The Returns to Elite Sports Programs: Signaling or Value-Added

Jordan Holbrook *

September 5th, 2024

Please click here for latest version.

Abstract

This study constructs a novel panel dataset of highly recruited high school athletes, analyzes their participation in college sports programs and subsequent professional sports career outcomes. Utilizing the matched applicant approach, or Dale and Krueger method, that exploits variation in enrollment decisions conditional on similar offer-sets, I estimate selection-corrected returns to attending an elite college sports programs on job placement in the NFL. The findings reveal that student-athletes from top-ranked football programs are significantly more likely to become professional athletes, with a one standard deviation increase in sports program ranking raising the likelihood of being drafted by 32% of the mean. The paper further explores whether these returns align with a human capital or signaling framework, concluding that the evidence supports the latter, particularly with heterogeneous effects across different position groups.

Key words: Education, Human capital, Signaling

JEL Codes: I21, J31, J45, O15

*University of Houston, Science and Research 2, Office 133, 3623 Cullen Boulevard, Houston, TX 77204. Email: jcholbrook@uh.edu. Grateful to Professor Chinhui Juhn and Professor Yona Rubinstein for their courses in labor economics and the economics of education; Professors Aimee Chin and Vikram Maheshri for their feedback and insight. I am thankful for comments, feedback, and help with programming from: Angelo Santos, Horacio Rueda, Yujie Zhang, & Charter Sevier.

1 Introduction

Each year over 1.2 million high school students compete in the most popular sport in the United States, American Football, many with the ambition of becoming a professional athlete. Yet, similarly, each year slightly over 250 athletes are selected by professional teams into the National Football League (NFL). In examining the college sports programs that professional athletes come from I find that almost 80% of athletes selected in the NFL draft come from only 20% of college sports programs. What sets these elite programs apart from other sports teams? This study focuses on measuring the individual returns to participation in elite college sports programs in terms of professional sports labor market outcomes.

A deeply rooted American ideology is that of meritocracy and ostensibly with sports being one of the most meritocratic institutions available. At every level of sports (high school, college, professionals) team incentives are aligned to field the most talented athletes to maximize the chances of winning. Thus, the best athletes should be recognized and "promoted" onto professional teams regardless of the college program in which they participated in, implying negligible program value added. This view contrasted with the evidence that a significant proportion of professional NFL athletes come from a small number of college teams motivate my central research questions: Does the 'quality' of the athlete's collegiate program matter in determining who gets selected to become a professional athlete? If the quality of a college sports program does indeed has an impact, can these returns be explained by a human capital or signaling framework?

I study these questions by building a novel panel dataset containing the following data: (1) high school athletes and measures of performance and athletic ability; (2) college football recruiting information and scholarship offer sets from the largest sports network in the US (ESPN); (3) measures of college sports program rankings and individual athlete collegiate performance measures; (4) post college data from the NFL. This dataset offers detailed longitudinal data on a wide range of pre-college attributes as well as post-college outcomes. Additionally, in contrast to many other research settings studying the effect of college programs on students, this dataset offers an abundance of tasked based performance measures while a student athlete is in college. These tasked-based performance measures are conveniently measured in the same units as the task-based performance that determines productivity and pay as a professional athlete.

Even with this unique dataset answering these research questions involves addressing some initial challenges. For more than 50 years economists have been interested in studying the returns to program participation. 'Programs' such as schools, colleges, or training programs have all generated speculation on obtaining accurate estimates of the return to participation. Obtaining unbiased estimates of the effects of college characteristics on student outcomes is particularly challenging as it involves understanding the effect of selection on outcomes for both observable and unobserved characteristics of students. The effect of selection on observable characteristics can be nuanced, with questions such as how predictive are test scores of students earning potential or are returns to education heterogeneous across racial or ethnic groups? Furthermore, the effect of selection on unobserved characteristics is a significant confounding factor as unobserved characteristics can influence both where students attend college, and subsequently where they are employed and how much they earn. Studies in the economics literature have shown divided results on this and similar topics.

In anticipation of many of the challenges of this type of research, this study proposes a focus on student athletes with the potential to add both simplicity and new data to this research setting. There is an abundance of data and measures on student athletes. Sports in and of itself is a data dominated industry and beginning in high school and on to college and the professional leagues, there are many objective performance measures that can be leveraged to address key questions.

In this paper, I build a novel panel data set of highly recruited high school athletes, observing the characteristics of the college teams they participate in as well as their professional athlete labor market outcomes. First, using data on these top high school athletes and the college teams they were extended scholarship offers I employ a dependable empirical strategy, known as the matched applicant or Dale and Krueger method, to address threats to identification from selection on unobservable characteristics. With this data and strategy, I provide selection-corrected estimates of the returns to elite sports programs on the main outcome variable of job placement. In contrast to Dale and Krueger (2002, 2011) but consistent with Chen, Grove, & Hussey (2012) I find substantial returns to higher quality college teams in terms of initial job placement. Student athletes that participate in top ranked college football programs are three to five times more likely to be employed professionally. A one standard deviation increase in college sports program ranking increases the likelihood of being drafted by 32% of the mean. I then show there are large heterogeneous effects for student athletes in the same program but in different position groups. The effect of college sports program quality is least important for Quarterbacks (QBs) while most important for offensive lineman (OG, OC, OT). These effect of an elite program is almost eight times larger for linemen than quarterbacks. Motivated by this large heterogeneity I turn to theoretical explanations of these effects. I examine if the large returns to elite sports programs are consistent with a human capital value added framework or a signaling framework. Using unique information available from college football performance data, I test implications of both frameworks by building a simple search model. I find the returns to elite sports programs are more consistent with a signaling model and that the signaling model has ex-ante predictions for the heterogeneous effects by position group.

I contribute to the literature in three key areas. First, I assemble one of the largest datasets of high school athletes to my knowledge, providing a useful resource for analysis. Second, I use unique data from college football to test signaling theory, offering new insights into how performance and group affiliation impact career outcomes. Third, I am the first to measure returns to elite sports programs at the extensive margin, specifically examining who reaches the NFL draft. My findings estimate a significant premium for elite programs in terms of job placement in the NFL.

The remainder of this paper is organized as follows. Section 2 describes the various data sources used to build the panel dataset of student athletes. Section 3 describes the methods and empirical strategy I employ to measure the payoffs to elite sports programs. Section 4 reports the results and main findings, and section 5 investigates the theoretical mechanisms underlying the results and section 6 summarizes and concludes.

2 Data

In this section, I report my data sources as well as the context of the study. I then define my sample and relevant variables of interest. Finally, I introduce summary statistics for the sample of student athletes and exhibit characteristics of the college football program in which these athletes participate.

2.1 High School Athletes

Starting in 2006 the largest sports network in the United States, the Entertainment and Sports Programming Network (ESPN), started collecting data and evaluating high school football athletes from all over the country. Professional scouts, analysts, and coaches employed by ESPN reviewed game film on top high school players and assigned each player a recruiting grade and national rank. These metrics were meant to assess the readiness of the high school player to compete at the collegiate level as well as a measure of the athletic skill and talent of the individual. This data has been recorded for each high school graduating class since 2006. Additionally, each high school student athlete in the database has a profile page with a detailed scouting report, recruiting activity, and player news in the media.

Along with detailed athletic ability information, the ESPN database consists of information on athletic scholarship offers. The player profile page lists each college football program that has extended an official scholarship offer. Additional information includes the status of the scholarship offers, i.e., whether the offer was accepted or not as well as if the athlete participated in an official campus visit. Other information included the student athlete's hometown and high school. I collect scholarship information on each athlete including the total number of scholarships offered, scholarship offers in athlete's home state, and which scholarship offer was ultimately accepted. This information is vital to my eventual empirical strategy.

The culmination of this high school student athlete information came to be known as the ESPN 300 and this publicly available data is displayed at www.espn.com. In subsequent years ESPN expanded the athlete rankings from only the best 300 high school players but ranked the top 100 players for each of the 16-18 position groups in American Football. I employ several web-scraping and data mining approaches to collect this public information and display it in a database appropriate for econometric analysis. The high school data set has on average 1,600 athletes graded by ESPN analysts for the years 2006-2022.

Table 2 provides statistical information on various characteristics of high school athletes. On average, the ESPN 300 high school ranking for these athletes is 46.42, with a standard deviation of 28.69. The range of rankings spans from the top-ranked athletes at 1 to the lowest-ranked athletes at 100. In terms of grades, the ESPN 300 high school athlete grade averages at 77.03, with a standard deviation of 4.49. The grades range from a minimum of 44 to a maximum of 95. This indicates that these athletes, as a whole, tend to be highly ranked and athletically talented.

2.2 Sample Construction & Data Sources

I construct a panel dataset starting with high school student athletes and following them into their collegiate and professional careers by linking six sources of data: (1) top high school athlete profiles from the ESPN 300 recruiting database 2006-2021, ESPN.com; (2) college football program rankings, Sports-reference.com; (3) college athletic department financial data, Equity in Athletics Disclosure Act EADA 2000-2021; (4) individual athlete college football performance statistics, collegefootballdata.com API 2000-2021; (5) data on professional athletes in the NFL 2000-2021, profootball-reference.com; (6) professional athlete salaries and contracts, spotrac.com. The target sample is all high school football players with a recruiting profile in the ESPN 300 database from 2006-2021 that can be linked to a college football program roster. High school athletes that cannot be linked to a college roster or college football programs that do not have a college football program ranking (e.g. new team with no historical performance) are dropped from the analysis sample.

2.3 Ranking College Football Programs

One of the unique challenges of this study is defining a metric evaluating college football program rankings. Ranking team and program performance has been the fixation for sports fans and analysts as long as sports teams have existed, and college football is no exception. Many of the large television and sports network providers have their own proprietary ranking of teams each season. There are many ways to measure college football program quality, however, a measure with two key attributes, time invariance and stability, are important for a multitude of reasons.

When comparing college football programs, one of the challenges lies in the fact that different programs have vastly different histories, with some teams playing for as few as 10 years, while others have been around for 50 or even 100 years. Additionally, these programs often compete in different conferences or leagues, each with varying levels of competitiveness. For these reasons, a ranking metric must be designed in a way that allows for meaningful comparisons across such diverse leagues with large variation in competitiveness.

This sports program ranking metric can be compared to the common selectivity measure of average SAT entrance scores used in the economics of education literature. Both metrics aim to provide a standardized way to assess the quality of institutions—sports programs in one case and academic institutions in the other—by relying on consistent, comparable data.

Time invariance in both cases is key. Just as the average SAT score provides a stable measure of a college's selectivity across different admission cycles, a time-invariant football ranking allows for comparisons of program strength over decades, unaffected by short-term fluctuations. This stability is crucial for tracking long-term trends and for understanding whether a strong performance is part of a lasting tradition or a brief peak.

Instead of developing an original metric to measure college sports team quality, I turn the sports analytics industry and use a well known rating system for American football teams. The Simple Rating System metric, is a least squares rating method developed by Massey (1997), estimates team ratings based on game outcomes, focusing on predicting the expected margin of victory between teams. The key assumption is that the expected margin of victory between two teams A and B is proportional to the difference in their ratings:

$$E[Y] = r_A - r_B \tag{1}$$

$$y_i = r_A - r_B + e_i \tag{2}$$

The observed outcome for each game includes a random error term, so the actual outcome of game i is modeled as equation (4). To estimate team ratings, for a given team outcomes are aggregated for all games in a season across all opponents. Then aggregated once more across seasons. The matrix of games and opponents can be written as X and to estimate team ratings, solve the normal form equations with an added scaling constraint r.

$$X^T X r = X^T y \tag{3}$$

$$\sum_{i=1}^{n} r_i = 0 \tag{4}$$

The least squares solution for the team ratings β is given by equation (7) and equation (8) decomposes the contributions of strength of schedule and average margin of victory in the rating formula.

$$\beta = (X^T X)^{-1} X^T y \tag{5}$$

$$\beta = \underbrace{(X^T X)^{-1}}_{\text{Strength of Schedule}} \cdot \underbrace{X^T y}_{\text{Average Margin of Victory}} \tag{6}$$

In the above equation: $(X^T X)^{-1}$ represents how the matrix X accounts for the matchups and adjusts for the strength of the teams' schedules. - $X^T y$ reflects the average margin of victory, as y is the vector of margins for each game and X^T sums these results for each team. So, β incorporates both the strength of the teams' schedules and their average performance (margin of victory), providing a comprehensive rating. For full details on this method see Massey (1997) and Meyers (1992). While these papers article the method for how to compute a team ranking system, the exact variables used in to create the SRS metric from Sports-Reference.com are proprietary as each sports analytics website, ESPN.com, NCAA.com, and etc potentially add in additional variables such as home field advantage or overtime weights to make their team ratings more precise.

