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

Jordan Holbrook *

November 24, 2024

Please click here for latest version.

Abstract

This study constructs a novel panel dataset of talented 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, by exploiting 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.

JEL: I21, J31, J45, O15
Keywords: Education, Human capital, Signaling

*Jordan Holbrook: Department of Economics, University of Houston, Houston, Texas, USA (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 and feedback from: Angelo Santos, Horacio Rueda, Yujie Zhang, & Charter Sevier.

1 Introduction

Each year, over 1.2 million high school students compete in the United States' most popular sport, American football, many aspiring to become professional athletes. [1] However, only about 250 athletes are drafted annually into the National Football League (NFL). Notably, nearly 80% of these drafted athletes come from just 20% of college sports programs. What distinguishes these elite programs from others, enabling them to consistently send their student-athletes to the professional level? Do these programs excel in player training and development, or do they have unique strategies for connecting their athletes with potential employers? This study examines the individual returns to participating in elite college sports programs, focusing on their impact on professional sports labor market outcomes.

A deeply rooted American ideology is that of meritocracy, with sports often seen as one of the most meritocratic institutions available (Riess 1980; Spearman, Norwood, and Waller 2016; Tolkin 2024) . At every level—high school, college, and professional—team incentives are ostensibly aligned to field the most talented athletes, maximizing the chances of winning. This implies that the best athletes should be recognized and "promoted" to professional teams based solely on talent, regardless of the college program they attended. However, evidence showing that a significant proportion of professional NFL athletes come from a small number of elite college teams challenges this assumption and motivates my central research questions: Does the quality of a college sports program influence NFL draft outcomes? If college program quality does indeed have an impact, can these returns be explained through a human capital or signaling framework?

One of the oldest and most debated questions in economics is measuring the returns to education across different programs. This question is conceptually and empirically challenging, as it requires addressing both observable and unobservable factors that influence the non-random selection of individuals into schools or colleges (Blackburn and Neumark 1993; Loury and Garman 1995; Black and Smith 2006; Brewer, Eide, and Ehrenberg 1999; Chevalier et al. 2004). Seminal work by Dale and Krueger (2002) introduced the matched applicant method to account for selection bias, using students' sets of college acceptance and rejection decisions to estimate the causal effects of college selectivity. Using this method, they found that initial estimates of a large earnings premium from attending elite colleges fade to zero

1. Source: "High School Football Participation Grows Above 1 Million Players," High School Football America. Available at: highschoolfootballamerica.com

once unobserved ability is accounted for. Since then, researchers have applied this method to elite programs with limited consensus: Mountjoy and Hickman (2021), find that earnings premiums diminish over time, others, such as Chetty, Deming, and Friedman (2023), highlight the significant advantages of Ivy-Plus colleges for students reaching top-earning positions.

The question of private returns to education is closely tied to another fundamental debate in economics: whether these returns are driven by human capital accumulation or signaling. Human capital theory, posits that education directly enhances productivity (Becker 1964), while signaling theory suggests that education primarily serves as a signal of inherent traits like ability or perseverance (Arrow 1973; Spence 1981). Both theories predict positive returns to education, making it difficult to disentangle their effects empirically. The debate remains an area of active research, perhaps due to the distinctive policy implications of each theory as well as by identifying mechanisms enhances our understanding of the labor market (Weiss 1995; Harmon and Oosterbeek 2000; Arcidiacono, Bayer, and Hizmo 2010; Arteaga 2018; Belleci 2022; Aryal, Bhuller, and Lange 2022).

The sports economics literature has focus on how talent evaluation, program affiliation, and market mechanisms influence career outcomes. A consistent challenge is distinguishing between individual productivity and group affiliation effects. Research by Gregory-Smith, Bryson, and Gomez (2023) and Ducking, Groothuis, and Hill (2015, 2017) finds no robust evidence of racial discrimination in the NFL, while earlier studies by Hendricks, DeBrock, and Koenker (2003) and Kitchens (2015) emphasize the role of statistical discrimination, showing that student athletes from highly ranked college teams are drafted earlier conditional on having similar levels individual ability. Similarly, Massey and Thaler 2013 document top draft picks are systematically overvalued in the NFL draft, showing there is precedent for potential inefficiencies in this labor market.

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 athletic 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 see for more on test scores and labor market outcomes (Currie and Thomas 2001; Rose 2006). 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 (Hedlund 2018).

In this paper, I build a novel panel data set of talented 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 from, I adapt the matched applicant or Dale and Krueger method, to address threats to identification from selection on unobservable characteristics. With this data and empirical strategy, I provide selection-corrected estimates of the returns to elite sports programs on the main outcome variable of job placement. Different from previous studies using the matched applicant method, 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 model of employer learning. 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. For position groups where productivity is easier to infer due to more accurate private signals, the effect of college program rank is less meaningful. In contrast, for position groups with less accurate private signals, attending an elite program plays a more significant role. In such cases, employers rely more heavily on the prestige of the athlete's college program as a proxy for ability, consistent with the predictions of the signaling model.

This study contributes to the literature on the effects of elite colleges and programs by examining the private returns to elite college sports programs. First, it demonstrates that selection plays a significant role among student-athletes, similar to traditional students, in determining college placement. Second, it shows that scholarship offer set portfolios serve as a strong proxy for individual fixed effects, sufficiently accounting for unobserved ability that drives the non-random selection of high school students to colleges. This finding confirms the validity of applying the Dale and Krueger method to student-athletes, aligning with prior research by Mountjoy and Hickman (2021) and Chen, Grove, and Hussey (2013). Third, the study extends the Dale and Krueger framework to a larger and more complex choice set, as student-athletes have more complex offer portfolios than traditional college applicants. Finally, even after controlling for selection bias using the matched applicant method, a substantial premium for attending elite sports programs persists. This result resonates with Chetty, Deming, and Friedman (2023), that the effects of elite programs are concentrated at the upper end of the distribution, particularly in job placement within the NFL.

This study contributes to the signaling and employer learning literature by leveraging the unique structure of the college football labor market, where athletic performance serves as a key signal used by employers in hiring decisions. The study demonstrates that signal accuracy varies systematically across position groups, creating a natural environment where the signaling and human capital frameworks generate differing predictions about the role of college programs in shaping student-athlete outcomes. By incorporating a model of employer learning, the analysis tests specific predictions from signaling theory and shows that these ex-ante predictions align with the observed heterogeneous effects of college program prestige on professional outcomes. This approach shows the value of signaling theory to generate accurate predictions in situations with differing information levels across groups. This evidence aligns with Aryal, Bhuller, and Lange (2022), showing that human capital and signaling can both be impactful as well as decomposed into measured effects.

My study builds on and extends the sports economics literature in three ways. First, I assemble one of the largest datasets of high school athletes, creating a unique resource to analyze athlete development and career outcomes. Second, I apply signaling theory to the context of college football, offering novel insights into how performance metrics and program affiliation influence professional outcomes, particularly in the NFL draft. My findings reveal a significant selection effect, with elite programs disproportionately represented in NFL job placements, supporting Kitchens 2015 but challenging the notion of purely meritocratic selection. Finally, I measure the returns to elite sports programs at the extensive margin—specifically, who reaches the NFL draft for all position groups— as opposed to most research on college football which only studies the sample of college athletes who will be drafted in the NFL or only looks at select position categories (Hendricks, DeBrock, and Koenker 2003; Pitts and Evans 2019; Keefer 2016).

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 5 reports the results and main findings, and section 6 investigates the theoretical mechanisms underlying the results and section 7 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:

$$E[Y] = r_A - r_B \tag{2.1}$$

$$y_i = r_A - r_B + e_i \tag{2.2}$$

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.

$$X^T X r = X^T y \tag{2.3}$$

$$\sum_{i=1}^{n} r_i = 0 \tag{2.4}$$

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

$$\beta = \underbrace{(X^T X)^{-1}}_{\text{Strength of Schedule}} \cdot \underbrace{X^T y}_{\text{Average Margin of Victory}} \tag{2.5}$$

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 and Keefer (2016, 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 positions and categories, performance measures are standardized using the formula:

$$Z_{ijg} = \frac{X_{ijg} - \mu_{jg}}{\sigma_{jg}}, \quad C_{ig} = \frac{1}{n_g} \sum_{j=1}^{n_g} Z_{ijg}, \quad C_i = \frac{1}{G} \sum_{g=1}^{G} C_{ig}, \quad S_i = \frac{1}{m} \sum_{s=1}^{m} C_{is}.$$

Here, Z_{ijg} is the standardized score for player i, measure j, and category g, where X_{ijg} is the raw performance measure, and μ_{jg} and σ_{jg} are the mean and standard deviation of measure j in category g. C_{ig} aggregates standardized scores within a category (n_g is the number of measures), while C_i averages category-specific scores to derive an overall performance score. Finally, S_i is the seasonal aggregation of scores over m seasons.

3 Setting and Descriptive Statistics

3.1 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.

3.2 Descriptive Statistics

3.2.1 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.

3.2.2 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 4 discussing how to overcome these challenges and provide selection-corrected estimates of the private returns to elite sports programs.

4 Empirical Strategy

My empirical analysis begins by adapting the matched-applicant model first developed in Dale and Krueger 2002 with similar frameworks used in Ge, Isaac, and Miller 2022, Chen, Grove, and Hussey 2013, Chetty, Deming, and Friedman 2023, and Mountjoy and Hickman 2021. 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:

$$y_{ij} = \beta_0 + \beta_1 SRS_j + \beta_2' X_{1i} + \underbrace{\beta_3' X_{2i} + \epsilon_{ij}}_{u_{ij}} \tag{4.1}$$

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 4.1 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.

4.1 Adapting to Sports Research Setting

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, Isaac, and Miller 2022, Chen, Grove, and Hussey 2013, Chetty, Deming, and Friedman 2023, and Mountjoy and Hickman 2021 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 and Krueger 2002 Ge, Isaac, and Miller 2022, Chen, Grove, and Hussey 2013, Chetty, Deming, and Friedman 2023, and Mountjoy and Hickman 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. Both Chetty, Deming, and Friedman 2023 and Mountjoy and Hickman 2021 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 exogenous, 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.

4.2 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 Matched Applicant Model. I define the Matched Scholarship Model as follows:

$$y_{ijg} = \beta_0 + \beta_1 SRS_j + \beta_2' X_{1i} + \sum_{1}^{m} \gamma_g Group_{ig} + \epsilon_{ijg} \tag{4.2}$$

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:

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

For comparison, Dale and Krueger 2002 and Ge, Isaac, and Miller 2022 set bins 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.

