purpose of this thesis was to examine drafting strategies in the National Football League and determine if teams were drafting inefficiently due to the use of cognitive heuristics. The results suggest that biases are affecting drafting decisions in a number of ways. However, it is unclear to what extent, if any, these biases are creating inefficiencies in the NFL draft. There remains a great deal of work to do, both in further analyzing the hypotheses presented here and exploring other mechanisms through which cognitive biases may be affecting football decisions.

## References

Berri, D.J. & Simmons, R. (2011). Catching a draft: On the process of selecting quarterbacks in the National Football League draft. *Journal of Productivity Analysis, 35*(1), 37-49.

Clayton, J. (2011, July 25). What new CBA means in football terms. *ESPN.com*.

Collective Bargaining Agreement (2011). Retrieved from http://images.nflplayers.com/media Resources/files/PDFs/General/2011_Final_CBA_Searchable_Bookmarked.pdf

Constitution and Bylaws of the National Football League (2006). Retrieved from http://static.nfl. com/static/content/public/static/html/careers/pdf/co_.pdf

Ferreira, M. B., Garcia-Marques, L., Sherman, S. J., & Sherman, J. W. (2006). Automatic and controlled components of judgment and decision making. *Journal of Personality and Social Psychology, 91*(5), 797-813.

Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology, 62*, 451-482.

Kosonen, P., & Winne, P. H. (1995). Effects of teaching statistical laws on reasoning about everyday problems. *Journal of Educational Psychology, 87*(1), 33-46.

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

Hirt, E. R., & Markman, K. D. (1995). Multiple explanation: A consider-an-alternative strategy for debiasing judgments. *Journal of Personality and Social Psychology, 69*(6), 1069-1086.

Massey, C., & Thaler, R. H. (2013). The loser's curse: Decision making and market efficiency in the national football league draft. *Management Science, 59*(7), 1479-1495.

McGee, K.J. & Burkett, L.N. (2003). The National Football League combine: A reliable predictor of draft status? *Journal of Strength and Conditioning Research, 17*(1), 6-11.

Mirabile, M.P. (2008). Intelligence and football: Testing for differentials in collegiate quarterback passing performance and NFL compensation. *TheSportJournal.org*.

Moxley, J. H., Ericsson, K. A., Charness, N., & Krampe, R. T. (2012). The role of intuition and deliberative thinking in experts' superior tactical decision-making. *Cognition, 124*(1), 72-78.

Pachur, T., & Marinello, G. (2013). Expert intuitions: How to model the decision strategies of airport customs officers? *Acta Psychologica, 144*(1), 97-103.

Pretz, J. E. (2008). Intuition versus analysis: Strategy and experience in complex everyday problem solving. *Memory & Cognition, 36*(3), 554-566.

Pro-Football-Reference. Retrieved from http://pro-football-reference.com/

Quinn, K.G., Geier, M., & Berkovitz, A. (2007). Passing on success? Productivity outcomes for quarterbacks chosen in the 1999-2004 National Football League player entry drafts. Working Papers 0711, International Association of Sports Economists & North American Association of Sports Economists.

Sports Reference – College Football. Retrieved from http://www.sports-reference.com/cfb/

Spotrac. Retrieved from http://spotrac.com/

Stanovich, K.E. (2012). Environments for fast and slow thinking. *Trends in Cognitive Sciences, 16*(4), 198-199.

Thomas, A. K., & Millar, P. R. (2012). Reducing the framing effect in older and younger adults by encouraging analytic processing. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 67B*(2), 139-149.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 195*, 1124-1131.

US News Best Colleges. Retrieved from http://colleges.usnews.rankingsandreviews.com/best-colleges/sitemap

Whitaker, K. (2013). Does a 'coattail effect' influence the valuation of players in the Major League Baseball draft? Paper presented at the MIT Sloan Sports Analytics Conference, Boston, MA.