I merge the high school athlete's dataset with another publicly available online database, sports-reference.com. Sports-reference.com is a premier online database for most collegiate and professional sports. I collect college characteristics for the teams where high school players were recruited, including information on the number of wins and losses for each team, team strength of schedule, and conference championships won. Table 2 reports the main college characteristics. Sports-reference.com has been used in other economic studies, including Foltice and Markus (2021) , Keefer (2016) , and Keefer (2017).

While there are many ways to evaluate college team quality, the SRS metric metric the useful properties discussed previously: time invariant— teams can be compared in terms of their SRS regardless of the number of years a college program has participated in college football, uniform across divisions – college football in the US has several tiers of leagues in which teams compete (Division I, Division II, Division III, etc.) under the SRS metric teams in different leagues can be compared, and finally stability – SRS is a relatively stable quality metric that changes little from year to year. Additional quality metrics are evaluated in the later section on robustness checks, including a discussion of the sensitivity of the findings to each quality ranking.

Figure 1 compares college football programs across four SRS (Simple Rating System) tiers, from -20 to greater than 10, highlighting team performance in terms of win percentage, bowl appearances, and conference championships. Lower-tier teams (e.g., Kent State and Massachusetts in the -20 to -10 tier) exhibit lower win percentages, fewer bowl appearances, and limited championship success. Mid-tier teams (e.g., BYU, Cincinnati, and Boise State) show moderate success across these metrics, while high-tier programs (e.g., Alabama, Michigan, and Ohio State in the > 10 tier) dominate with high win percentages, frequent bowl appearances, and numerous conference championships. The overall trend illustrates that teams with higher SRS ratings tend to perform better across all metrics.

2.4 Measuring Athletic Performance

In this section, I discuss the structure and methods used for evaluating athletic performance across various positions and categories within a college football team. The performance of athletes is measured using a wide array of statistical performance measures that are standardized and aggregated into composite scores.

2.4.1 Team Structure and Performance Measures

The team is structured into 33 unique positions, divided into 3 units and further grouped into 10 position categories. Each position category has between 3 to 6 distinct performance measures, with the exception of the **Offensive Lineman** category, which has no official performance measures recorded at the collegiate level. Another notable exception is the lack of standardized defensive performance measures for 11 defensive positions before 2016. For 9 out of the 10 position groups, established performance measures offer more than 1 million potential combinations for analysis.

The relevant performance measures for each category are outlined in Table 1. Defensive categories include measures such as *QB HUR*, *SACKS*, and *TFL*, while offensive categories feature metrics like *YDS*, *TD*, and *YPC*. These measures are critical in assessing player contributions in different aspects of the game.

Standardization of Performance Measures

To ensure comparability across different positions and categories, performance measures are standardized. The standardized score Z_{ijg} for player i, measure j, and category g is calculated as follows:

$$Z_{ijg} = \frac{X_{ijg} - \mu_{jg}}{\sigma_{jg}}$$

Where X_{ijg} represents the raw performance measure for player i, and μ_{jg} and σ_{jg} are the mean and standard deviation of measure j within category g.

Composite Score by Category

Within each category, standardized scores are aggregated into a composite score C_{ig} for player i, using the following formula:

$$C_{ig} = \frac{1}{n_g} \sum_{j=1}^{n_g} Z_{ijg}$$

Where n_g is the number of performance measures in category g. This score reflects the player's overall performance within a specific category.

Overall Player Score

To obtain an overall performance score for each player, the category-specific scores are averaged across all categories:

$$C_i = \frac{1}{G} \sum_{g=1}^{G} C_{ig}$$

Where G represents the number of categories in which the player has recorded performance measures.

Seasonal Aggregation

Player performance over multiple seasons is aggregated by averaging the overall player scores across all seasons:

$$S_i = \frac{1}{m} \sum_{s=1}^{m} C_{is}$$

Where m represents the number of seasons. This seasonal aggregation allows for a better evaluation of a player's performance over consecutive seasons.

2.5 College Football Recruiting

The college recruiting process is a structured method that coaches use to identify, evaluate, and eventually recruit student-athletes. It typically begins with coaches gathering a large pool of prospective recruits through recruiting websites, third-party services, recommendations from high school coaches, and showcases. From there, coaches narrow down the list by sending recruiting letters, questionnaires, and camp invites to athletes who meet basic requirements, such as height, weight, and academic performance. As athletes respond, coaches begin in-depth evaluations that focus on both athletic and academic abilities, as well as character, to create a ranked list of top prospects. This list continues to shrink as coaches conduct further assessments, including calls with high school coaches and watching athletes compete in tournaments or at camps.

Once coaches have a final list, they extend verbal offers and scholarships to their top recruits, aiming to fill open roster spots. The final step involves recruits signing official offers and ensuring they meet eligibility requirements. Throughout this process, athletes must be proactive, sending updated performance videos, contacting coaches, and maintaining strong academic records to ensure they remain eligible to compete at the college level. Recruiting timelines vary by sport and division, but student-athletes are encouraged to start the process early, build relationships with coaches, and be prepared to make decisions about scholarships and offers when the time comes.

The timeline for when college coaches can officially contact athletes and when athletes need to sign their offers is governed by NCAA recruiting rules, which vary by sport and division level. For most Division I and Division II sports, coaches can start proactively reaching out to recruits on June 15 after their sophomore year or September 1 of their junior year. However, student-athletes can begin reaching out to coaches earlier, sending emails, video, and academic transcripts, though coaches may not respond until the official contact period begins. Once offers are extended, athletes have two primary signing periods to formalize their commitment. The **Early Signing Period** typically occurs in November of an athlete's senior year and allows those who have already decided on a college to sign early. The **Regular Signing Period** begins in April of the senior year and extends into the summer, giving athletes more time to finalize their college decision if they didn't commit early. These deadlines are important to keep in mind as athletes progress through the recruiting process.

Unlike regular students, who typically apply to a few colleges and can introduce selection through their choice of schools based on academic fit, athletes only apply for admission after they have officially accepted a scholarship offer from a college. This potentially mitigates some of the selection bias that can occur in the general college application process, where students self-select into certain schools based on various factors, such as perceived chances of admission, academic preferences, and financial considerations. Regular students often have to pay application fees, which can limit the number of schools they apply to and influence the types of schools they consider. For athletes, this is not an issue, as the recruitment process bypasses the traditional application stage; once they accept an offer, there is an application process there are different admission requirements for athletes and the likelihood of being rejected is extremely low. This streamlined process for athletes focuses more on the match between athletic talent and team needs rather than the broader selection of non-athletes.

2.6 Descriptive Statistics

Selection in Recruiting

Figure 2 shows the relationship between high school athletic ability, as measured by ESPN 300 rankings, and the quality of college football programs into which athletes are selected. The x-axis represents high school athletic ability, while the y-axis indicates the quality of the college program, with higher values representing more elite sports programs. The positive slope of the line suggests a strong correlation between an athlete's high school athletic ability and the quality of the college football program they attend. The linear fit line highlights that athletes with higher ESPN 300 rankings are more likely to be recruited into top-tier programs, indicating a clear selection mechanism based on athletic talent. This suggests that elite programs tend to recruit the highest-performing athletes from high school. It is important to understand this selection mechanism in order to address selection when estimating the causal effects of elite sports programs.

Figure 3 illustrates the distribution of scholarship offers received by high school athletes from the ESPN 300 rankings, plotted against their high school ability. The x-axis represents high school athletic ability scores, while the y-axis shows the total number of scholarship offers received by each athlete. The scatterplot reveals that athletes with higher ability scores tend to receive more scholarship offers, with the number of offers peaking around the middle of the distribution (ability scores around 80). Athletes at the extreme upper end of the ability scale (above 90) still receive a substantial number of offers, but the concentration of offers tends to taper off slightly. This figure demonstrates the high demand for top-performing athletes, where a significant number of offers are concentrated for those with above-average ability.

Interestingly, athletes at the very top of the ability distribution (ability scores above 90) do not receive the highest number of total offers. This is likely because top-tier schools focus their recruitment on these elite athletes, while mid-tier and lower-ranked schools avoid recruiting them, knowing they have little chance of securing their commitment. Instead, athletes in the mid-range of the ability spectrum, with scores around 80, tend to receive the most offers. This suggests that mid-tier programs are more actively competing for recruits in this range, as they are more likely to be within reach, while higher-ability athletes are targeted primarily by elite programs.

Elite Program Concentration

Over the past two decades, a striking pattern has emerged among NFL players and the college programs from which they are drafted. Nearly 80% of NFL athletes have come from just 20% of college football programs. This points to a fairly concentrated top-heavy distribution of talent and offers. There are currently over 900 colleges and universities with official football programs, but less than 30 schools—produce the majority of NFL players. This concentration illustrates the important influence of elite programs in the NFL.

Figure 4 visualizes this phenomenon, showing that a small subset of college programs dominates the NFL draft. These elite schools provide a disproportionate number of athletes who make it to the NFL, creating a clear hierarchy within college football. This suggests that athletes aiming for professional careers often cluster in programs with better resources, coaching, and visibility, further concentrating opportunities in the hands of a few institutions. This pattern is not unique to college football. Similar trends exist in other fields, such as academia. A study by Wapman et al. (2022) reveals that a small number of prestigious universities produce a significant share of tenure-track faculty in the U.S. With these descriptive statistics in mind, I highlight the importance of this research question addressing the role of elite sports programs on the career trajectory of student athletes. Keeping in mind a highly concentrated college sports industry and large selection in college sports recruiting I turn to section 3 discussing how to overcome these challenges and provide selection-corrected estimates of the private returns to elite sports programs.

3 Causal Effects of Elite Sports Programs

3.1 Statistical Model

My empirical analysis begins by adapting the matched-applicant model first developed in Dale and Krueger (2002) with similar frameworks used in Ge et al, (2023), Chetty (2022), Mountjoy (2021), and Chen et al, (2012). This model uses a selection-on-observables method to account for the nonrandom allocation of highly recruited student-athletes to college football programs.

The model linking student athlete characteristics to labor market outcomes such as job placement and performance I will assume takes the following form:

$$y_{ij} = \beta_0 + \beta_1 SRS_j + \beta_2' X_{1i} + \underbrace{\beta_3' X_{2i} + \epsilon_{ij}}_{u_{ij}} \tag{7}$$

y_{ij} represents outcomes for individual student athlete i, on team j. Team, j, has a sports program rating of SRS_j, measured by the Simple Rating System metric developed in Massey (1997). The term SRS_j is the key independent variable and is intended to measure the quality of the college sports program. Common in this literature is the use of average SAT score as a measure of selectivity with the assumption that selectivity is synonymous with school quality. In adapting these principles to my setting of collegiate sports and considering both the college enrollment process and requirements are vastly different related to traditional students, I treat the Simple Rating System program rating as interchangeable with college sports program quality and interpret β_1 the coefficient on the SRS_j variable as estimating the return to participating in a sports program of a given quality level. The vector contains X_1 are student athlete observable characteristics (height, weights, measures of athletic skill and ranking, position group and etc,. . .). I am, however, unable to observe all information relevant to outcomes and subsequently model the error term u_{ij} in equation 7 as the sum of two factors unobservable in the data. The first factor being X_{2ij}, this is information used in the recruiting process by college scouts, coaches, and recruiters during the recruiting season (usually football season of the athlete's junior and senior year of high school) and ϵ_{ij} the error term orthogonal to the other independent variables.