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:

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

In other words, treatment variation occurs when, within a group g, there exists meaningful heterogeneity in the colleges attended by individuals with identical scholarship offer sets. Bringing this condition to the data, a useful approach to identifying meaningful heterogeneity within matching scholarship offer set groups is the examine the variance of sports program quality measures (SRS) within a given offer set. More formally, let j_i represent the college team j that individual i chooses to attend. Treatment variation exists in group g if, for a matching set θ_i, the variance of the sports quality measure (SRS) of the colleges attended by individuals is strictly greater than zero. Mathematically, this condition is defined as:

$$\text{Var}\left(\{\mathsf{SRS}_{j_i} \mid i \in g, \theta_i = \theta_g\}\right) > 0$$

This condition implies that at least one individual within the group chose a college with an SRS score different from the others, ensuring variation in treatment. Conversely, if the variance equals zero, all individuals in the group attended colleges with the same SRS score, and no treatment variation exists.

Let ν_g be a binary indicator reflecting whether treatment variation exists within group g. The indicator is defined as:

$$\nu_g = \begin{cases} 1 & \text{if Var}\left(\{\mathsf{SRS}_{j_i} \mid i \in g, \theta_i = \theta_g\}\right) > 0, \\ 0 & \text{if Var}\left(\{\mathsf{SRS}_{j_i} \mid i \in g, \theta_i = \theta_g\}\right) = 0. \end{cases}$$

This definition ensures that groups with $\nu_g = 1$ exhibit treatment variation, making them valid for analysis, while groups with $\nu_g = 0$ are excluded due to lack of meaningful differences in college choices.

Now, I define Group_{ig}, the matching sets of dummy variables dependent on two factors: first, scholarship offers from the same school or same tier of school, and second, treatment variation within the set of athletes with similar offer sets. The group indicator variable Group_{ig} is defined as:

$$\mathsf{Group}_{ig} = \begin{cases} 1 & \text{if } \theta_i = \theta_g \text{ and } \nu_g = 1, \\ 0 & \text{if } \theta_i \neq \theta_g \text{ and } \nu_g = 1, \\ \text{undefined} & \text{if } \nu_g = 0. \end{cases}$$

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 regression model.

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. In other words, conditional on similar offer sets the individual student athletes college enrollment decision is random. This may seem to a be stringent assumption as there are many potential factors that can influence athlete enrollment decision such as home bias, coaching relationships, or team style of play. Factors that could prove problematic for this identification strategy have to effect individual enrollment decisions as well as meaningfully influence the key outcome studied in this paper, the likelihood of becoming a professional athlete in the NFL draft. Showing any individual factor satisfies both of these conditions is extremely challenging, thus the enrollment decision in this study is taken as given and seen to be plausibly exogenous.

5 Empirical Results

5.1 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) and Dale and Krueger (2014), 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.

5.2 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), 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.

5.3 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.

5.4 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 seemingly orthogonal to athletic performance as coefficient value for sports program quality, β_{SRS} remains persistent and unchanged.

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.

5.5 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.

6 Exploring Mechanisms

6.1 Signaling Theory & Employer Learning

The debate over whether education and credentials function primarily as signals of ability or as mechanisms for human capital accumulation has persisted for decades. Human capital theory, posits that education enhances productivity by equipping individuals with skills and knowledge (Becker 1964). In contrast, signaling theory, argues that education serves as a marker of inherent traits, such as ability or perseverance, which are unobservable to employers (Arrow 1973; Spence 1981). Both theories predict positive returns to education, making it challenging to disentangle their effects empirically (Weiss 1995).

Numerous studies have sought to address this question, with mixed results. The persistence of this literature is perhaps due to the policy implications if education or college programs serve mainly as a signal or whether human capital development drives much labor market success. Huntington-Klein (2021) highlights the difficulties in separating human capital and signaling effects due to overlapping mechanisms. For instance, while Chevalier et al. (2004) reject pure signaling models, they find that human capital explanations alone are insufficient. Testing these theories is further complicated by confounding factors such as accounting variation in individual ability, employer practices, and labor market conditions.

Researchers have attempted to isolate signaling effects by focusing on the availability and accuracy of information about workers. Early work by Riley (1979) and Albrecht (1981) demonstrated that employers rely on educational signals when information about candidates is limited. They focused on variation related to information about workers as information is the key distinguishing feature of the signaling hypothesis. Human capital theory does not have predictions about how information accuracy effects returns to education. In the literature, it is difficult to find comparable groups of workers with different levels of information availability that are not also dissimilar in other ways (e.g., occupation or education level) (Riley 1979).

The employer learning literature has provided a framework to investigate how information about workers changes over time. This body of work examines how employers update their beliefs about worker productivity over time based on observable performance (Farber and Gibbons 1996; Altonji and Pierret 2001; Altonji 2005). For instance, these models predict that as employers gather more information (learn) about workers, the importance of observable characteristics like education diminishes, and unobserved traits become more influential (Schönberg 2007). While these studies offer valuable insights, they often rely on time-series data, which introduces confounding factors such as changes in worker productivity, job training, and employee-employer matching (Pinkston 2009; Ge, Moro, and Zhu 2021).

Sports, provides a unique setting to gain traction exploring these ideas. Unlike traditional labor markets, sports data offers highly specialized performance metrics that are publicly observed. Importantly, these metrics are measured consistently across all levels of play—high school, college, and professional leagues—eliminating concerns about industry or major-specific biases that complicate other studies. These unique tasked based performance measures alone, however, do not resolve all the challenges in studying the signaling hypothesis. American Football has a unique production function, called by some ”the only true team sport”[2], because not every player is eligible to score points. Only some position groups get the opportunity to score points and players can even be penalized for scoring points if they are not part of an eligible position group. Thus, this unique production function of the game provides significant variation in the types and quantities of performance data recorded for different position groups. For example, the quarterback position group often has 10 to 15 distinct performance metrics recorded, while offensive linemen may have only a few. This variation allows for precise comparisons of how information availability affects labor market outcomes.

The structure of American football also creates unique information asymmetries. Different position groups have varying levels of performance data available, which allows for within-program variation in the same year and before hiring occurs. To replicate this same type a variation with traditional students a hypothetical scenario would be where all students take an exit exam that predicts job performance, but the exam is graded with varying levels of accuracy for different groups. Such controlled variation in information accuracy is absent in other sports, like basketball or soccer, where all players can score in similar ways.

Building on these insights, this section shows the value of a simple model of employer learning and statistical discrimination. There are two main goals: (1) to assess whether the large return to elite sports programs aligns more closely with human capital or signaling framework, and (2) to explain the significant heterogeneity in program effects across position groups as discussed in 5.5. This model draws on the theoretical foundations of employer learning (Farber and Gibbons 1996; Altonji and Pierret 2001; Altonji 2005; Pinkston 2009), while focusing on the role of information accuracy, discussed in Riley (1979) and Albrecht (1981). The next section formalizes this approach.

6.2 Model Setup

I begin with a basic model of employer learning 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.

2. https://hairinabiscuit.com/blog/f/football-the-only-team-sport

Employers observe a set of applicants with two pieces of information: a private signal S_i, and the group affiliation V_G of potential employee's. The productivity of applicant i from school G is defined as:

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

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

Where $V_{G,i}$ is the productivity of applicant i, and $\epsilon_{G,i}$ represents the idiosyncratic productivity above the group effect. For simplicity, productivity process follows a normal distribution.

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:

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

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

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

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

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

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 value $S_{G,i}$ when predicting productivity. Conversely, if the signal is less accurate, the principal will place more weight on the group affiliation V_G.

This model represents the tradeoffs faced by employers when trying to make hiring decisions. The main challenge for employers being: 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.

6.3 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 that a model of employer learning can incorporate both signaling elements and human capital accumulation.

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 redefined in this extended context. I 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. The adapted equation is modeled as,

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

where 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, and finally $H_{G,i,t}$ 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 of this model is given by:

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

In this equation, $H_{G,i,0}$ represents the initial human capital of individual i, while $I_{G,i,s}$ denotes the investment in human capital, such as training or education, made at time s. The parameter α captures the rate at which these investments translate into increases in human capital. Lastly, $\eta_{G,i,t}$ represents the random shocks affecting human capital accumulation at time t.

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:

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

with $\sigma^2_{H,t}$ is the variance in human capital accumulation at time t and $\sigma^2_{G\eta,t}$ 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, this is possibly due to 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.

6.4 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 variation in performance measures emphasizes 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.

6.5 Signaling Model 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 Jackson 2024 for a full discussion.

6.5.1 Evidence Supporting Signaling

Figure 11 provides strong evidence that the returns to attending elite sports programs are consistent with a signaling model. 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 defensive tackles (DT)—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.

A purely human capital model struggles to account for key patterns observed in the data. First, it cannot explain the significant heterogeneity in the effects of college program quality on student-athletes who are part of the same program, in the same year, on the same team, facing the same opponents, and receiving the same coaching and training. Despite these shared conditions, the impact of college program quality varies widely across position groups. Second, the human capital model predicts that position groups spending more time on the field—and thus receiving more "on-the-job training"—should experience greater benefits from elite program quality. Under this framework, offensive linemen (OT, OG) and quarterbacks (QB), who are both on the field for comparable durations, would be expected to exhibit similar program effects. However, the data reveal opposing effects, with offensive linemen experiencing much larger benefits than quarterbacks.

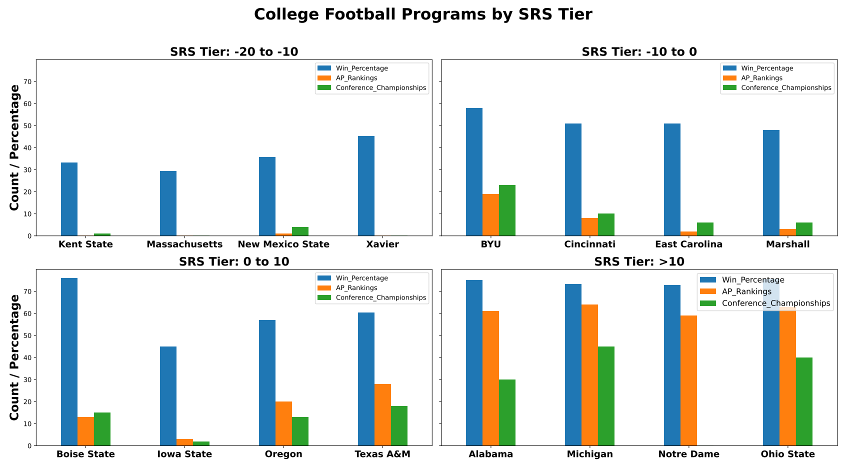
Furthermore, as discussed in Sections 5.4 and 7, controlling for individual athletic performance did not alter the coefficient estimates of college program quality. This suggests that elite programs have a limited impact on improving individual performance metrics, further challenging the applicability of the human capital model in this context.