In the recruiting process for athletes, one of the challenges to estimating labor market returns is that student athlete characteristics causing different schools to extend a scholarship offer are not observed by the researcher. The recruiting process for student athletes is multifaceted and incorporates both the observed measures of athletic ability such as: points scored in a game, number of tackles recorded in a season, or strength and speed. Additionally, unobserved individual traits such as coach-ability, teamwork, performance under high stakes pressure are certainly important to college sports programs considering how to allocate scholarship offers. Furthermore, if any of these unobserved characteristics are correlated with the college program rating, then our estimate of the returns to participation would be biased. Specifically, if one believes the correlation to be positive, for example more talented or ambitious players are recruited by higher ranked college sports programs, then our estimate will be biased upwards. These unobservable characteristics are proxied for in the X_{2ij} vector under the assumption that the number of schools that extend scholarship offers as well the sports program quality of these scholarship offers reveals critical information used by in the college sports recruiting process.

3.2 Model Assumptions

To address concerns of selection on observation characteristics the common method is to introduce robust sets of controls variables that allow observations of similar or identical characteristics to be compared. Similarly, exploiting the information revealed in scholarship offersets the objective is to match student athletes together who were recruited by the same or similar sets of college sports programs. Thus, taking advantage of variation in college enrollment decisions while still comparing individuals with near identical observable and unobservable characteristics is the genius behind the matched applicant method developed in Dale and Krueger (2002) and subsequently furthered by Ge et al, (2023), Chetty (2022), Mountjoy (2021), and Chen et al, (2012) applied to differing populations of college students. My matching framework differs slightly from Dale and Krueger (2002) on two key dimensions: (1) application sets versus scholarship offer sets; (2) the exogenous nature of the application set. Considering the first difference, in Dale & Krueger (2002, 2014), Ge et al, (2023), Chetty (2022), and Mountjoy (2021) these studies have information on the set of colleges high school students apply to as well as subsequent acceptance and rejection information. Thus, three pieces of information are available for matching and to use to proxy for individual unobserved ability. As discussed in section 2.5 the recruiting process for high school student athletes is slightly different. Normally, it is college athletic programs that first reach out to students, establish contact, and offer an athletic scholarship; then, an athlete determines which college team to play for by accepting the scholarship offer and signing a Nation Letter of Intent (NLI) during an official signing period. Thus, while scholarship offer sets are different that application sets their purpose in the modeling framework is identical that of admission and rejection decisions. Chetty and Mountjoy show that having the admittance sets is similar to having the application, admittance, and rejection sets.

The aforementioned studies thus must take the application sets as exogenous and only model the college admissions process. In this research setting however, because athletes apply to college programs only after receiving an offer I do not need to rely on the this assumption. Critiques of this matching strategy argue that much of the selection between students is not in the colleges they are accepted to, but lies in the set of colleges students apply for admittance. For example, perhaps students from disadvantaged backgrounds might not even consider applying for some elite college programs even despite a high likelihood of acceptance because they have no information in their social network about what education is like at these types of institutions. Thus, the assumption to take the application sets as exogenously given is exceeding strong and information obtained from the application and rejections sets conditional on applications is not accounting for the individual selection by student in which schools to they submit an application. This critique is circumvented in this research setting because there no individual application process of which schools to seek make offers from. An athletic scholarship offer is a stronger, independent evaluation of ability and talent in the student athlete, signaled by a. college athletic department.

3.3 Matched Scholarship Model

Building on Dale and Krueger (2002) and adapting the matching framework to the college sports setting, I develop the matched scholarship model. This model accounts for the unique setting of athletic scholarship offers, which differ from traditional college admissions while still keeping true to the original innovation of the Dale and Krueger, Matched Applicant Model.

$$y_{ijg} = \beta_0 + \beta_1 SRS_j + \beta_2' X_{1i} + \sum_{1}^{m} \gamma_g Group_{ig} + \epsilon_{ijg} \tag{8}$$

where y_{ijg} represents labor market outcomes for individual i, associated with college team j and matching group g. The group indicator variables Group_{ig} capture the effect of belonging to a specific matching group, and ϵ_{ijg} represents the error term, accounting for unexplained variation in labor market outcomes.

Defining Matching Groups:

I begin by dividing college programs into bins according to their SRS score, with each bin corresponding to a different level of program quality. Let SRS_j represent the SRS score for college j, and let B be the number of bins (e.g., $B = 5$ for quintiles, $B = 3$ for terciles). Each college j is assigned to a bin based on its SRS score:

$$\text{Bin}_j = \begin{cases} 1 & \text{if } SRS_j \in \text{Quintile 1 (highest)}, \\ 2 & \text{if } SRS_j \in \text{Quintile 2}, \\ \vdots & \vdots \\ B & \text{if } SRS_j \in \text{Quintile } B \text{ (lowest)}. \end{cases}$$

For comparison, Dale & Krueger and Ge et al. (2023) bin sets of schools using a fixed 25-point interval on the average SAT score selectivity variable.

Generating the Matching Sets:

$$\theta_i = \left(\sum_{j \in \text{Bin } 1} O_{ij}, \sum_{j \in \text{Bin } 2} O_{ij}, \ldots, \sum_{j \in \text{Bin } B} O_{ij} \right)$$

where O_{ij} is an indicator variable equal to 1 if individual i received a scholarship offer from college j, and 0 otherwise.

Generating the Group ID: For each individual i, let O_{ij} be an indicator variable equal to 1 if individual i received a scholarship offer from college j, and 0 otherwise. The Group ID for individual i is constructed as a vector or sequence of digits, where each digit represents the count of offers received from colleges within each bin.

We redefine $Group_{ig}$ as conditional on the treatment variation condition, denoted by ν_g. Let ν_g be a binary indicator that reflects whether treatment variation exists within group g. Mathematically, this is defined as:

$$\nu_g = \begin{cases} 1 & \text{if treatment variation exists within group } g, \\ 0 & \text{if no treatment variation exists within group } g. \end{cases}$$

Condition of Treatment Variation: Treatment variation occurs when, within a group g, there exists at least one pair of individuals i and i' who received the same set of scholarship offers but chose to attend different colleges. Let j_i represent the college team j that individual i chooses to attend. Treatment variation exists in group g if:

$$\exists i, i' \in g \text{ such that } (\theta_i = \theta_{i'}) \text{ and } (j_i \neq j_{i'})$$

This condition ensures that the groups used in the regression analysis are those with meaningful comparisons, thereby allowing us to estimate the effect of different college choices on labor market outcomes. This approach refines the inclusion criteria for groups in our regression analysis, enhancing the validity of our estimates by focusing only on groups with substantive differences in college choices.

Defining $Group_{ig}$ Conditional on ν_g: Now, we define $Group_{ig}$, the group indicator variable, as conditional on ν_g. If $\nu_g = 0$, meaning no treatment variation exists within the group, that group is dropped from the analysis. The group indicator variable $Group_{ig}$ is defined as:

$$Group_{ig} \mid \nu_g = \begin{cases} 1 & \text{if } \theta_i = \theta_g \text{ and } \nu_g = 1, \\ 0 & \text{if } \theta_i = \theta_g \text{ and } \nu_g = 1, \\ \text{undefined} & \text{if } \nu_g = 0. \end{cases}$$

Thus, $Group_{ig}$ is only defined when treatment variation exists in group g ($\nu_g = 1$). If no treatment variation exists ($\nu_g = 0$), the group is dropped from the analysis and does not contribute to the regression model. This approach ensures that only groups with meaningful treatment variation are included in the analysis, enhancing the validity of the estimates.

Identification Assumption: The key identification assumption in this analysis is that, conditional on similar scholarship offer sets, the decision to accept a scholarship and join a particular team is uncorrelated with the error term ϵ_{ijg}. This assumption can be formalized as:

$$\mathbb{E}[\epsilon_{ijg} \mid Group_{ig} = 1, j_i] = \mathbb{E}[\epsilon_{ijg} \mid Group_{ig} = 1]$$

Where:

- ϵ_{ijg} is the error term capturing unobserved factors that affect the outcome y_{ijg}.
- $Group_{ig} = 1$ indicates that the individual i received a set of similar scholarship offers as other individuals in the group g.
- j_i represents the specific college team chosen by individual i.

Violation of the Assumption: The assumption is violated if:

$$\mathbb{E}[\epsilon_{ijg} \mid Group_{ig} = 1, j_i] \neq \mathbb{E}[\epsilon_{ijg} \mid Group_{ig} = 1]$$

This suggests that unmeasured characteristics may influence both the college choice (represented by j_i) and the outcome y_{ijg}, as discussed by Hoxby (2009). Such a violation implies that the selection of a college team may not be independent of unobserved factors that also affect the outcome, which could bias the results.

4 Empirical Results

4.1 The Effect of Elite Programs on Job Placement

Table 4 presents the main model specifications estimating the effect of college sports participation on job placement as a professional athlete. The outcome variable for this model is whether a student athlete was selected by a professional team in the NFL draft. I include four specifications for this model that highlight the progression of the empirical strategy. First, is baseline specification with minimal controls. I investigate the effect of college team quality with minimal controls for student athlete height and weight. Specification (2) adds measures of student athlete athletic skill pre-college as measured by the ESPN 300 analysts; these measures are the equivalent of student's own standardized test score but for athletic ability. With specification (3), I address the impact of peer quality. There are two potential sources of peer effects, one is the quality of teammates on a college football team before the incoming college freshman join the team. Second, is the quality of teammates who were recruited together as high school students, and all will be joining a particular college team at the start of a new season. Specification (3) seeks to capture the later source of peer effects by including the number of top high school athletes recruited to the same college team for each individual student athlete observation. There is sustainable variation in number of top ESPN high school athletes recruited to college team rosters. The quality of peers that enter the college program with a student athlete could affect athletic development and the outcome of being selected in the NFL draft.

Finally, specification (4) incorporates all previous control variables as well as two control variables that seek to mitigate concerns of unobserved factors that influence recruiting and bias college team quality. These variables are the total number of scholarships offered to the high school athlete, and the average college team quality of all the teams in the scholarship offer set for each high school player. Average team quality of the scholarship offer set is computed by first matching each school in the offer set to its related quality measure of winning percentage, then compute the average winning percentage from each high school athlete's scholarship offer set. Finally, I create quartile bins on the continuous average scholarship offer set winning percentage variable and include dummy variables for each quartile bin.

Matching Dale and Krueger 2002, 2011, all explanatory variables are determined prior to when the student athlete begins college. Looking at model (1) we see that for a 1 standard deviation increase in college program quality as measured by the Simple Rating System metric (SRS) increases the likelihood of being drafted by a professional team 0.043 percentage points. When we add in measures of athletic skill, measured in high school, this job placement premium on college team quality shrinks substantially to 0.027 percentage points. Additionally, accounting for incoming peer compositions further decreases the coefficient of interest to 0.024 percentage points. Finally, I add in the variables from the scholarship offer set model (4) to similarly replicate the self-revelation model of Dale and Krueger 2002. The coefficient of interest, the effect of participation in a college sports program, again diminishes when the additional controls are added into model three but still captures a large and significant effect. For a 1 standard deviation increase in college team quality (SRS), the likelihood of being drafted into the NFL increases by 0.018 percentage points.

The average likelihood of being drafted is reflected in the mean drafted term with a value of 0.056 percent in specification (4). Thus, for high school athlete participating in a college sports program one standard deviation higher in college team quality (SRS), increases the likelihood of being drafted by 32% of the mean. Moving from the lowest ranked school to the highest ranked school would 192% change in the likelihood of being drafted, or changing the likelihood of being drafted from 0.056 to 0.164. Moving from a median ranked school to a top ranked school results in a 96% percent increase in the likelihood of being drafted with the likelihood change from 0.056 to 0.11. This effect is larger than the impact of an individual student athlete's incoming peer group, and smaller but of similar magnitude as the impact of the athletes own athletic skill as measured by the ESPN 300 analyst grade and rank. Athletic skill is intuitively the largest determinant of a professional athletic career, yet the impact participation in a more elite college football program has a significant return in terms of job placement as a professional athlete.

Figure 6 illustrates impact of participating in an elite college football program. Taking the predicted probability of being drafted from model (4) of Table 4 and graphing if along with the measure of college team quality (SRS), going from the bottom quartile to the top quartile is associated with a three to five time increase in the predicted probability of being selected in the NFL draft. There is extreme variation in salaries between college athletes selected first in the NFL draft versus being selected in the later rounds, however, at the time of writing "Mr. Irrelevant", the affectionate title for the college athlete selected last in the NFL draft each year had a salary of over $700,000 for each year of the four-year rookie contract. Thus, just being selected by an NFL team to play football professionally increases earnings dramatically.

4.2 Robustness Check – Sensitivity of Returns to College Quality Measure

As stated previously there are many ways to measure college program quality and I discuss the sensitivity of my preferred specification results (Table 4, col 4) to alternative measures of college program quality. I re-estimate equation (2) six times, only varying the college program quality measure. I choose six other program quality measures with similar attributes as those of my preferred quality measure, Simple Rating System (SRS), including: total program winning percentage (total wins / total games), Strength of Schedule (SOS) (see equation 5), number of professional players from a particular college program, number of years a program was ranked in the top 25 in the nation, number of conference championships, and winning percentage of post-season or national competitive tournament games. Each of these alternative quality measures captures some dimension of what it means to be a national competitive or elite college football program.

As demonstrated in Table 5, each of the alternative measures is associated with a positive and significant impact on the outcome of being selected in the NFL draft. Estimates of the impact of participation in more elite college football program range from 0.003-0.044 percent. Thus, my preferred quality measure is a towards the median of all the quality measure estimates, what I consider a conservative estimate of the return to an elite sports program regarding job placement.

4.3 Matching Group Specification Sensitivity

Table 6 show five different matching model specifications compared to the baseline results. There are multiple methods for generating matching groups when evaluating the impact of scholarship offers on athlete outcomes, each involving a trade-off between group homogeneity and sample size. More homogeneous groups, such as those created by exact matching, result in smaller sample sizes because of the treatment variation condition, where only groups with variation in the treatment are kept. On the other hand, less restrictive matching models retain more of the sample but create less homogeneous groups. For example, in the *Matching Model 1*, athletes are grouped based on a 5-digit binary ID indicating the presence or absence of scholarship offers from schools ranked in different quintiles. *Matching Model 2* and *Matching Model 3* extend this by counting offers within each quintile, while *Matching Model 4* and *Matching Model 5* use deciles and terciles, respectively, with the number of offers per group capped at different thresholds.

The trade-off between within-group homogeneity and sample size is documented in the summary table. For instance, *Exact Matching* creates 20,500 groups with only 177 having treatment variation, yielding a significantly smaller sample of 422 observations. In contrast, *Matching Model 2* groups the athletes into 4,127 groups, retaining 2,326 groups with treatment variation and a sample size of 21,109. Despite the differences in matching specifications, the main effect of college quality (measured by SRS) remains consistent across models, showing a positive and statistically significant relationship with athlete outcomes. The effect size is stable, with coefficients ranging between 0.014 and 0.018 across the models, reinforcing the robustness of the findings regardless of the matching method used.

Exact matching results are not reported because 98% of the sample is lost under this method, rendering the remaining matches insufficient for meaningful analysis. The exact matches tend to lack the necessary variation in treatment between elite and non-elite programs, which is crucial for addressing the research question. Since exact matching groups do not capture the diversity in college quality offers that is central to the study, they are not representative of the broader athlete population and provide limited insight into the effects of attending elite programs. As such, less restrictive matching models are more appropriate for this analysis.

4.4 Accounting for Athletic Performance

In this section, I address the possibility that the large premium associated with attending an elite college football program may be due to differences in individual athletic performance rather than the inherent value of the program itself. Up to this point, I have accounted for selection on observed characteristics, controlled for peer effects, and accounted for high school ability. Furthermore, the use of scholarship offer sets has helped mitigate potential bias from unobserved characteristics. However, a key question remains: Are elite programs merely proxies for superior athletic performance, with their apparent impact on outcomes disappearing once I account for actual individual performance?

To explore this, I introduce a potentially endogenous variable—individual athletic performance at the college level—as a robustness check. While including this variable might complicate the causal interpretation of our analysis, the primary purpose here is to examine whether differences in athletic performance explain the observed effects. If athletic performance accounts for a significant portion of the results, it would suggest that elite programs do not directly influence outcomes but rather recruit players who are already better performers. This check will help identify any omitted variable bias that may be influencing or skewing the results.

To measure athletic performance, I draw from the methodology outlined in section 2, where player performance is standardized and aggregated into composite scores across various position groups. College football teams are divided into 33 unique positions across 3 units, further grouped into 10 position categories. These categories include both offensive and defensive positions, each of which has 3 to 6 distinct performance measures (with exceptions for offensive linemen and certain defensive positions). Performance metrics, such as sacks, tackles for loss, passing yards, touchdowns, and rushing averages, are critical indicators of an athlete's contribution to the game.

To ensure comparability across different positions and categories, each raw performance measure is standardized using a z-score formula, where the player's performance is adjusted relative to the mean and standard deviation of that measure within the player's category. These standardized scores are then aggregated into composite performance scores for each player, capturing their overall athletic contribution. By incorporating these composite performance measures into our baseline model, we aim to assess whether the premium for elite college programs persists once individual athletic performance is accounted for. This additional robustness check will help determine whether the observed premium is driven by the program itself or simply reflects the superior athletic performance of its recruits.

Table 7 presents results from several regression models examining the effect of attending an elite college football program on the likelihood of being selected in the NFL Draft. The key result is that the return to college program quality (as measured by the SRS) remains significant and consistent across all models, even after accounting for individual athletic performance. In model (1), where only college quality is included, the coefficient for program quality is 0.058, and this remains significant at the 1% level as additional controls are added. In the final model (5), which accounts for both high school ability, peer effects, and individual college performance, the coefficient for college quality is still positive and significant (0.027), indicating that elite programs confer a premium even when considering athletic performance.

Additionally, when college athletic performance is introduced in model (5), the results show a positive and highly significant coefficient (0.055), suggesting that athletic performance plays an important role in draft selection. However, the fact that the college program quality variable remains significant implies that attending an elite program offers advantages beyond just individual performance on the field.

It is also worth noting that the sample changes slightly when merging data from the ESPN 300 high school athletes with performance measures from the CollegefootballData API. In particular, all defensive athletes before 2016 are dropped from the analysis due to the lack of recorded defensive performance measures before that year. This change primarily affects defensive positions, reducing the overall sample size but not altering the core results regarding the impact of elite programs on draft selection.

4.5 Heterogeneous Effects by Position Group

Similar to many other sports, athletic performance in American Football is measured differently for different position groups. There is substantial variation in how performance is measured as well as the types of measures available for different position groups. There are three positions groups where performance is easiest and most transparent to measure. These positions are quarterback, running backs, and wide receivers, collectively known as "offensive skill positions." As these position groups are those most likely to score offensive points during a football game. Generally, speaking points are scored by advancing the ball forward as measured by positive yards gained. I investigate whether the returns to elite sports programs are homogeneous across position groups or heterogeneous by position type.

Figure 7 illustrates the heterogeneous effects of college program rank on the probability of being drafted into professional football across different position groups. Each blue point represents the coefficient for a specific position group, showing how the rank of the college program impacts draft prospects for players in that position. The error bars reflect the standard errors, giving a sense of the uncertainty around these estimates. The red dashed line represents the average effect (coefficient from Table 4), providing a benchmark for comparison across position groups.

This analysis of heterogeneous effects is useful in the setting of football due to American football's highly specialize nature. Athletes train for years in specific positions, such as quarterback (QB), offensive tackle (OT), or cornerback (CB), and rarely switch positions. As a result, the skills and performance expectations are tailored to the demands of each position group, making it crucial to understand how factors like college program rank affect different positions uniquely. For example, the figure shows that offensive tackles (OT) and offensive guards (OG) benefit more from attending higher-ranked programs, whereas quarterbacks (QB-DT and QB-PP) show a negative or neutral relationship with college program rank, suggesting that individual performance measures may matter more for them than the prestige of their college program.

Heterogeneous effects are also a key consideration in the broader literature on economic returns to elite educational programs. For instance, Brewer, Edie, and Ehrenberg (1999) found significant returns to attending elite private institutions, while Dale and Kruger (2002, 2011) reported that returns to attending elite colleges were indistinguishable from zero when measuring long-term earnings. These mixed findings highlight that the effects of elite education may vary significantly across different student groups, just as the effects of elite athletic programs vary across position groups in football. Similarly, Chetty, Demming, and Friedman (2023) found that attending an Ivy-Plus college significantly increased the chances of reaching the top 1% of earners, which emphasizes the potential for substantial variation in outcomes based on student background and program type. As noted by these studies this is a literature where it is important to be mindful of heterogeneous effects.

5 Theoretical Explanations of Returns

Understanding the returns to elite sports programs can be approached through different theoretical frameworks, such as human capital and signaling models. Each framework offers distinct implications for how returns are generated and what they mean for student-athletes and employers (professional teams). A simple model allows us to compare these theoretical perspectives, providing a structured way to interpret empirical findings.

For example, returns consistent with a human capital model suggest that the value added by the college program enhances the athlete's inherent abilities, which leads to better professional outcomes. On the other hand, a signaling model implies that the returns are driven by the information conveyed about the athlete's pre-existing abilities, rather than any value added by the program itself. Using this simple model, we can guide ex-ante predictions of heterogeneous effects in baseline regressions. It also serves as a foundation for counterfactual analysis, particularly in exploring the impact of recent policy changes in college sports, such as Name, Image, Likeness (NIL) deals and the open transfer portal. These analyses help us understand how the returns to elite sports programs might shift under different scenarios.

5.1 Model Setup

Simple Search Model We begin with a basic search model where employers (professional teams) assess the productivity of athletes based on two pieces of information: a private signal reflecting individual performance and group affiliation, which is tied to the athlete's college program.

Employers observe a set of applicants with two pieces of information:

- A private signal S_i
- Group affiliation V_G

Thus, the productivity of applicant i from school G is defined as:

$$V_{G,i} = V_G + \epsilon_{G,i}$$

where $V_{G,i}$ is the productivity of applicant i, and $\epsilon_{G,i}$ represents the idiosyncratic productivity above the group effect. The productivity follows a normal distribution:

$$V_{G,i} \sim \mathcal{N}(V_G, \sigma_v^2)$$

Employers wish to access the true productivity of applicants, $V_{G,i}$, but are unable to do so and must instead rely on a private signal, $S_{G,i}$. The private signal is a function of the true productivity of the athlete and an idiosyncratic error term $\eta_{G,i}$. The private signal can be thought of as a "score" from a type of examination or performance measure during an athletic competition. The private signal is driven by the underlying productivity of the individual but is measured imperfectly with some error.

The private signal is modeled as:

$$S_{G,i} = V_{G,i} + \eta_{G,i}$$

where $\eta_{G,i}$ is distributed normally with mean 0 and variance σ_η^2:

$$\eta_{G,i} \sim \mathcal{N}(0, \sigma_\eta^2)$$

Thus, we can rewrite the private signal as a function of three elements: group effect, measurement error (luck) of the signal, and $\epsilon_{G,i}$, which represents individual effort above and beyond the group effect.

Private Signal and Employer's Dilemma

The "principal," or in this setting professional football teams hiring college athletes, observe three pieces of information: (1) The private signal $S_{G,i}$; (2) The distribution of true productivity $V_{G,i}$ and the distribution of the private signal $S_{G,i}$; (3) Each individual athlete's group affiliation G_i.

Professional teams aim to infer the expected future productivity of college athletes based on the private signal and group affiliation information in order to make hiring decisions. The expected productivity $\mathbb{E}[V_{G,i} \mid S_{G,i}]$ is modeled as a weighted combination of the group average productivity V_G and the private signal $S_{G,i}$:

$$\mathbb{E}[V_{G,i} \mid S_{G,i}] = (1 - \gamma_G)V_G + \gamma_G S_{G,i}$$

Using properties of statistics, one can solve for the optimal weights in the principal's or firm's hiring problem (see Appendix C.4 for proof). The optimal weights are defined as follows:

- The optimal weight γ_G ranges between 0 and 1:

$$\gamma_G = \frac{\sigma^2_{G\epsilon}}{\sigma^2_{G\epsilon} + \sigma^2_{G\eta}}$$

$$1 - \gamma_G = \frac{\sigma^2_{G\eta}}{\sigma^2_{G\epsilon} + \sigma^2_{G\eta}}$$

This model has several important implications. The weight γ_G lies between 0 and 1, reflecting how much the principal trusts the private signal relative to the group average. As the accuracy of the private signal increases (i.e., $\sigma^2_{G\eta}$ decreases), γ_G approaches 1, leading the principal to rely more heavily on the private signal $S_{G,i}$ when predicting productivity. Conversely, if the signal is less accurate, the principal will place more weight on the group affiliation V_G.