7 Conclusions

In conclusion, this study addresses the challenges of estimating the returns to elite college sports programs by focusing on highly recruited high school athletes and their labor market outcomes. Using a unique dataset of these athletes and a robust empirical strategy, I find substantial returns to elite college football programs in terms of job placement in the NFL. The analysis shows that athletes from top-ranked programs are significantly more likely to be drafted, with the effect varying widely by position group. For positions with less individual performance data, such as offensive linemen, the prestige of the college program plays a larger role, consistent with a signaling framework. On the other hand, for positions with more detailed performance measures, like quarterbacks, the effect of college program rank is less pronounced. These findings suggest that the returns to elite sports programs are driven more by signaling than by human capital.

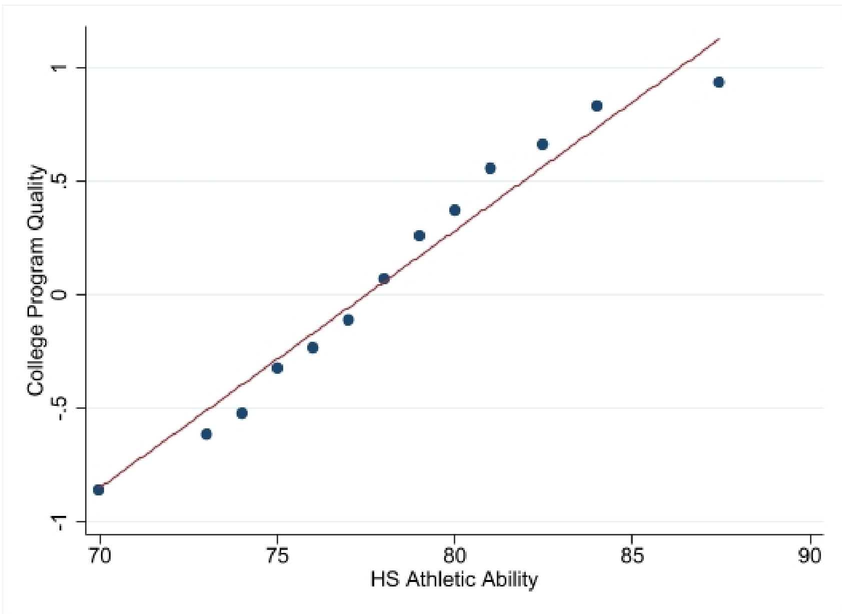
Tables and Figures

Fig. 1. College Football Programs by Simple Rating System Tier



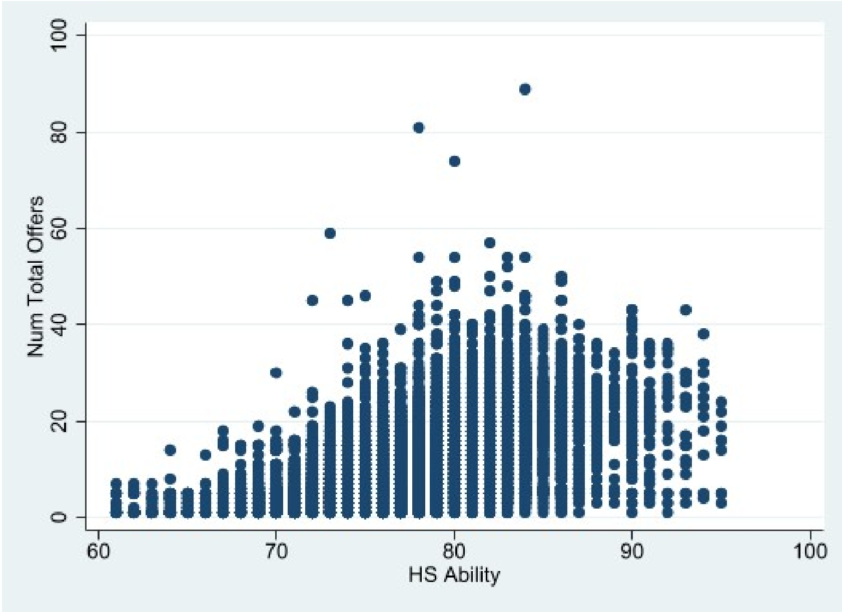
Note: This figure presents the distribution of college football programs categorized by their Simple Rating System (SRS) tiers. SRS is a composite measure of team quality that adjusts for strength of schedule and other factors.

Fig. 2. Selection into College Football Programs by ESPN 300 HS Athletes



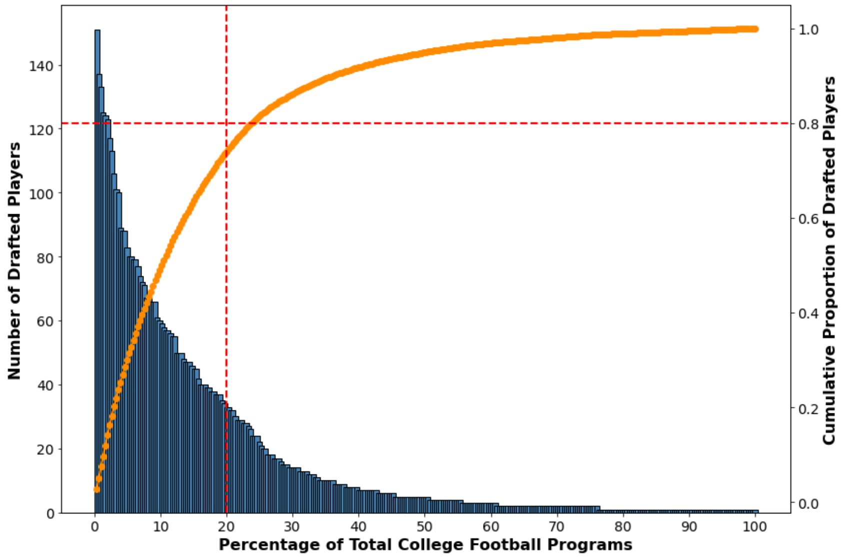
Note: This figure shows the distribution of selection into college football programs from ESPN 300 high recruited athletes. Bin scatter plot created in Stata with automatic bin widths across the distribution of athletic ability.

Fig. 3. Distribution of Scholarship Offers to ESPN 300 HS Athletes



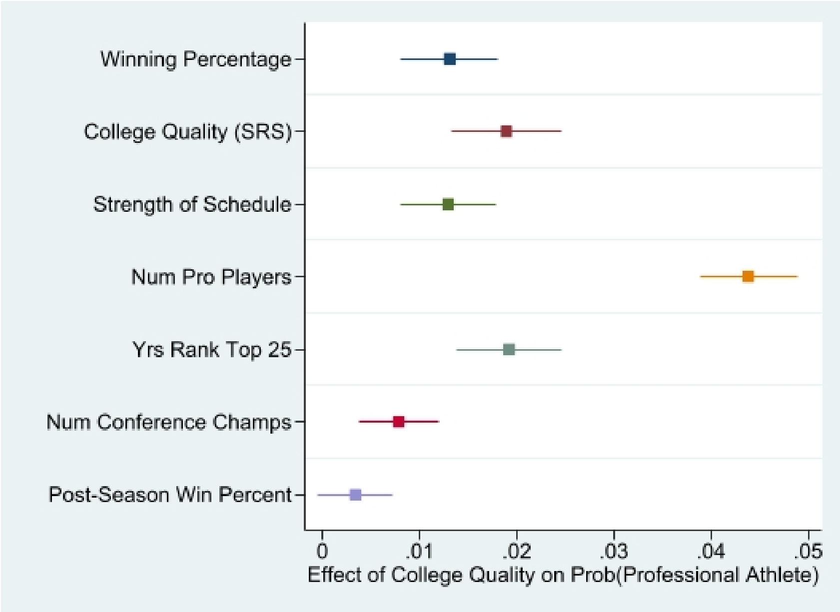
Note: This figure shows the distribution of scholarship offers received by ESPN 300 high school athletes, disaggregated by the quality of the college football programs offering the scholarships.

Fig. 4. Concentration of NFL Talent by College Football Program



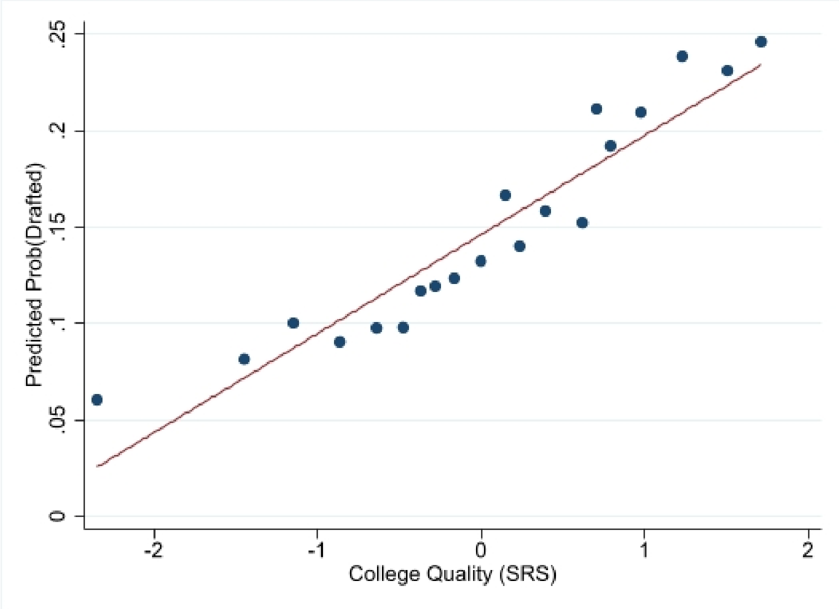
Note: This figure illustrates the concentration of NFL talent across college football programs, highlighting the disproportionate contribution of a small percentage of programs to the overall NFL talent pool.

Fig. 5. Sensitivity Analysis - College Program Quality Measures



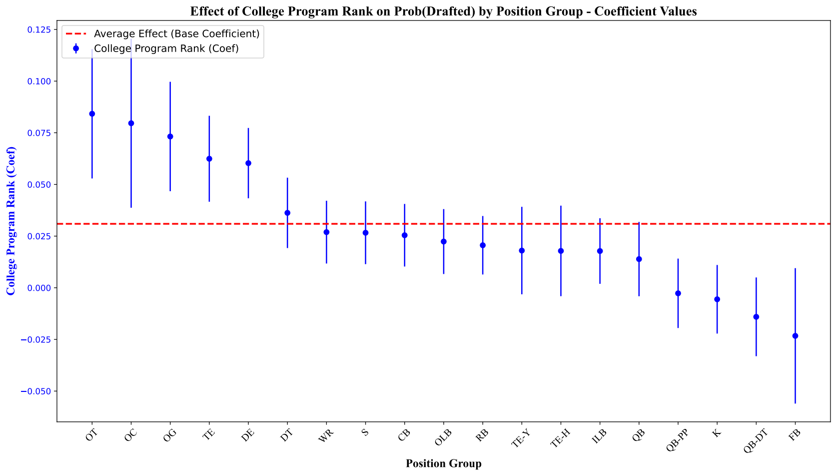
Note: This figure presents a sensitivity analysis of various college program quality measures and their impact on athlete outcomes, showing the robustness of results across different specifications.