This model presents a challenge for employers: how much weight should be placed on the private signal $S_{G,i}$ versus the group affiliation V_G when predicting the productivity of the athlete? The answer depends on the accuracy of the private signal. As the signal becomes more accurate (i.e., $\sigma^2_{G\eta}$ decreases), employers are expected to place more weight on $S_{G,i}$ and less on V_G, effectively down-weighting the importance of the athlete's college affiliation.

Figure 8 demonstrates that as signal accuracy improves (i.e., $\sigma^2_{G\eta}$ decreases), the weight placed on group affiliation $(1-\gamma_G)$ decreases. Group affiliation has a diminishing role as individual performance measures become more accurate.

5.2 Extension to Include Human Capital

I extend the model to incorporate human capital accumulation, where an athlete's productivity evolves over time based on their investment in human capital. The extended model allows for dynamic analysis, showing how changes in human capital affect the weight employers place on individual performance signals versus group affiliation.

5.2.1 Human Capital Accumulation Process

Human capital $H_{G,i,t}$ accumulates over time, and its variance $\sigma^2_{H,t}$ can influence the overall weight placed on the private signal. The private signal is modified to account for this human capital accumulation, and the principal's problem is revisited in this extended context.

Let's denote the human capital of an individual i from group G at time t as $H_{G,i,t}$. The productivity of the individual $V_{G,i,t}$ at time t would then depend on both their initial group affiliation and their accumulated human capital:

$$V_{G,i,t} = V_G + \epsilon_{G,i} + \beta H_{G,i,t}$$

Here:

- V_G is the average productivity of the group (college program).
- $\epsilon_{G,i}$ is the initial individual deviation from the group average productivity.
- β is a parameter that measures the return on human capital investment.

- $H_{G,i,t}$ is the accumulated human capital of individual i at time t.

Human capital $H_{G,i,t}$ can be modeled as a function of time, investment in education or training, and other factors. A simple linear form might be:

$$H_{G,i,t} = H_{G,i,0} + \sum_{s=1}^{t} \alpha I_{G,i,s} + \eta_{G,i,t}$$

Where:

- $H_{G,i,0}$ is the initial human capital of individual i.
- $I_{G,i,s}$ is the investment in human capital (e.g., training, education) at time s.
- α is the rate at which investment translates into human capital.
- $\eta_{G,i,t}$ is the random shock to human capital accumulation at time t.

5.2.2 Adjusting the Principal's Problem

Given that productivity now depends on human capital, the private signal $S_{G,i,t}$ observed by employers at time t should reflect this:

$$S_{G,i,t} = V_{G,i,t} + \eta_{G,i,t} = V_G + \epsilon_{G,i} + \beta H_{G,i,t} + \eta_{G,i,t}$$

The principal (employer) must now predict the expected productivity $\mathbb{E}[V_{G,i,t} \mid S_{G,i,t}]$ based on both the initial group affiliation and the accumulated human capital. The expected productivity at time t becomes:

$$\mathbb{E}[V_{G,i,t} \mid S_{G,i,t}] = (1 - \gamma_{G,t})V_G + \gamma_{G,t} S_{G,i,t}$$

Where the weight $\gamma_{G,t}$ on the private signal now depends on the variance of the human capital accumulation process:

$$\gamma_{G,t} = \frac{\sigma^2_{G\epsilon} + \beta^2\sigma^2_{H,t}}{\sigma^2_{G\epsilon} + \beta^2\sigma^2_{H,t} + \sigma^2_{G\eta,t}}$$

Here:

- $\sigma^2_{H,t}$ is the variance in human capital accumulation at time t.
- $\sigma^2_{G\eta,t}$ is the variance of the noise or "luck" component at time t.

This extension introduces a dynamic component where the weight on the private signal $S_{G,i,t}$ may change over time as human capital accumulates. Early in the career, group affiliation V_G may play a larger role in predicting productivity, but as human capital $H_{G,i,t}$ accumulates, the private signal $S_{G,i,t}$ (which now includes the effect of human capital) becomes more informative.

As a result, the model can capture how the importance of college affiliation decreases over time as the athlete's individual performance, driven by accumulated human capital, becomes the dominant factor in predicting future success. This dynamic framework allows for examining the long-term returns to college programs and the role of human capital in shaping career trajectories.

The model compares two scenarios related to the impact of human capital variance $\sigma^2_{H,t}$ on the weight placed on group affiliation. Figure 9 illustrates these two potential scenarios.

In the flat line scenario (green), shocks to human capital $\sigma_{H,t}$ do not affect the weight placed on group affiliation. The weight remains constant, implying that regardless of variations in human capital, employers' reliance on group affiliation for evaluating an athlete's potential remains unchanged. This suggests that group affiliation continues to play a consistent role in decision-making.

In contrast, the increasing line scenario (red) represents a case where the weight placed on group affiliation rises as the variance in human capital $\sigma_{H,t}$ increases. As the variability in human capital grows, employers may place greater importance on group affiliation when assessing an athlete's potential, likely due to the increased uncertainty in individual performance signals.

Figures 8 and 9 show that this simple model has very different predictions for how employers should weight group affiliation depending on changes to the parameters of the weighting function. Changes in information reflect a signaling mechanism, while changes in "job training" are more consistent with human capital accumulation.

5.3 Measuring Private Signal Accuracy

In the context of evaluating athletes in American football, the accuracy of the private signal—an essential component in assessing an individual's potential—is influenced by the number and quality of performance measures available for each player. American football, with its highly specialized roles, provides a rich production function that results in significant variation in the types and quantities of performance data recorded for different position groups. For example, quarterbacks (QBs) might have 10 to 15 distinct performance metrics recorded in a single game, capturing various aspects of their play, such as completions, passing yards, and touchdowns. In contrast, offensive lineman, whose roles are more limited, might only have 2 to 4 performance metrics available, reflecting a much narrower set of activities.

This variation extends beyond the type of performance measures to the number of observed plays per game and per season, which further affects the accuracy of the private signal. Quarterbacks, who are central to most offensive plays, might be observed in 50 to 75 plays in a typical game, providing a wealth of data points that enhance the accuracy of their performance signal. On the other hand, kickers may only participate in 5 plays per game, leading to a more limited and potentially less accurate signal.

The underlying assumption in measuring private signal accuracy is straightforward: as the number of observed performance measures increases, the private signal becomes more accurate. Similarly, increased playing time, resulting in more observed plays, also contributes to a more precise measurement of an athlete's performance. These differences in the availability and quantity of data across positions suggest that the accuracy of private signals can vary significantly depending on the role a player occupies on the field.

Table 1 provides an overview of the unique statistical performance measures available for various categories in football, ranging from defensive statistics like sacks and solo tackles to offensive metrics such as passing yards and completions. Each position group has a distinct set of metrics, reflecting the specialized nature of their roles in the game. This diversity in performance measures underscores the importance of accounting for position-specific data when evaluating the accuracy of private signals in the context of professional sports.

Figure 10 presents the average number of performance measures recorded for various football position groups, with error bars representing the standard error of the mean (SEM). The data includes positions such as quarterback (QB), running back (RB), and wide receiver (WR), among others, and highlights the variation in the number of performance metrics available for different position groups.

A few key observations from the figure include the quarterback positions (both QB-DT and QB-PP) having the highest average performance measure counts, with 11.34 and 10.96 measures, respectively. This shows the high level of scrutiny placed on quarterback play. In contrast, positions like offensive tackle (OT) and center (OC) show significantly fewer recorded metrics, with averages of 2.78 and 2.48, respectively, suggesting fewer specialized performance metrics are tracked for these positions. These differences highlight the role-specific demands and the variability in available data for each position group.

5.4 Theoretical Predictions & Heterogeneous Effects

Figure 11 presents the relationship between two key variables: (1) the coefficient representing the effect of college program rank on player outcomes (plotted as blue points with error bars), and (2) the average number of performance measures recorded for each high school position group (represented by red bars with error bars indicating the standard error of the mean). The x-axis lists the position groups, such as quarterbacks (QB), wide receivers (WR), and offensive linemen (OL), with both variables plotted to highlight the effect of college program rank alongside the average number of performance measures for each position.

The blue points (with error bars) show how college program rank influences player outcomes across various positions. For instance, offensive tackles (OT) and offensive guards (OG) exhibit the highest positive coefficients, indicating that college program rank has a more substantial impact on outcomes for these positions. In contrast, positions like fullbacks (FB) and defensive tackles (DT) have negative coefficients, suggesting that college program rank plays a lesser or even negative role in predicting outcomes for these groups.

The red bars illustrate the average number of performance measures recorded for each position. Quarterbacks (both QB-DT and QB-PP) show the highest average performance measure counts, with over 10 measures, reflecting the greater complexity and scrutiny applied to these positions. On the other hand, positions like offensive tackle (OT) and center (OC) have significantly fewer performance measures, with averages below 3 measures.

Offensive tackles (OT) exhibit a high positive coefficient of 0.053, paired with a relatively low average performance count of 2.78, suggesting that the college program rank significantly impacts outcomes for this position, likely due to the limited availability of individual performance measures. In contrast, quarterbacks (QB-DT and QB-PP), despite having the highest average performance measure counts—around 11 and 10, respectively—show negative coefficients for college program rank, indicating that the prestige of the college program is less influential for these positions, where more performance data is available. Fullbacks (FB), with a notably negative coefficient of -0.054, show that college program rank may negatively influence outcomes, most likely reflecting changes to NFL offensive personnel. The fullback position group has become less important for many teams due to the increased emphasis on passing among many NFL teams in recent years, see Charlton (2024)for a full discussion.

Evidence Supporting the Signaling Framework

This figure provides strong evidence that the returns to attending elite sports programs are consistent with a signaling framework. The signaling model offers ex ante predictions of heterogeneous effects of college program rank by position group. Specifically, for position groups with more available performance information—such as quarterbacks (QB)—the effect of college program rank is less meaningful, as the large amount of measured data allows employers to make more accurate assessments based on individual performance. In contrast, for position groups with limited measured information—such as offensive linemen (OL) or fullbacks (FB)—attending an elite program plays a much more significant role. The lack of extensive individual performance data means employers may rely more heavily on the prestige of the athlete's college program as a proxy for ability, consistent with the signaling model.

References

Anderson, Michael L. 2012. "The Benefits of College Athletic Success: An Application of the Propensity Score Design with Instrumental Variables. NBER Working Paper No. 18196." National Bureau of Economic Research. Publication Title: National Bureau of Economic Research ERIC Number: ED533309.

Arteaga, Carolina. 2018. "The effect of human capital on earnings: Evidence from a reform at Colombia's top university." *Journal of Public Economics*, 157: 212–225.

Barrera-Osorio, Felipe, and Hernando Bayona-Rodríguez. 2019. "Signaling or better human capital: Evidence from Colombia." *Economics of Education Review*, 70: 20–34.

bcsguestwriter. 2012. "Top-50 Largest Athletics Department Spenders."

Berg Dale, Stacy, and Alan B. Krueger. 1999. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables."

Bordón, Paola, and Breno Braga. 2020. "Employer learning, statistical discrimination and university prestige." *Economics of Education Review*, 77: 101995.

Brewer, Dominic J., Eric R. Eide, and Ronald G. Ehrenberg. 1999. "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *The Journal of Human Resources*, 34(1): 104–123. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System].

Case, Anne, Angela Fertig, and Christina Paxson. n.d.. "The lasting impact of childhood health and circumstance." 24(2): 365–389.

Chen, Weiwei, Wayne A. Grove, and Andrew Hussey. 2013. "The Payoff to School Selectivity: An Application of Dale and Krueger's Method to MBA Programs."

College Athletics Spending And The Movement Towards Revenue Sharing. 2021. Archive Location: World Last Modified: 2022-02-10T01:56:22-07:00 Publisher: https://athleticdirectoru.com/ Section: Articles.

Conti, Gabriella, and James J Heckman. n.d.. "Understanding the Early Origins of the Education–Health Gradient." 5(5): 585–605.

Dale, Stacy, and Alan B. Krueger. 2011*a*. "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data."

Dale, Stacy, and Alan Krueger. 2011*b*. "Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data." National Bureau of Economic Research w17159, Cambridge, MA.

Dale, Stacy B., and Alan B. Krueger. 2014. "Estimating the Effects of College Characteristics over the Career Using Administrative Earnings Data." *The Journal of Human Resources*, 49(2): 323–358. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System].

Dale, Stacy Berg, and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *The Quarterly Journal of Economics*, 117(4): 1491–1527. Publisher: Oxford University Press.

Desrochers, Donna M. n.d.. "Academic Spending Versus Athletic Spending: Who Wins?" 16.

Dickson, Matt, and Colm Harmon. 2011. "Economic returns to education: What We Know, What We Don't Know, and Where We Are Going—Some brief pointers." *Economics of Education Review*, 30(6): 1118–1122.

Huntington-Klein, Nick. 2021. "Human capital versus signaling is empirically unresolvable." *Empirical Economics*, 60(5): 2499–2531.

Mabel, Zachary, C. J. Libassi, and Michael Hurwitz. 2020. "The value of using early-career earnings data in the College Scorecard to guide college choices." *Economics of Education Review*, 75: 101958.

Massey, Cade, and Richard H. Thaler. 2013. "The Loser's Curse: Decision Making and Market Efficiency in the National Football League Draft." *Management Science*, 59(7): 1479–1495. Publisher: INFORMS.

Staff, Sportico. 2021. "Sportico's Intercollegiate Finance Database."

Weiss, Andrew. 1995. "Human Capital vs. Signalling Explanations of Wages." *Journal of Economic Perspectives*, 9(4): 133–154.

Zimmerman, Seth D. 2019. "Elite Colleges and Upward Mobility to Top Jobs and Top Incomes." *American Economic Review*, 109(1): 1–47.