Fig. 6. Predicted Probability of Selection in NFL Draft by College Quality



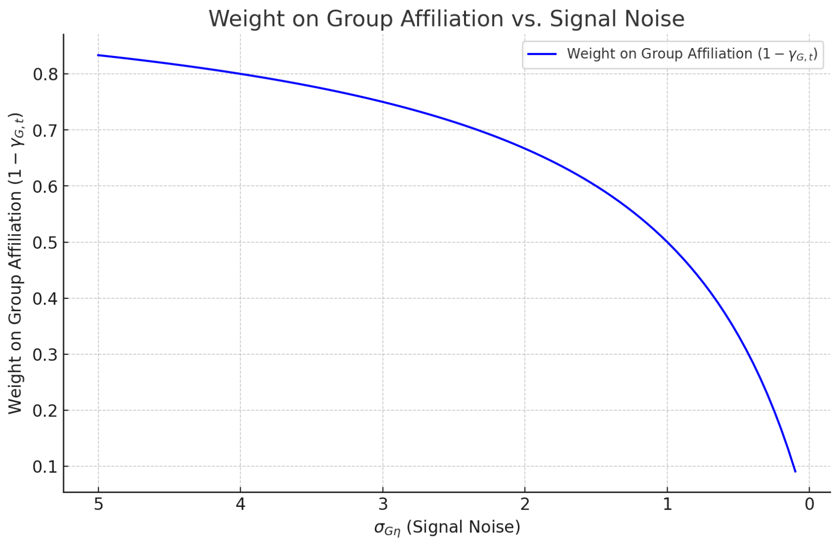
Note: This figure plots the predicted probability of being selected in the NFL Draft based on the quality of the college football program attended, with higher-quality programs associated with a greater likelihood of selection.

Fig. 7. Effect of College Program Rank on Probability of NFL Draft by Position Group - Coefficient Values



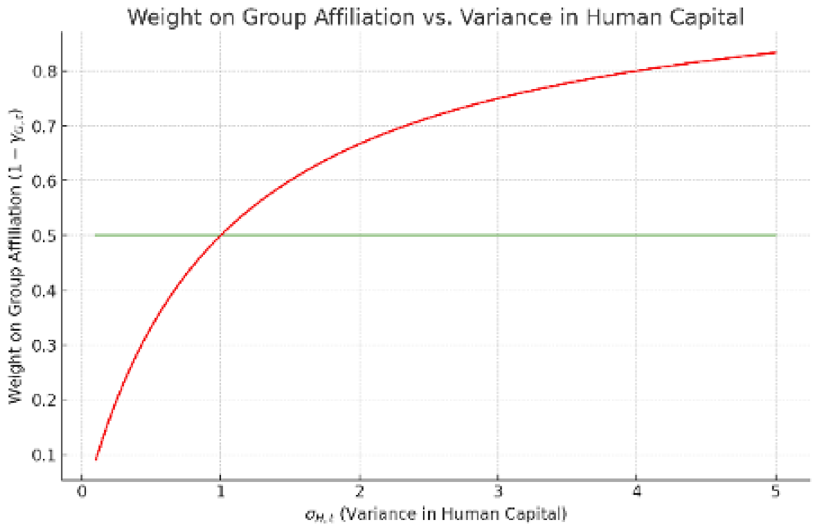
Note: This figure shows the effect of college program rank on the probability of being drafted into the NFL, disaggregated by position group, highlighting heterogeneity in draft outcomes by playing position.

Fig. 8. Model Prediction - Changes in Information Quality



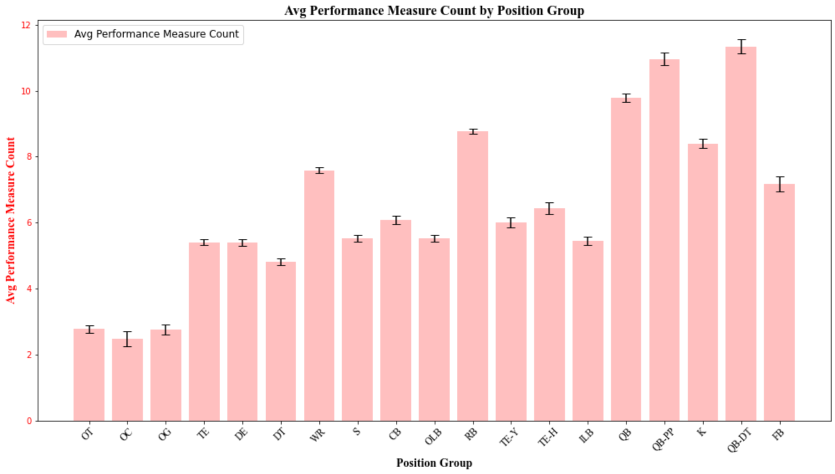
Note: This figure presents a model prediction of how changes in information quality affect outcomes in the decision-making process, demonstrating the impact of signal accuracy on selection into elite programs.

Fig. 9. Model Prediction - Changes in Training Quality



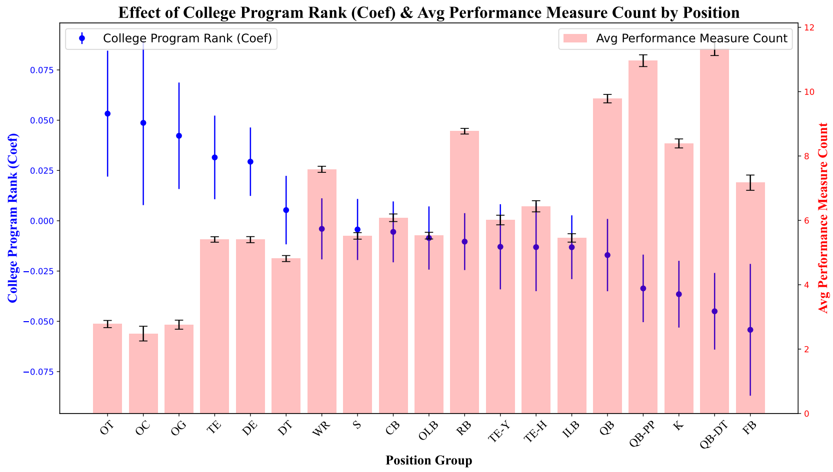
Note: This figure presents a model prediction of how changes in training quality impact performance and subsequent outcomes, illustrating the role of human capital development in athletic success.

Fig. 10. Variation in Measured Performance by Position Group



Note: This figure shows the variation in measured performance across different position groups in college football, providing insights into the differences in performance metrics by playing position.

Fig. 11. Heterogeneous Effects & Variation in Information Accuracy



Note: This figure illustrates the heterogeneous effects of varying levels of information accuracy on athlete outcomes, particularly in predicting NFL Draft selection probabilities based on different performance measures.

Table 1: College Football Athletic Performance Measures by Category

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

Note: The table outlines different categories of athletic performance statistics in football, providing unique stat types for each category. These performance measures track player activities and are typically collected during games by official scorekeepers and analysts. For example, defensive stats include measures like QB HUR (quarterback hurries), SOLO (solo tackles), and SACKS (quarterback sacks), which assess defensive players' impact on the opposing team's offense. Fumble stats like REC (fumble recoveries) and LOST (fumbles lost) track how teams handle ball security. Offensive stats such as Passing YPA (yards per attempt), Rushing CAR (carries), and Receiving YPR (yards per reception) quantify a player's ability to advance the ball.

For example, a player with 1,000 receiving yards (YDS) from 50 receptions (REC) would have a YPR (yards per reception) of 20, showing their efficiency in gaining yards per catch. These statistics offer a way to analyze player contributions and help teams evaluate performance across different aspects of the game.

Table 2: Summary Statistics ESPN 300 HS Athletes, 2006-2021

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

Note: The table presents summary statistics for high school athletes ranked in the ESPN 300 from 2006 to 2021, highlighting key characteristics such as rank, scholarship offers, height, weight, and outcomes like NFL draft selection. The data was sourced from the ESPN 300 Recruiting Database (https://www.espn.com/college-sports/football/recruiting/rankings), which ranks top high school athletes by position, height, weight, grade, and commitment status to NCAA programs. Each year, ESPN evaluates and grades the top 300 high school football recruits based on their performance, potential, and recruitment status.

Table 3: Summary Statistics College Football Program Characteristics

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

Note: Summary statistics cross tabulate for various variables. Table shows mean, standard deviation, min, and max values for each variable measured at the college football program level. The table provides summary statistics on the characteristics of college football programs, focusing on historical performance, rankings, and achievements. The Simple Rating System (SRS), a metric reflecting team performance adjusted for strength of schedule, has a mean value of 5.35. Strength of schedule, which measures the relative difficulty of a team's opponents, averages 2.35. The data comes from Sports Reference (https://www.sports-reference.com/cfb/schools/), which compiles detailed historical statistics on college football teams, including performance metrics, game outcomes, rankings, and championship counts, based on available records from as early as 1869.

Table 4: Returns to Elite Sports Programs

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

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

Notes: The table titled "Returns to Elite Sports Programs" presents the results of four regression models analyzing the factors influencing the likelihood of a high school athlete being selected in the NFL Draft. Each column represents a different model specification, progressively including more controls. In all models, College Team Quality (SRS), measured using the Simple Rating System (SRS), is positively and significantly associated with being selected in the NFL Draft across all specifications. For example, in column (1), a one-unit increase in College Team Quality increases the likelihood of being drafted by 0.043 percentage points, holding other factors constant. This effect remains significant and positive even after accounting for additional variables like high school athletic ability, peer recruits, and scholarship offer sets. Columns (2) and (3) introduce High School Athletic Ability, which is also a significant predictor. The presence of top recruits in the same cohort is introduced in columns (3) and (4), showing a smaller but still significant effect. Scholarship offer controls are added in column (4), suggesting that athletes who received more offers are more likely to be drafted.

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

	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

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

Notes: This table presents a robustness check for the main specification results, replicating the model from column (4) of the main results table but substituting different measures of college quality to assess how they influence the likelihood of becoming a professional athlete. The column College Quality Measure Coefficient reports the coefficients for each alternative measure, while the standard errors and the R^2 values reflect model fit and explanatory power.