Figures and Tables

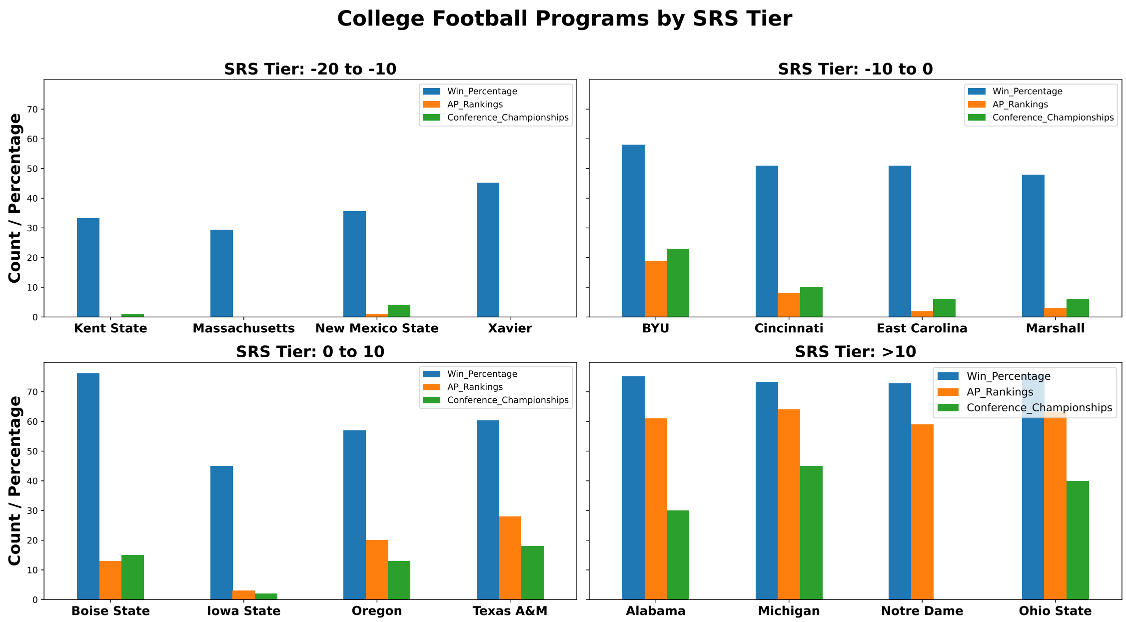

Figure 1: College Football Programs by Simple Rating System Tier

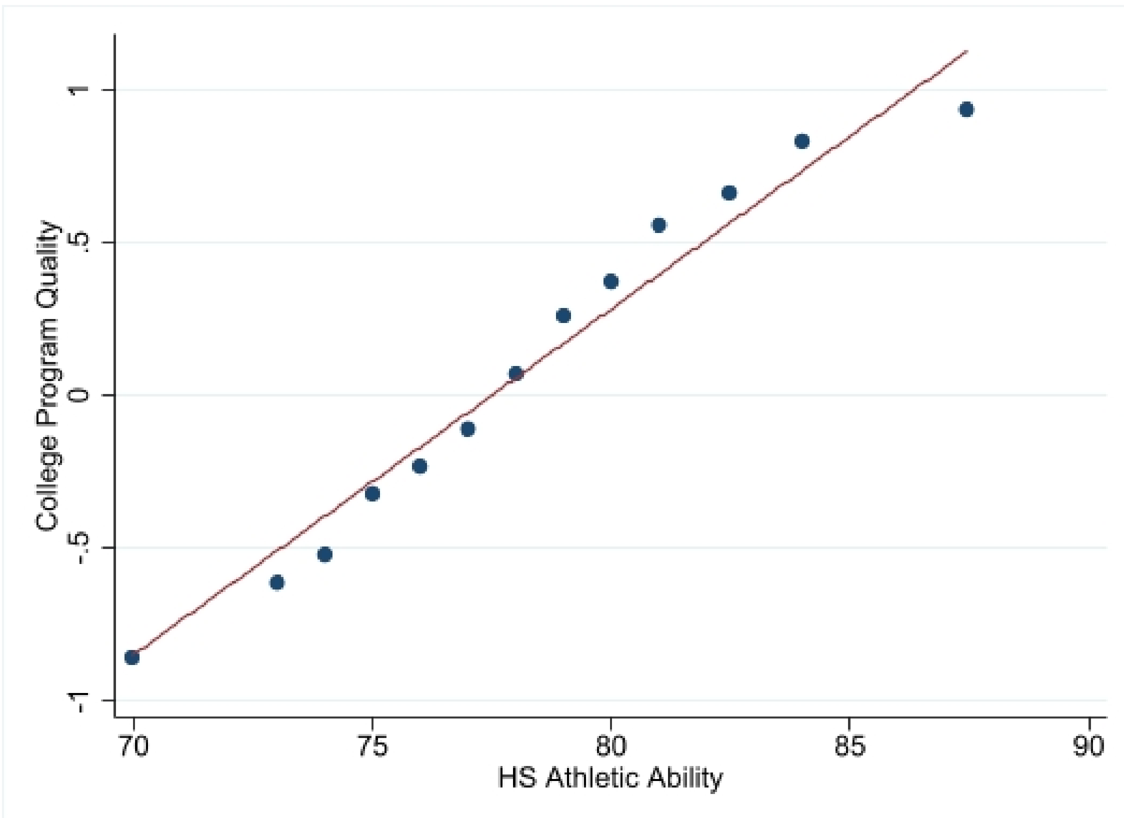

Figure 2: Selection into College Football Programs by ESPN 300 HS Athletes

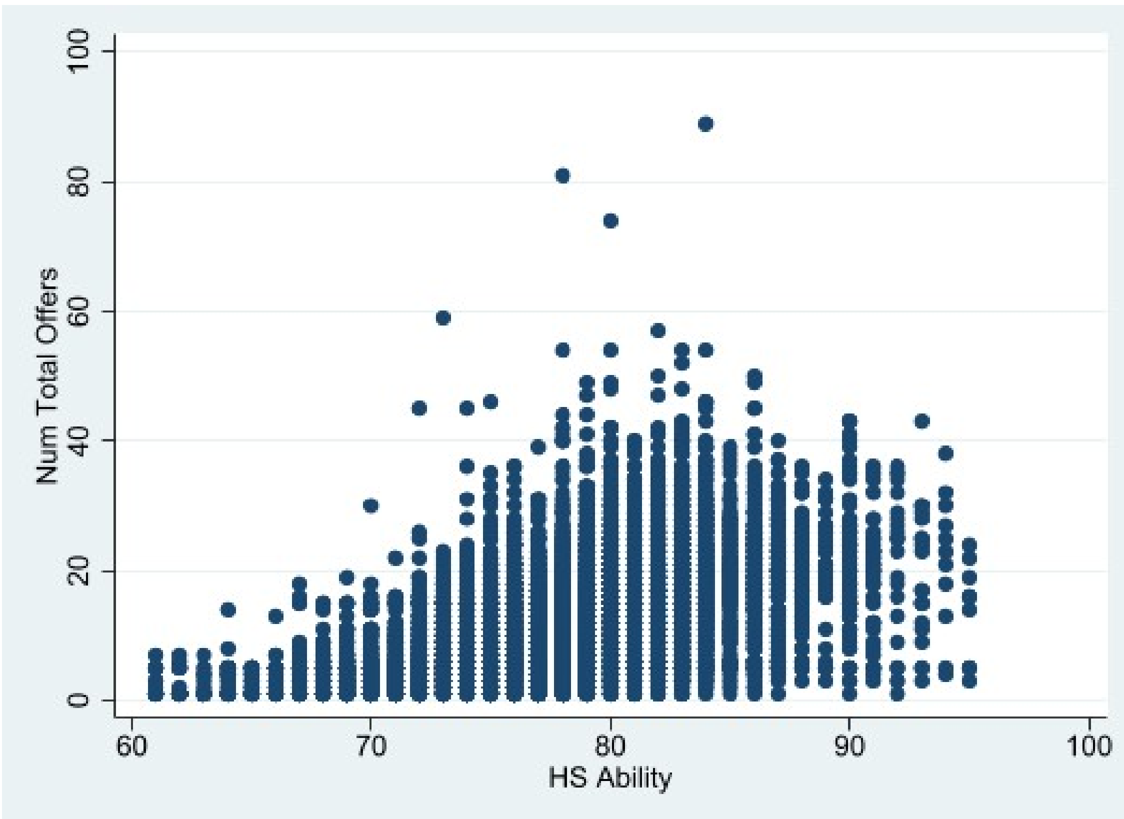

Figure 3: Distribution of Scholarship Offers to ESPN 300 HS Athletes

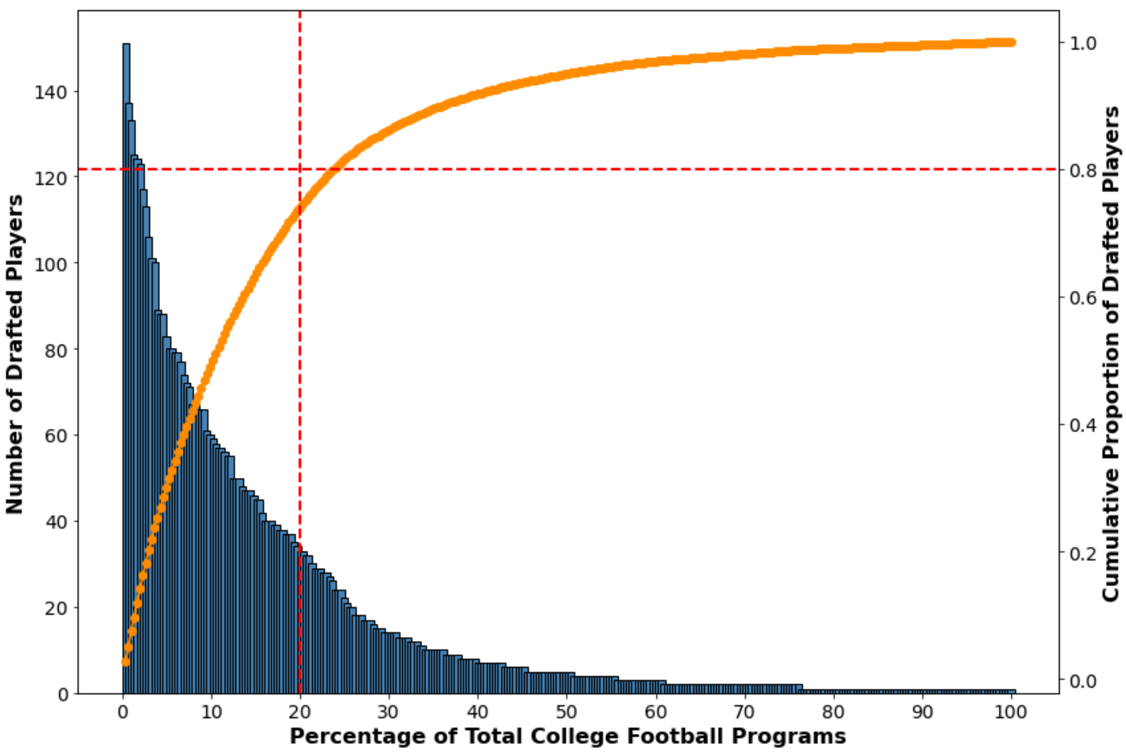

Figure 4: Concentration of NFL Talent by College Football Program

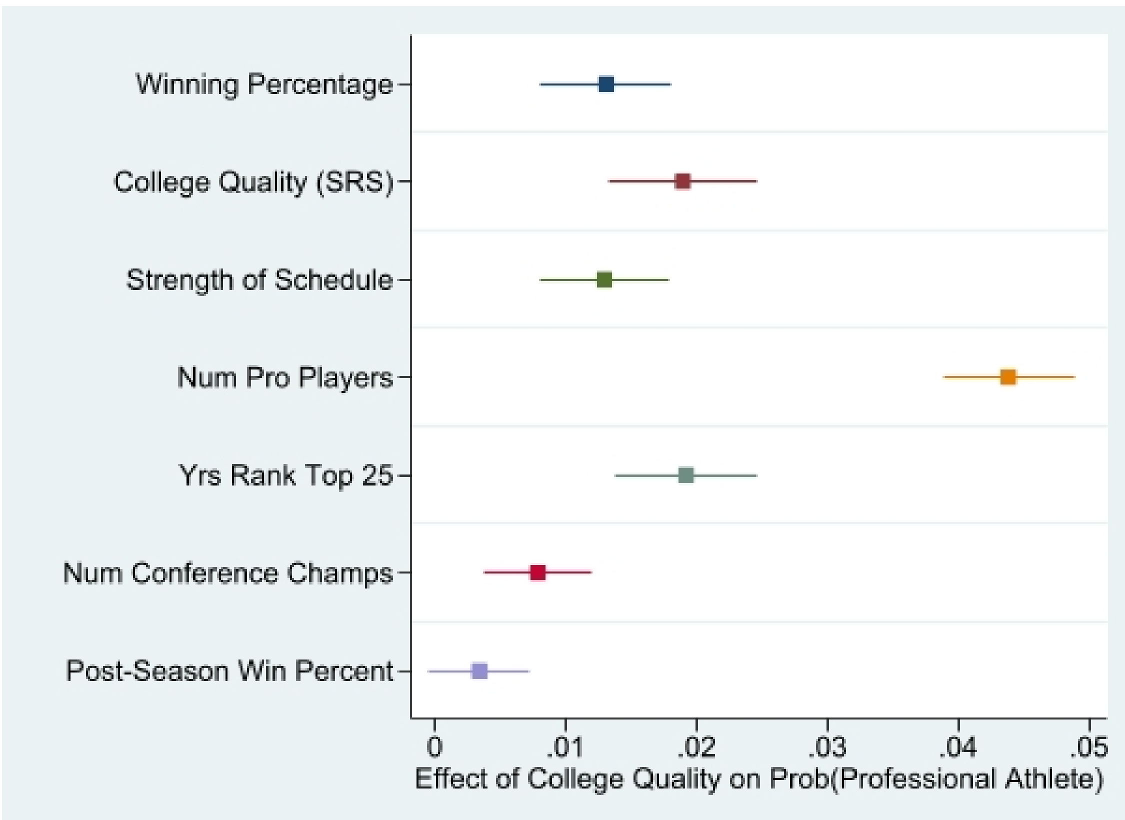

Figure 5: Sensitivity Analysis - College Program Quality Measures

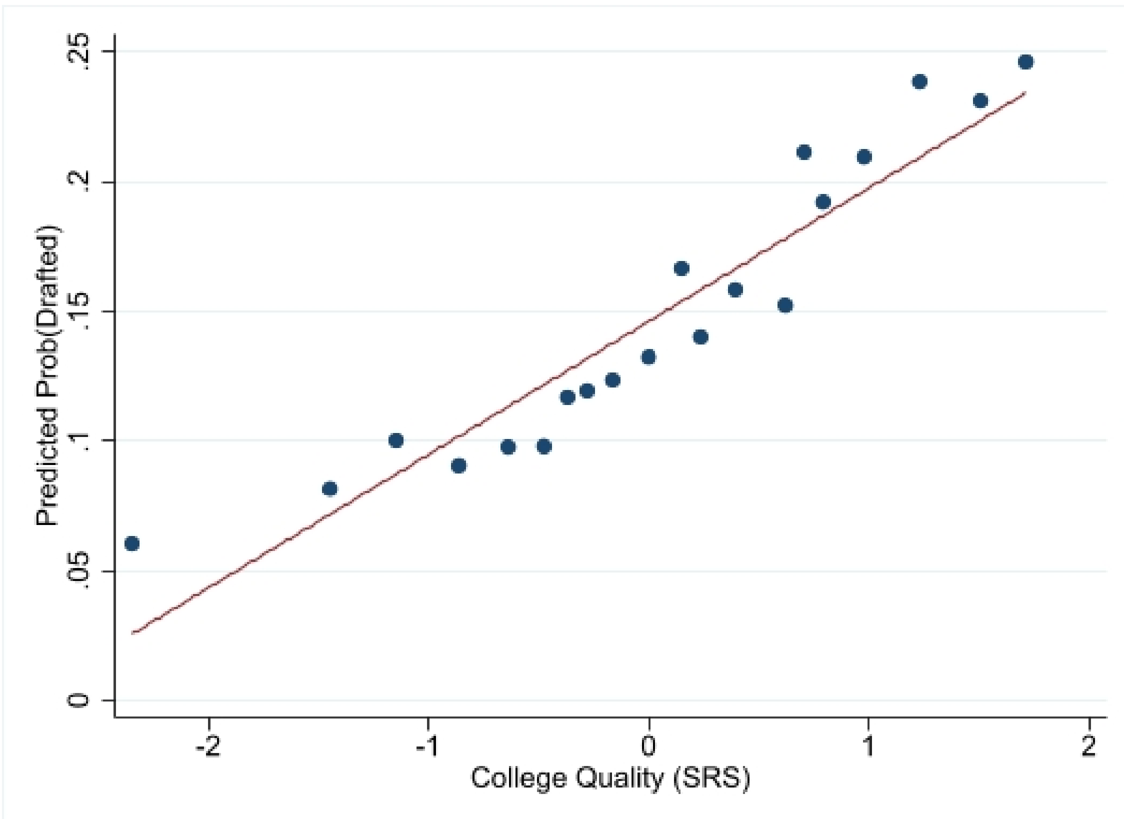

Figure 6: Predicted Probability of Selected in NFL Draft by College Quality

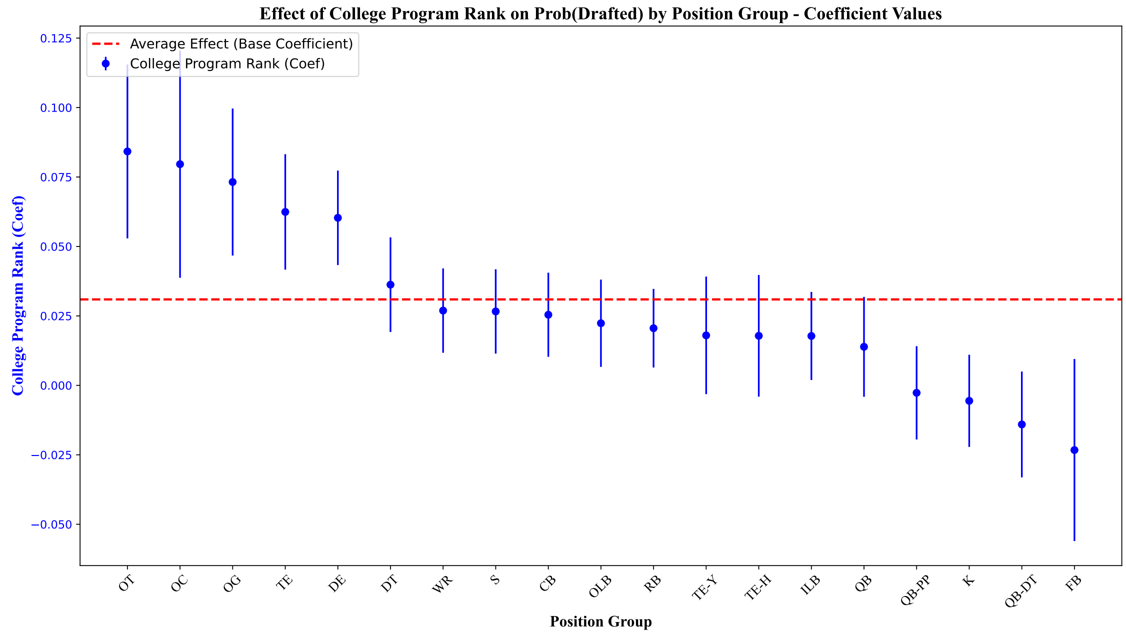

Figure 7: Effect of College Program Rank on Prob(Drafted) by Position Group - Coefficient Values