Table 6: College Sports Program 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
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: * 0.10 ** 0.05 *** 0.01. This table compares five different matching model specifications with the baseline results to evaluate the impact of scholarship offers on athlete outcomes. Each matching model involves a trade-off between creating more homogeneous groups (which reduces sample size) and retaining larger sample sizes (which leads to less homogeneous groups). For example, Matching Model 1 groups athletes based on a 5-digit binary ID representing the presence or absence of offers from schools ranked in different quintiles, while Models 2, 3, 4, and 5 use increasingly complex methods, such as counting offers and using deciles or terciles, with varying thresholds.
The trade-off is evident in the group and sample sizes, with more restrictive matching models resulting in smaller samples but more homogeneous groups. Despite these differences, the effect of college quality (measured by the Simple Rating System, SRS) on athlete outcomes remains consistent across models, with coefficients ranging from 0.014 to 0.018, confirming the robustness of the findings. Exact matching is not included in the analysis due to a significant loss of sample size (98%), which leads to insufficient treatment variation and limits its usefulness in addressing the research question. Thus, less restrictive matching models provide a more representative analysis of the relationship between college quality and athlete outcomes.

Table 7: Matched Scholarship Model with Athletic Performance

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

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

Note: * 0.10 ** 0.05 *** 0.01. This table presents the results of five regression models estimating the likelihood of being selected in the NFL Draft, with varying controls for college program quality, high school (HS) ability, scholarship offers, and college performance. Each model builds upon the previous one by incorporating additional variables.

References

Albrecht, James W. 1981. "A Procedure for Testing the Signalling Hypothesis." *Journal of Public Economics* 15, no. 1 (February): 123–132. https://doi.org/10.1016/0047-2727(81)90057-8. https://www.sciencedirect.com/science/article/pii/0047272781900578.

Altonji, Joseph G. 2005. "Employer Learning, Statistical Discrimination and Occupational Attainment." *American Economic Review* 95, no. 2 (April): 112–117. https://doi.org/10.1257/000282805774670455. https://pubs.aeaweb.org/doi/10.1257/000282805774670455.

Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination*." *The Quarterly Journal of Economics* 116, no. 1 (February): 313–350. https://doi.org/10.1162/003355301556329. https://doi.org/10.1162/003355301556329.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* 2, no. 4 (October): 76–104. https://doi.org/10.1257/app.2.4.76. https://www.aeaweb.org/articles?id=10.1257/app.2.4.76.

Arrow, Kenneth J. 1973. "Higher Education as a Filter." *Journal of Public Economics* 2, no. 3 (July): 193–216. https://doi.org/10.1016/0047-2727(73)90013-3. https://www.sciencedirect.com/science/article/pii/0047272773900133.

Arteaga, Carolina. 2018. "The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia's Top University." *Journal of Public Economics* 157 (January): 212–225. https://doi.org/10.1016/j.jpubeco.2017.10.007. https://www.sciencedirect.com/science/article/pii/S0047272717301809.

Aryal, Gaurab, Manudeep Bhuller, and Fabian Lange. 2022. "Signaling and Employer Learning with Instruments." *American Economic Review* 112, no. 5 (May): 1669–1702. https://doi.org/10.1257/aer.20200146. https://www.aeaweb.org/articles?id=10.1257/aer.20200146.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition.* NBER. https://www.nber.org/books-and-chapters/human-capital-theoretical-and-empirical-analysis-special-reference-education-first-edition.

Belleci, Robert. 2022. "Correlational Study Examining the Recruitment of High School Football Players in California and the Teams They Play For." Ph.D., Northcentral University. https://www.proquest.com/docview/2741309608/abstract/7F421CDDF6804B8EPQ/1.

Black, Dan A., and Jeffrey A. Smith. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *Journal of Labor Economics* 24 (3): 701–728. https://doi.org/10.1086/505067. https://www.jstor.org/stable/10.1086/505067.

Blackburn, McKinley L., and David Neumark. 1993. "Omitted-ability Bias and the Increase in the Return to Schooling." *Journal of Labor Economics* 11, no. 3 (July): 521–544. https://doi.org/10.1086/298306. https://www.journals.uchicago.edu/doi/abs/10.1086/298306.

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. https://doi.org/10.2307/146304. https://www.jstor.org/stable/146304.

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.* SSRN Scholarly Paper. Rochester, NY, March. https://papers.ssrn.com/abstract=2240305.

Chetty, Raj, David Deming, and John Friedman. 2023. *Diversifying Society's Leaders? The Determinants and Causal Effects of Admission to Highly Selective Private Colleges.* Technical report w31492. Cambridge, MA: National Bureau of Economic Research, July. https://doi.org/10.3386/w31492. http://www.nber.org/papers/w31492.pdf.

Chevalier, Arnaud, Colm Harmon, Ian Walker, and Yu Zhu. 2004. "Does Education Raise Productivity, or Just Reflect It?" *The Economic Journal* 114, no. 499 (November): F499–F517. https://doi.org/10.1111/j.1468-0297.2004.00256.x. https://doi.org/10.1111/j.1468-0297.2004.00256.x.

Currie, Janet, and Duncan Thomas. 2001. "Early Test Scores, School Quality and SES: Longrun Effects on Wage and Employment Outcomes." In *Worker Wellbeing in a Changing Labor Market,* 20:103–132. Emerald Group Publishing Limited, December. https://doi.org/10.1016/S0147-9121(01)20039-9. https://www.emerald.com/insight/content/doi/10.1016/s0147-9121(01)20039-9/full/html.

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. https://www.jstor.org/stable/23799087.

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. https://www.jstor.org/stable/4132484.

Ducking, Johnny, Peter A. Groothuis, and James Richard Hill. 2015. "Exit Discrimination in the NFL: A Duration Analysis of Career Length." *The Review of Black Political Economy* 42, no. 3 (January): 285–299. https://doi.org/10.1007/s12114-014-9207-9. https://doi.org/10.1007/s12114-014-9207-9.

———. 2017. "Compensation Discrimination: An Analysis of Linebackers, Defensive Linemen, and Defensive Backs in the National Football League." *The Review of Black Political Economy* 44, nos. 3-4 (January): 251–260. https://doi.org/10.1007/s12114-017-9256-y. https://doi.org/10.1007/s12114-017-9256-y.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *The Quarterly Journal of Economics* 111 (4): 1007–1047. https://doi.org/10.2307/2946706. https://www.jstor.org/stable/2946706.

Foltice, Bryan, and Justin Markus. 2021. *Finding Value in NFL Draft Picks.* SSRN Scholarly Paper. Rochester, NY, November. https://doi.org/10.2139/ssrn.3836012. https://papers.ssrn.com/abstract=3836012.

Ge, Suqin, Elliott Isaac, and Amalia Miller. 2022. "Elite Schools and Opting In: Effects of College Selectivity on Career and Family Outcomes." *Journal of Labor Economics* 40, no. S1 (April): S383–S427. https://doi.org/10.1086/717931. https://www.journals.uchicago.edu/doi/abs/10.1086/717931.

Ge, Suqin, Andrea Moro, and Beibei Zhu. 2021. "Testing for Asymmetric Employer Learning and Statistical Discrimination." *Applied Economics* 53, no. 12 (March): 1361–1377. https://doi.org/10.1080/00036846.2020.1830939. https://doi.org/10.1080/00036846.2020.1830939.

Gregory-Smith, Ian, Alex Bryson, and Rafael Gomez. 2023. "Discrimination in a Rank Order Contest: Evidence from the NFL Draft." *J Econ Race Policy* 6, no. 3 (September): 215–231.

https://doi.org/10.1007/s41996-023-00118-y. https://doi.org/10.1007/s41996-023-00118-y.

Harmon, Colm, and Hessel Oosterbeek. 2000. "The Returns to Education: A Review of Evidence, Issues and Deficiencies in the Literature." *CEE Discussion Papers* (December). https://ideas.repec.org//p/cep/ceedps/0005.html.

Hedlund, David P. 2018. "Performance of Future Elite Players at the National Football League Scouting Combine." *The Journal of Strength & Conditioning Research* 32, no. 11 (November): 3112. https://doi.org/10.1519/JSC.0000000000002252. https://journals.lww.com/nsca-jscr/fulltext/2018/11000/Performance_of_Future_Elite_Players_at_the.14.aspx.

Hendricks, Wallace, Lawrence DeBrock, and Roger Koenker. 2003. "Uncertainty, Hiring, and Subsequent Performance: The Nfl Draft." *Journal of Labor Economics* 21 (4): 857–886. https://doi.org/10.1086/377025. https://doi.org/10.1086/377025.

Huntington-Klein, Nick. 2021. "Human Capital Versus Signaling is Empirically Unresolvable." *Empir Econ* 60, no. 5 (May): 2499–2531. https://doi.org/10.1007/s00181-020-01837-z. https://doi.org/10.1007/s00181-020-01837-z.

Jackson, Charlton. 2024. "NFL Positional Value In Terms of Salary Cap Spending: Why Are Running Backs Undervalued?" *CMC Senior Theses* (January). https://scholarship.claremont.edu/cmc_theses/3640.

Keefer, Quinn A. W. 2016. "Rank-based Groupings and Decision Making: A Regression Discontinuity Analysis of the NFL Draft Rounds and Rookie Compensation." *Journal of Sports Economics* 17, no. 7 (October): 748–762. https://doi.org/10.1177/1527002514541448. https://doi.org/10.1177/1527002514541448.

———. 2017. "The Sunk-Cost Fallacy in the National Football League: Salary Cap Value and Playing Time." *Journal of Sports Economics* 18, no. 3 (April): 282–297. https://doi.org/10.1177/1527002515574515. https://doi.org/10.1177/1527002515574515.

Kitchens, Carl T. 2015. "Are Winners Promoted Too Often? Evidence from the Nfl Draft 1999–2012." *Economic Inquiry* 53 (2): 1317–1330. https://doi.org/10.1111/ecin.12165. https://onlinelibrary.wiley.com/doi/abs/10.1111/ecin.12165.

Loury, Linda Datcher, and David Garman. 1995. "College Selectivity and Earnings." *Journal of Labor Economics* 13 (2): 289–308. https://www.jstor.org/stable/2535105.

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, no. 7 (July): 1479–1495. https://doi.org/10.1287/mnsc.1120.1657. https://pubsonline.informs.org/doi/10.1287/mnsc.1120.1657.

Mountjoy, Jack, and Brent R. Hickman. 2021. *The Returns to College(s): Relative Value-Added and Match Effects in Higher Education.* Working Paper, September. https://doi.org/10.3386/w29276. https://www.nber.org/papers/w29276.

Pinkston, Joshua C. 2009. "A Model of Asymmetric Employer Learning with Testable Implications." *The Review of Economic Studies* 76, no. 1 (January): 367–394. https://doi.org/10.1111/j.1467-937X.2008.00507.x. https://doi.org/10.1111/j.1467-937X.2008.00507.x.

Pitts, Joshua D., and Brent Evans. 2019. "Drafting for Success: How Good Are NFL Teams at Identifying Future Productivity at Offensive-Skill Positions in the Draft?" *The American*

Economist 64, no. 1 (March): 102–122. https://doi.org/10.1177/0569434518812678. https://doi.org/10.1177/0569434518812678.

Riess, Steven A. 1980. "Sport and the American Dream." *Journal of Social History* 14 (2): 295–303. https://www.jstor.org/stable/3786796.

Riley, John G. 1979. "Testing the Educational Screening Hypothesis." *Journal of Political Economy* 87, no. 5, Part 2 (October): S227–S252. https://doi.org/10.1086/260830. https://www.journals.uchicago.edu/doi/abs/10.1086/260830.

Rose, Heather. 2006. "Do Gains in Test Scores Explain Labor Market Outcomes?" *Economics of Education Review,* School-to-Work and Educational Reform Symposium, 25, no. 4 (August): 430–446. https://doi.org/10.1016/j.econedurev.2005.07.005. https://www.sciencedirect.com/science/article/pii/S0272775706000264.

Schönberg, Uta. 2007. "Testing for Asymmetric Employer Learning." *Journal of Labor Economics* 25, no. 4 (October): 651–691. https://doi.org/10.1086/522905. https://www.journals.uchicago.edu/doi/abs/10.1086/522905.

Spearman, Lequez, Dawn Norwood, and Steven Waller. 2016. "DOES THE CREAM RISE TO THE TOP? NCAA DIVISION I PLAYERS' PERCEPTIONS OF MERITOCRACY IN BASEBALL." *Journal of contemporary athletics* 10 (2): 119–119.

Spence, Michael. 1981. "Signaling, Screening, and Information." In *Studies in Labor Markets,* 319–358. University of Chicago Press. https://www.nber.org/books-and-chapters/studies-labor-markets/signaling-screening-and-information.

Tolkin, Alex. 2024. "Field of (American) Dreams: Sports and Belief in Meritocracy."

Wapman, K. Hunter, Sam Zhang, Aaron Clauset, and Daniel B. Larremore. 2022. "Quantifying Hierarchy and Dynamics in US Faculty Hiring and Retention." *Nature* 610, no. 7930 (October): 120–127. https://doi.org/10.1038/s41586-022-05222-x. https://www.nature.com/articles/s41586-022-05222-x.

Weiss, Andrew. 1995. "Human Capital Vs. Signalling Explanations of Wages." *Journal of Economic Perspectives* 9, no. 4 (December): 133–154. https://doi.org/10.1257/jep.9.4.133. https://www.aeaweb.org/articles?id=10.1257/jep.9.4.133.

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

A Additional Data Details (Online)

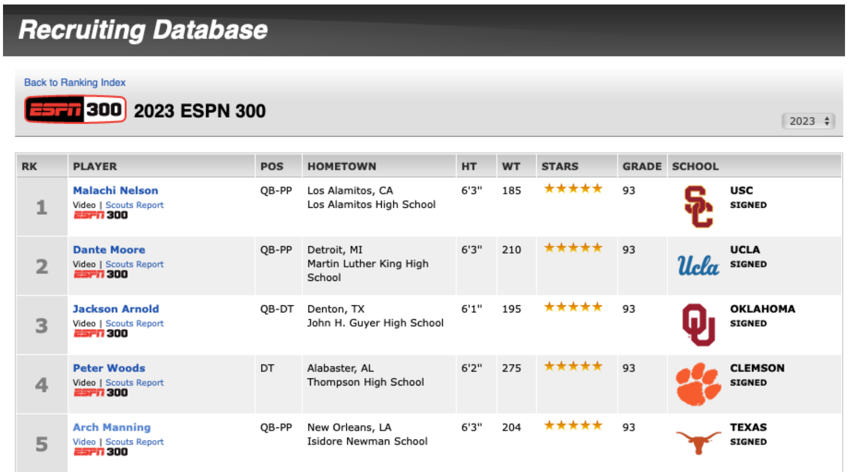
A.1 ESPN 300

The ESPN 300 College Football recruiting database represents one of the most comprehensive and rigorously compiled evaluations of high school football talent in the country. This database, spearheaded by ESPN's national recruiting director Tom Luginbill, is built upon a foundation of meticulous game film analysis. The evaluation process begins with breaking down several games into a "hit tape," which features a balanced collection of an athlete's exceptional plays, mistakes, and average performances [A.1]. This method ensures scouts get a complete picture of a player's capabilities, avoiding a bias toward only highlight-worthy moments. Scouts use these hit tapes to grade recruits according to a detailed and structured grading scale that spans from "Rare prospects" (90-100), who possess game-changing skills and immediate college-level impact, to "Solid prospects" (60-69), whose strengths may be outmatched against elite opponents but can contribute meaningfully at the non-BCS or FCS levels. Unranked players are left under review until their film evaluation is completed (AL.com, 2015).

What sets the ESPN 300 apart is the caliber of its talent evaluators. Luginbill emphasizes that only individuals with coaching experience at the collegiate or professional level are permitted to assess players. He strongly believes that breaking down film and accurately identifying talent requires a trained eye and an intimate understanding of the game, skills honed through years of professional experience. Luginbill himself brings credibility to the process, having played college football at three schools, including Georgia Tech, and coached professionally in the XFL, NFL Europe, and Arena Football League. This professional rigor ensures that the ESPN 300 reflects not only the athletic potential of recruits but also their projected ability to compete at the next level of football (AL.com, 2015).

A.1. AL.com. (2015). *How are recruiting rankings determined? An inside look at the process.* Retrieved from https://www.al.com/sports/2015/02/how$_a$re$_r$ecruiting$_r$ankings$_d$e.html.

Fig. 12. ESPN 300 Website Info



Recruiting Database

Back to Ranking Index

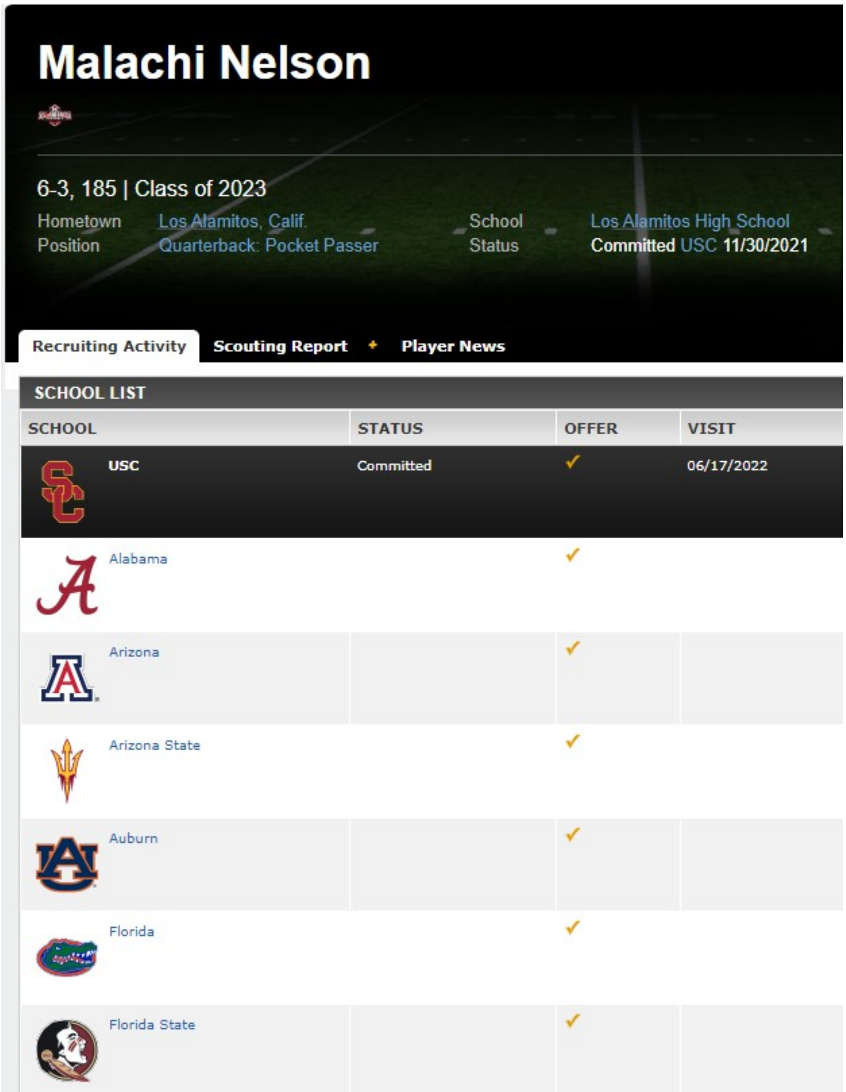
ESPN 300 2023 ESPN 300

2023

RK	PLAYER	POS	HOMETOWN	HT	WT	STARS	GRADE	SCHOOL
1	Malachi Nelson Video | Scouts Report ESPN 300	QB-PP	Los Alamitos, CA Los Alamitos High School	6'3"	185	★★★★★	93	USC SIGNED
2	Dante Moore Video | Scouts Report ESPN 300	QB-PP	Detroit, MI Martin Luther King High School	6'3"	210	★★★★★	93	UCLA SIGNED
3	Jackson Arnold Video | Scouts Report ESPN 300	QB-DT	Denton, TX John H. Guyer High School	6'1"	195	★★★★★	93	OKLAHOMA SIGNED
4	Peter Woods Video | Scouts Report ESPN 300	DT	Alabaster, AL Thompson High School	6'2"	275	★★★★★	93	CLEMSON SIGNED
5	Arch Manning Video | Scouts Report ESPN 300	QB-PP	New Orleans, LA Isidore Newman School	6'3"	204	★★★★★	93	TEXAS SIGNED

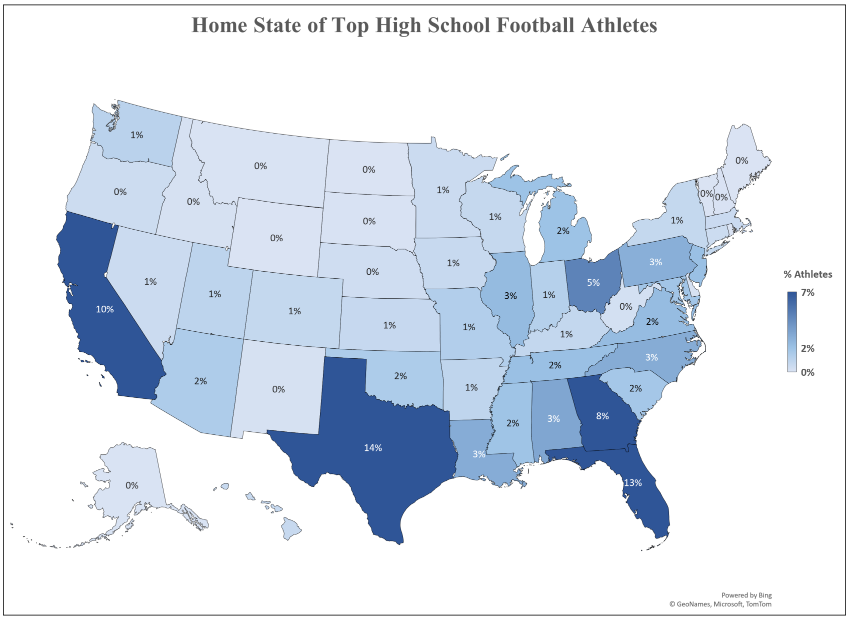
Note: This figure shows an example of the ESPN 300 High School Football Recruiting Rankings.