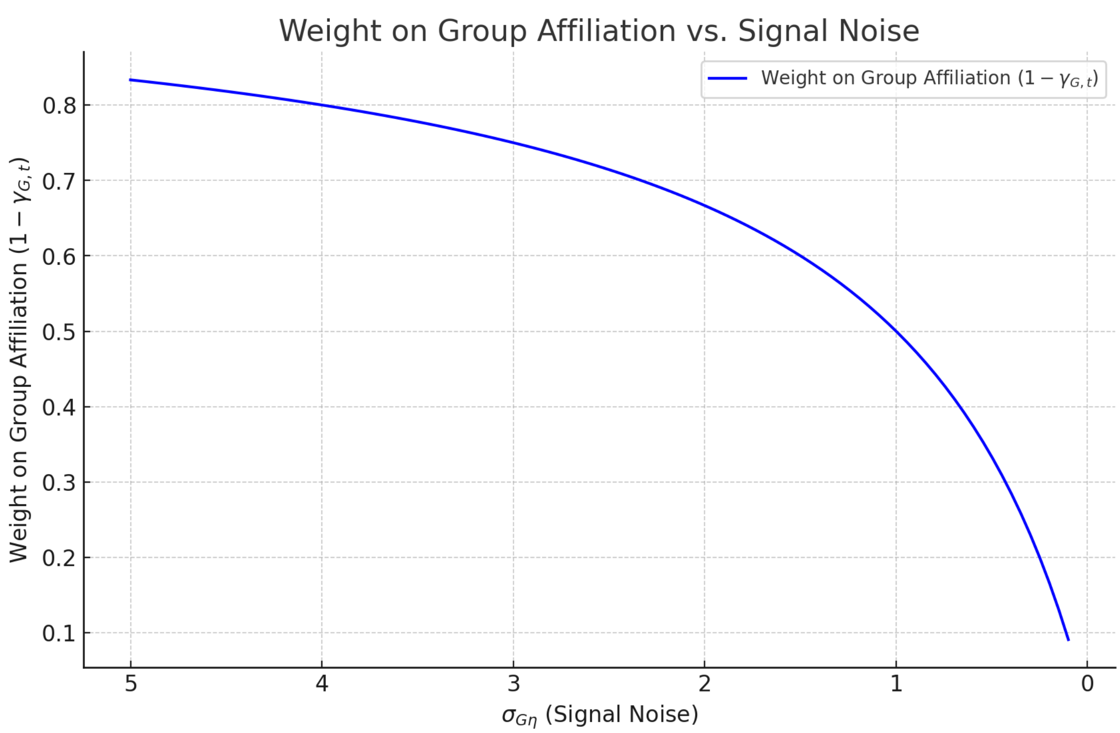

Figure 8: Model Prediction - Changes in Information Quality

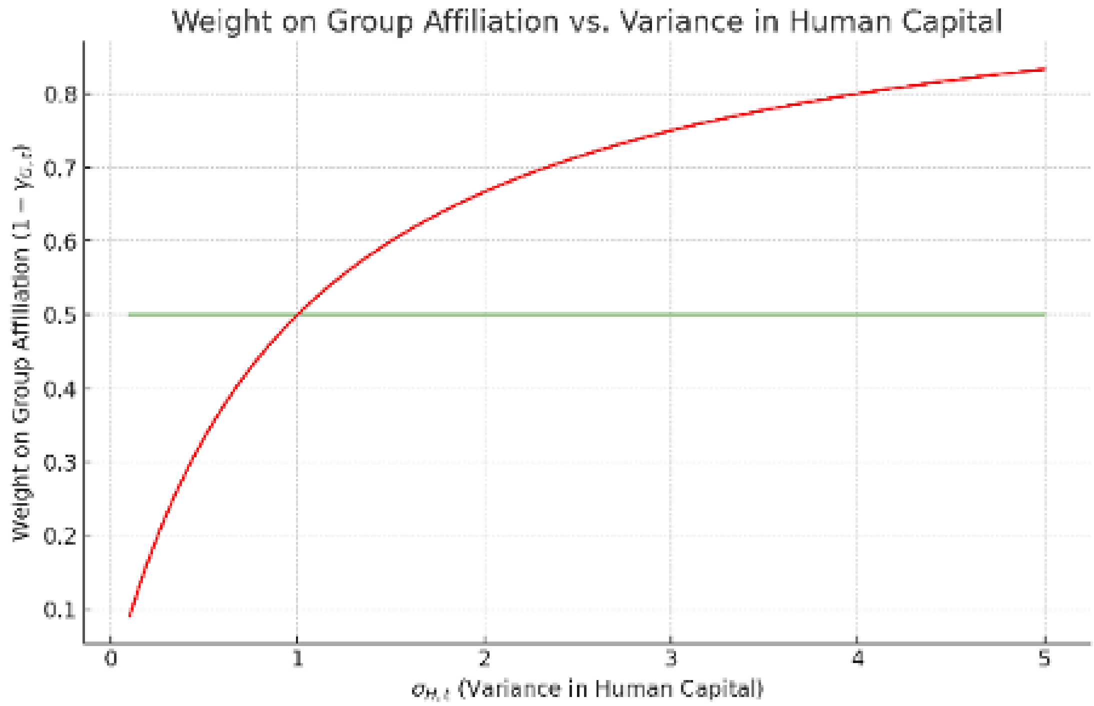

Figure 9: Model Prediction - Changes in Training Quality

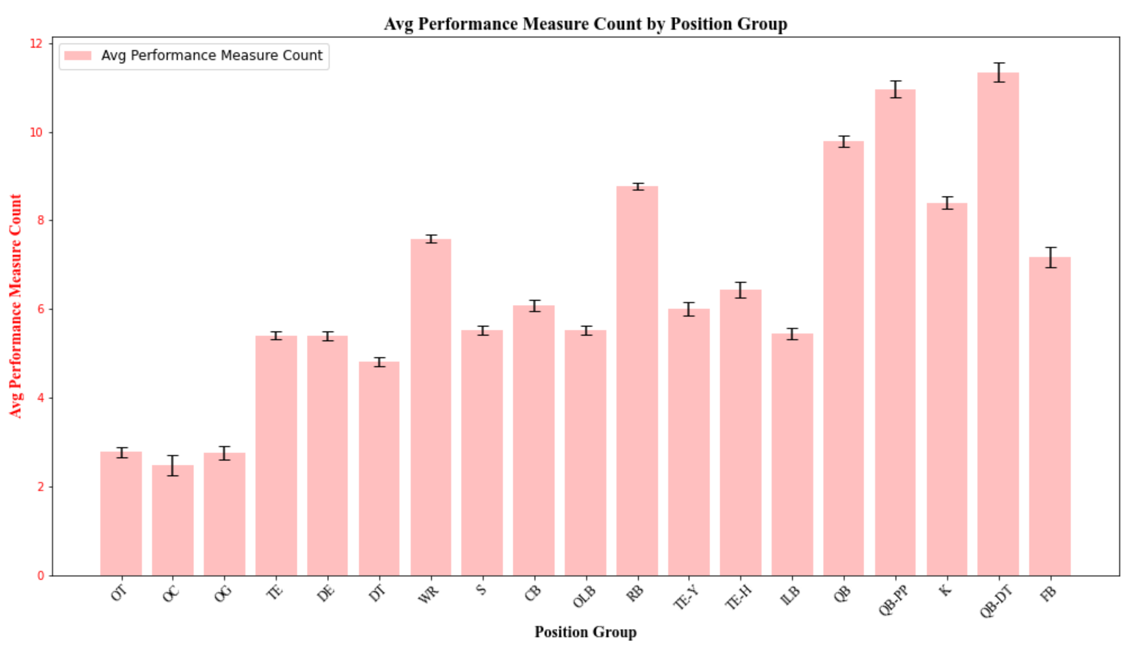

Figure 10: Variation in Measured Performance by Position Group

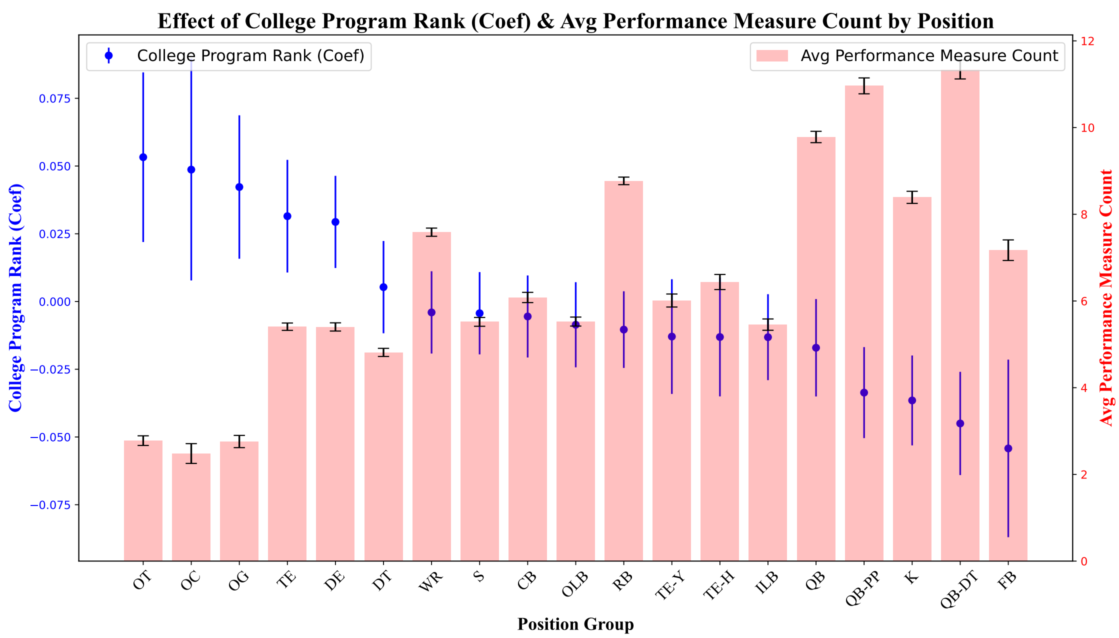

Figure 11: Heterogeneous Effects & Variation in Information Accuracy

Category	Unique Stat Types
Defensive	QB HUR, SOLO, SACKS, PD, TFL, TOT, TD
Fumbles	REC, LOST, FUM
Interceptions	YDS, AVG, TD, INT
Kick Returns	YDS, AVG, NO, LONG, TD
Kicking	XPA, FGM, PCT, LONG, FGA, XPM, PTS
Passing	YPA, COMPLETIONS, INT, PCT, ATT, YDS, TD
Punt Returns	YDS, NO, AVG, LONG, TD
Punting	YDS, LONG, TB, YPP, In 20, NO
Receiving	YPR, YDS, REC, TD, LONG
Rushing	CAR, YDS, TD, YPC, LONG

Table 1: Statistical Performance Measures by Category

High School Athlete Characteristics	Mean	Std. dev	Min	Max
ESPN 300 HS Rank	46.42	28.69	1	100
ESPN 300 HS Athlete Grade	77.03	4.49	44	95
HS Graduation Year	2014	4.84	2006	2022
Total Scholarship Offers	8.67	7	1	89
Height	73.95	2.46	65	82
Weight	221.74	43.5	43	396
Num Top Recruit Peers	12.08	8.05	0	30
Accepted Scholarship Offer	0.90	0.29	0	1
Selected in NFL Draft	0.06	0.24	0	1

Table 2: Summary Statistics ESPN 300 HS Athletes

College Football Program Characteristics	Mean	Std. dev	Min	Max
College Team Start Year	1912.74	20.94	1869	1975
Number of Years	106.67	21.18	19	133
Total Games Played	1127.53	201.67	218	1356
Wins	638.15	170.72	105	961
Loss	451.07	107.66	82	675
Win/Loss Percentage	0.58	0.09	0.348	0.764
Simple Rating System	5.35	5.46	-13.41	14.73
Strength of Schedule	2.35	3.06	-7.75	6.21
Years Ranked in Top 25	24.44	17.30	0	62
Conference Championships	13.84	11.27	0	49

Table 3: Summary Statistics College Football Program Characteristics

Selected in NFL Draft	(1) **College Quality**	(2) **HS Ability + College Quality**	(3) **HS Ability + Peers + College Quality**	(4) **Scholarship Offer sets**
College Team Quality (SRS)	0.043***	0.027***	0.024***	0.018***
	(0.00)	(0.00)	(0.00)	(0.00)
HS Athletic Ability		0.026***	0.025***	0.028***
		(0.00)	(0.00)	(0.00)
Num Top Recruits in Cohort			0.007**	0.007**
			(0.00)	(0.00)
Athlete Controls (Height, Weight)	✓	✓	✓	✓
Scholarship Offerset Controls				✓
Mean Drafted	0.072	0.069	0.067	0.056
	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.058	0.077	0.078	0.088
N	20,260	20,260	20,260	20,260

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Returns to Elite Sports Programs

	College Quality Measure Coefficient	Std. Error	R^2
Winning Percentage	0.013***	(0.00)	0.043
Simple Rating System	0.019***	(0.00)	0.043
Strength of Schedule	0.013***	(0.00)	0.043
Num Pro Players	0.044***	(0.00)	0.054
Years Rank Top 25	0.019***	(0.00)	0.044
Num Conference Champs	0.008***	(0.00)	0.042
Post-Season Win Percent	0.003*	(0.00)	0.041

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Sensitivity Analysis - Alternative Measures of College Program Quality

Table 6: College Quality and Matching Models

	(1) Self Revelation	(2) Match Model 1	(3) Match Model 2	(4) Match Model 3	(5) Match Model 4	(6) Match Model 5
College Quality (SRS)	0.018***	0.018***	0.017***	0.016***	0.018***	0.014***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Athlete Controls	✓	✓	✓	✓	✓	✓
Scholarship Offer-set Controls	✓					
R^2	0.045	0.051	0.092	0.100	0.183	0.194
N	20,298	20,331	20,256	20,289	18,589	16,075
Groups	-	32	908	1,011	4,127	6,575

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Regression Results for Selection in NFL Draft

Selected in NFL Draft	(1) **College Quality**	(2) **HS Ability + College Quality**	(3) **HS Ability + Peers + College Quality**	(4) **Scholarship Offer sets**	(5) **College Performance**
College Program Quality	0.058***	0.039***	0.032***	0.025***	0.027***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
HS Ability		0.034***	0.032***	0.038***	0.042***
		(0.00)	(0.00)	(0.01)	(0.01)
Num Top Recruit Peers			0.015***	0.014**	0.017***
			(0.01)	(0.01)	(0.01)
College Performance					0.055***
					(0.01)
Athlete Controls (Height, Weight)	✓	✓	✓	✓	✓
Scholarship Offer-set Controls				✓	✓
Mean Drafted	0.11	0.11	0.11	0.11	0.11
r2	0.033	0.043	0.044	0.052	0.091
N	10,859	10,859	10,859	10,859	10,859

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$