Fig. 13. ESPN 300 Scholarship Info



Note: This figure shows an example of the ESPN 300 High School Football Recruiting Rankings. For each ranked high school athlete the each college football scholarship offer is recorded as well as the dates of official visits.

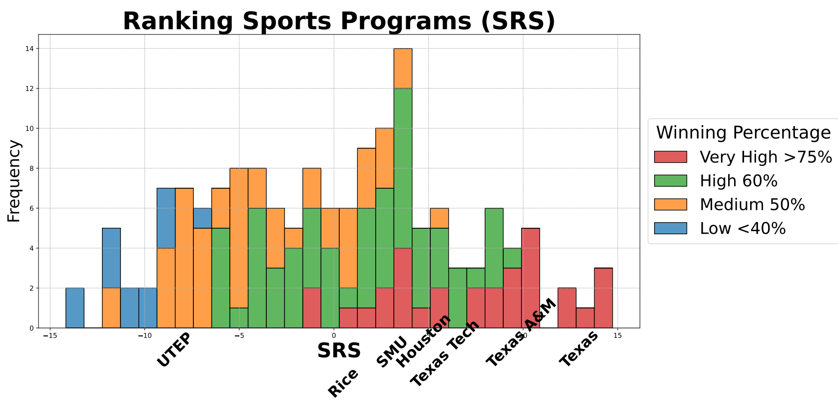
Fig. 14. ESPN 300 Recruit State Distribution



Note: This map shows the percentage of ESPN 300 Top Recruits from each US State.

A.2 Measuring Elite Sports Programs

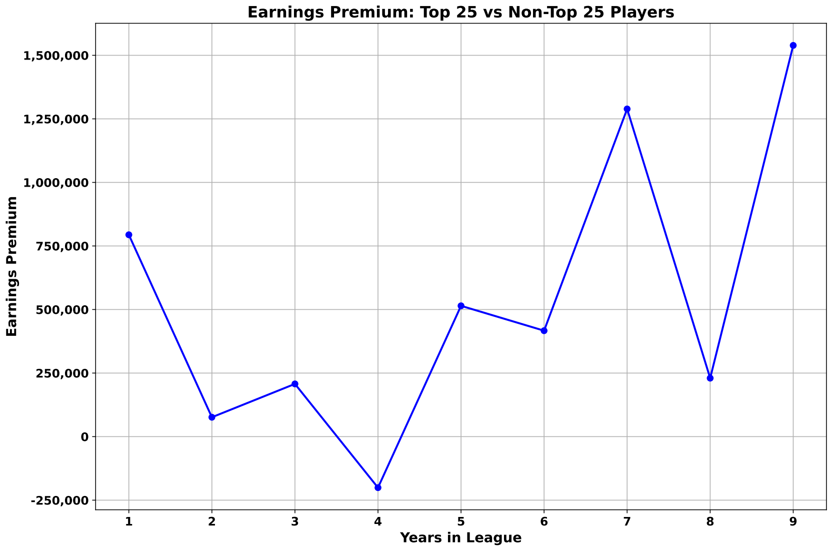
Fig. 15. College Sports Quality (SRS) & Winning Percentage



Note: This shows the correlation between how sports quality measure simple rating system (SRS) relates to other well-known measures of sports quality, winning percentage.

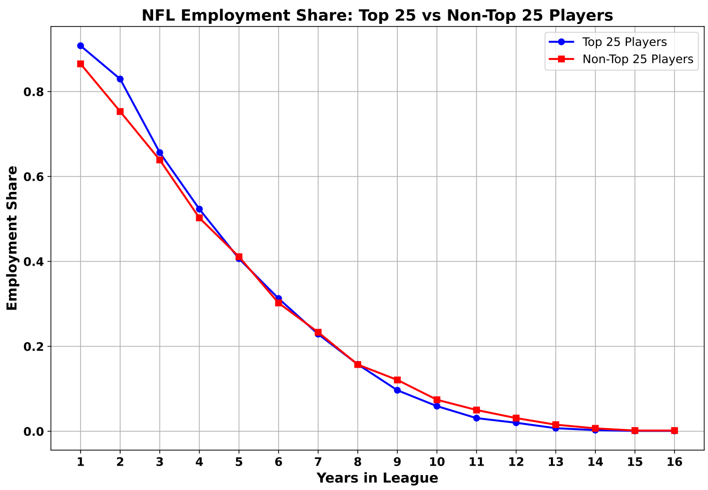
A.3 NFL Athletes Career Descriptive Statistics

Fig. 16. Earnings Premium: Top 25 vs Non-Top 25 Players (Non-Balanced Panel)



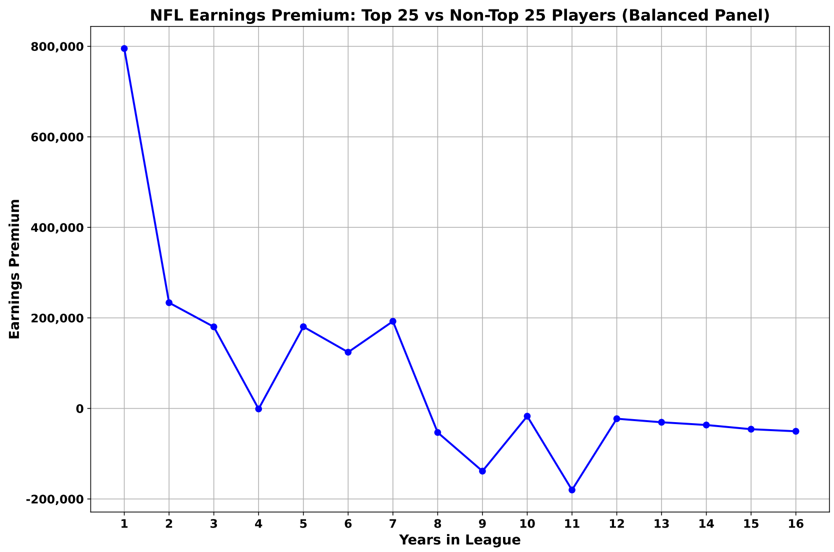
Note: This figure shows the earnings premium for players from Top 25 colleges compared to Non-Top 25 colleges across years in the NFL, using a non-balanced panel.

Fig. 17. NFL Employment Share: Top 25 vs Non-Top 25 Players



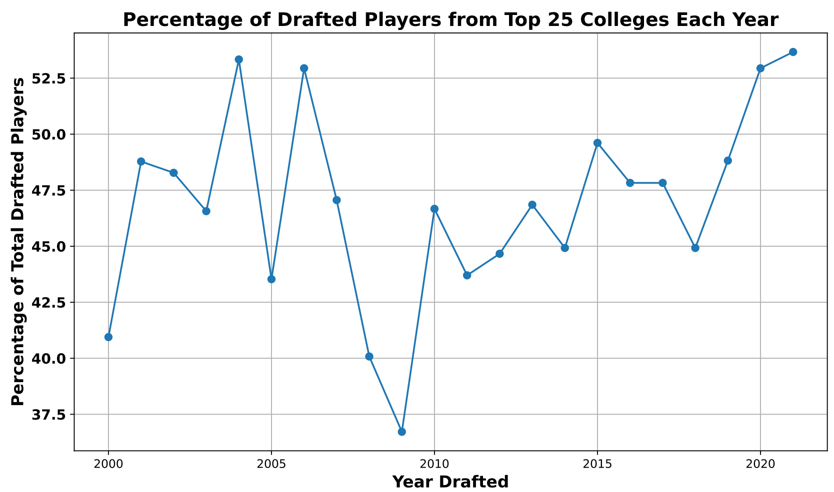
Note: This figure displays the employment share in the NFL for players from Top 25 colleges compared to Non-Top 25 colleges across years in the league.

Fig. 18. NFL Earnings Premium: Top 25 vs Non-Top 25 Players (Balanced Panel)



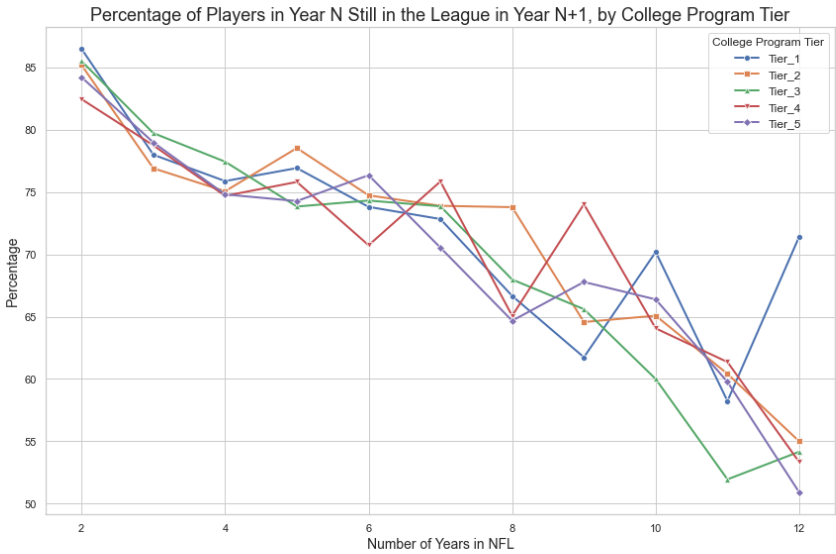
Note: This figure presents the earnings premium for players from Top 25 colleges compared to Non-Top 25 colleges across years in the NFL, using a balanced panel.

Fig. 19. Percentage of Drafted Players from Top 25 Colleges Each Year



Note: This figure shows the percentage of players drafted from Top 25 colleges each year, illustrating trends over time.

Fig. 20. Percentage of Players in Year N Still in the League in Year N+1, by College Program Tier



Note: This figure highlights the retention rates of players in the NFL from one year to the next, broken down by college program tier, providing insights into player longevity by program quality.

B Additional Method Details (Online)

B.1 College Choice in the Matched Scholarship Sample

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:

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

Where:

- ϵ_{ijg} is the error term capturing unobserved factors that affect the outcome y_{ijg}.
- $\mathrm{Group}_{ig} = 1$ 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:

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

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.

B.2 Matching Groups Details & Robustness

The matching group generation methods use different strategies to assign unique IDs to athletes based on the colleges that offered them scholarships. These methods focus on categorizing colleges into quality tiers (e.g., quintiles, deciles, or terciles) and summarizing the offers each athlete received. The goal is to create standardized group identifiers that reflect the diversity and distribution of scholarship offers while maintaining a manageable structure for analysis.

MatchingModel_v1 and *MatchingModel_v4* take a binary approach, where each digit in the ID indicates whether an athlete received an offer from a college within a specific quality tier. For *MatchingModel_v1*, this is done for five quintiles, creating a 5-digit binary ID. *MatchingModel_v4* expands this concept to 10 deciles, resulting in a 10-digit binary ID. These methods prioritize simplicity but may lose some detail about the number of offers within each tier.

Other models, such as *MatchingModel_v2*, *MatchingModel_v3*, and *MatchingModel_v5*, use counts instead of binary indicators. *MatchingModel_v2* and *MatchingModel_v3* summarize the number of offers within each quintile, capping the counts at 5 or 9, respectively, to ensure consistent ID length. *MatchingModel_v5* operates similarly but focuses on three terciles instead of quintiles. These methods provide more granular information about the distribution of offers, allowing for a richer representation of the scholarship offer sets while introducing slightly more complexity in ID generation.

Table B.1: Tradeoff Between within Group Homogeneity & Power-Size

Matching Estimator Model	Group Count	Groups w Treatment Variation	Observations Count
MatchModel1	32	32	23021
MatchModel2	4,127	2,326	21,109
MatchModel3	6,,575	2,182	18,521
MatchModel4	1,011	966	22,963
MatchModel5	908	833	22,940
ExactMatching	20,500	177	422

Note: The table summarizes the performance of different matching estimator models by comparing their group count, the number of groups with treatment variation, and the total number of observations. The key tradeoff highlighted in this table is between within-group homogeneity (achieved by more precise matching) and statistical power (dependent on the size and number of groups).

Table B.2: Athlete Scholarship Offers and Match Group IDs

Athlete Name	Schools Offered (SRS)	School Tiers	ID v1	ID v2
Jacob Smith	ASU (3.7), Ohio St (9.4), LSU (8.7)	[4, 1, 2]	11010	11010
Mike Chen	Alabama (9.7), UH (5.4), UT (7.3), TT (5.7)	[1, 3, 2, 3]	11100	11200
Evan Hill	SMU (3.2), MIT (2.3), Duke (4.1)	[4, 5, 4]	00011	00021

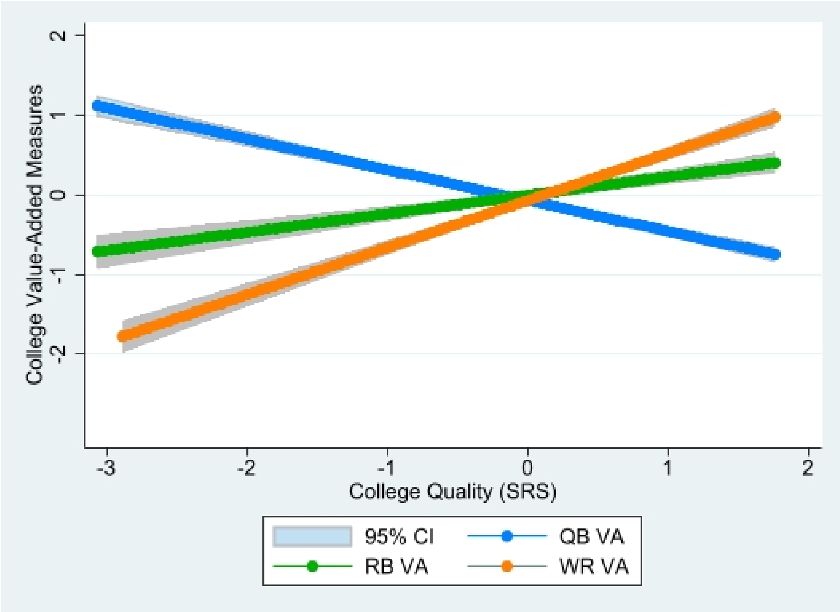
Note: The table demonstrates how scholarship offers received by athletes are transformed into matching group IDs based on the quality of the colleges (measured by SRS) and their respective tiers. Each row represents an athlete, listing the colleges that offered scholarships along with their SRS scores and the corresponding tiers. The final columns show how this information is encoded into different versions of matching group IDs. These group IDs serve as dummy variables in the regression model, where each unique ID corresponds to a group of athletes with similar offer sets. This encoding enables the model to account for treatment variation (differences in college choices) within these groups.

C Additional Estimation Results (Online)

C.1 Additional Results & Robustness

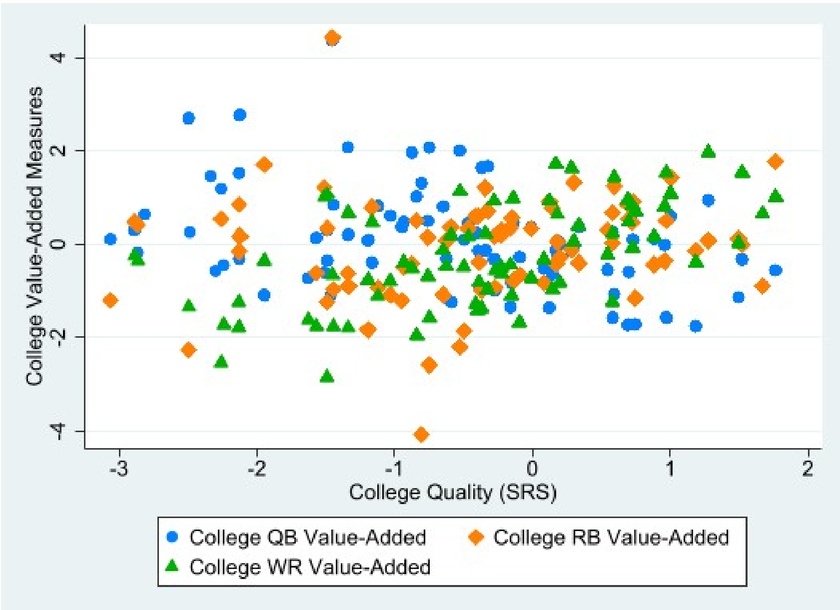
D Appendix Figures and Table (Online)

Fig. 21. College Value-Added Measures vs. College Quality (SRS)



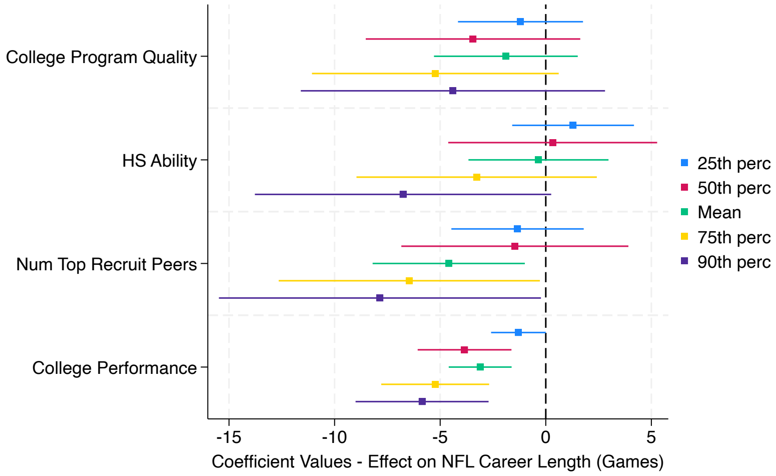
Note: This figure demonstrates the relationship between college quality (SRS) and value-added measures for different positions, including quarterbacks, running backs, and wide receivers, along with confidence intervals.

Fig. 22. Scatterplot of College Value-Added Measures by College Quality (SRS)



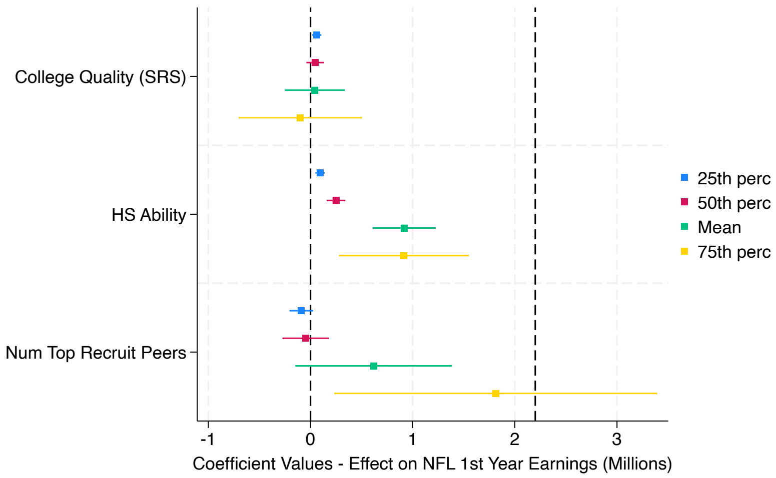
Note: This scatterplot illustrates the relationship between college quality (SRS) and value-added measures for quarterbacks, running backs, and wide receivers, highlighting the spread of data points across positions.

Fig. 23. Effect of College and Player Attributes on NFL Career Length



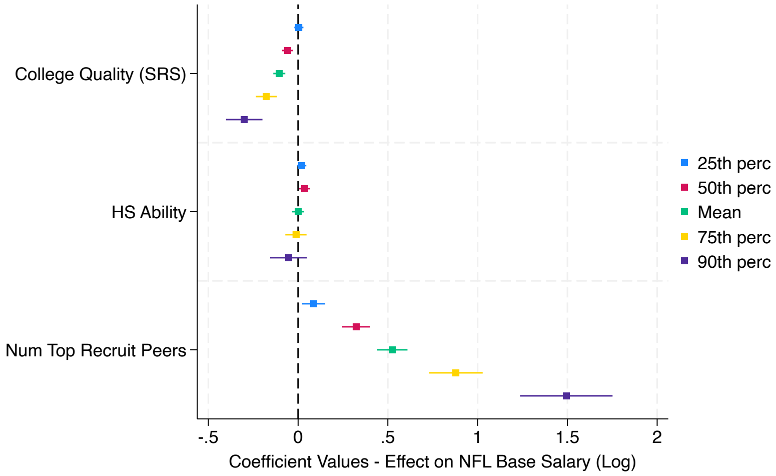
Note: This figure shows coefficient values from quantile regressions analyzing the effect of college quality, high school ability, peer recruits, and college performance on NFL career length. Percentiles reflect the distribution of effects.

Fig. 24. Effect of College and Player Attributes on NFL Rookie Earnings



Note: This figure presents coefficient values from quantile regressions examining how college quality, high school ability, and peer recruits affect first-year NFL earnings.

Fig. 25. Effect of College and Player Attributes on NFL Base Salary



Note: This figure displays coefficient values from quantile regressions showing the impact of college quality, high school ability, and peer recruits on NFL base salary. Percentiles illustrate distributional effects.