Leverage over Life Outcomes:
Measuring Agency among College Quarterbacks

Edward Lucas Hock

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Arts

University of Washington
2025

Committee:
Kyle Crowder
Tyler McCormick

Program Authorized to Offer Degree:
Sociology

University of Washington

Abstract

Leverage over Life Outcomes: Measuring Agency among College Quarterbacks

Edward Lucas Hock

Chair of the Supervisory Committee:
Kyle Crowder
Department of Sociology

The sociological literature has produced a litany of definitions for *agency*. These ideas are diverse, hotly debated, and have become increasingly elaborate as theorists have attempted to reconcile often incompatible concepts. This theoretical morass has made it difficult for scholars to conduct research related to agency, namely to investigate the impact of key life choices on key life outcomes. In this paper, I present a framework to untie this particular knot, establishing the concept of "leverage," a measure of how much difference a particular actor's choices make on an outcome in the social world. Using high quality data from college football, I conduct a study on the impact players' choice of college program has on their eventual probability of being drafted into the NFL. I show that leverage is distributed heterogeneously across the population of highly recruited quarterbacks, and I present strong evidence that the "middle class" of talented but non-elite players have the most leverage over their desired outcome in this case. I discuss the implications of these findings for sociological thinking about agency in general, and suggest social milieux where research into leverage may be relevant.

Leverage over Life Outcomes:
Measuring Agency among College Quarterbacks

Eddie Hock, Department of Sociology, University of Washington
Defended May 20, 2025

INTRODUCTION

We tell ourselves that the courses of our lives hinge on important decisions. In day-to-day experiences, we are constantly confronted with simple questions like what to eat for dinner, what to do over the weekend, and what route to take on our commute. These decisions may affect our futures in small, predictable, path dependent ways. But from time to time, we find ourselves at crossroads that feel much more consequential. Should we move to a new city or new country for a job or a relationship? Apply to graduate school or stay employed in the private sector? Take a risk on a business venture, an investment, or a passion project? Key life choices like these feel so impactful because we understand that choosing one option over another cuts off whatever future paths we might have had if we had gone the other way; that is, there are countless potential life courses behind door B that we will never experience if we open door A, and vice versa. These dilemmas dominate the narratives we tell about our lives (Ewick and Silbey 1995). We chart our personal growth in terms of what we consider when making consequential choices. We delineate the chapters of our life stories by these moments and the changes they entail. We believe they make extraordinary impacts on our lives.

It seems sociologically clear that people are exercising something called *agency* when they make important life choices like these. They have discretion between their options, and are differentially constrained regarding what options are plausible for them, and experience different outcomes depending on which outcome they choose. Yet it has proven difficult to measure the impacts the pivotal, definitive decisions people make actually have on their lives, largely because it has not at all seemed sociologically clear what exactly agency is. In the aptly titled *What Is Agency?* (1998), Emirbayer and Mische write that "the term *agency* itself has maintained an elusive, albeit resonant, vagueness; it has all too seldom inspired systematic analysis, despite the long list of terms with which it has been associated" (962). When we speak of agency, are we talking about an abstract sense of self-actualization or empowerment? Are we talking about how free a person is from structural constraints in the resources and experiences to which they have access? Or are we talking about whether the choices they have actually matter to their life outcomes or to the world? We struggle to investigate the role of agency in important life choices, and more specifically to measure how important choices actually are, because we have not actually agreed on what concept we are trying to quantify.

In this paper, I present a framework for measuring the importance of specific life decisions for specific life outcomes, and in order to do this, I adopt a particular

definition of agency, which I call the "difference-making" or "leverage" perspective. Giddens (1984) writes that "Action depends upon the capability of the individual to 'make a difference' to a pre-existing state of affairs or course of events. An agent ceases to be such if he or she loses the capability to 'make a difference,' that is, to exercise some sort of power" (14). To measure agency, then, is to measure the difference someone's actions make in the world; to measure the impact of a key life decision on some outcome is to measure how much variation is induced in that outcome depending on what choices an actor makes. I introduce the novel concept of "leverage," defined as the degree of variation induced in an outcome by an actor's discrete choices, to operationalize this idea. The goal of this project is to define leverage, situate it in the conversation about agency, and conduct an empirical study exploring how it manifests in a specific social context.

The case I focus on is the college commitment decision for high school football players, particularly those who are highly regarded by scouts and heavily recruited by division 1 coaches. This is a major life decision that closely mirrors one many Americans make in their own lives (where to go to college), but may promise truly transformative returns in social and financial capital for players, their families, and potentially their entire communities if the player is drafted into the National Football League (NFL) at the end of their college career. While these recruits must consider many of the same factors their non-sports playing peers do when deciding where to go to school, the allure of an NFL contract is impossible to ignore. Even if making it to the league is not their top priority — which, for most recruits, it is — the possibility that they will end up "playing on Sundays" is at the front of their minds as they decide which scholarship offer to accept (Crabtree 2022). The pressurized, dramatic environment of college football exhibits numerous characteristics that make it appealing for sociological study. It is a crucible where talented young people compete for status, money, and fame that only a few of them will enjoy for their entire lives. The standards of this competition are established and altered by institutions like the National Collegiate Athletic Association (NCAA), college athletic conferences, and university administrations. Players and their families must navigate these changing rules of engagement or risk receiving nothing for their enormous investments of time, energy, and money into a game that is likelier to result in permanent health problems than fame and fortune. Many of us make decisions that resemble this one — where to go to college is a question millions of Americans face every year — but rarely are the stakes so clear, the agents so aware of their options, the rewards so obvious, the risks so dangerous. This is an ideal context to explore how leverage is distributed across the lives of a rarified stratum of young people.

The data concerning this process is promising because it satisfies theoretical conditions that most sociological data does not, allowing us to examine the effects of discrete, intentional life choices on discrete, measurable life outcomes. Since recruiting information is publicly available, we know what schools gave which recruits scholarship offers each year, along with where they chose to go, what happened during their playing

career in college, and whether they eventually made it to the NFL. Division 1 football programs have limited recruiting resources and can offer only a limited number of scholarships each year, so if a school extends an offer to a player, it means that they believe he has a role to play on their team and that they think they have a shot at landing him. This means that I know what choices (scholarship offers) were available to each individual, what option (school) they chose, and whether they got their desired outcome (being drafted) in the factual case. There is also high quality data regarding how players were rated by scouts when they were in high school, their statistical performance in college football, and similar status/performance statistics for the schools themselves. Since I have information of this granularity for each individual and each school, I can make plausible claims about the leverage associated with the college decision across the population of recruited quarterbacks.

In my analysis, I first identify a model for draft probability. This model takes into account player-side statistics including race, recruit ranking out of high school, and in-game performance metrics like passing yards and touchdowns. It also involves school-side statistics including Colley matrix rating (a statistical method used to measure a team's performance in the context of their competition, which can be highly asymmetrical in college football), whether the school is in a "power" conference, and offensive scheme. Both player- and school-side statistics impact the outcome considerably. By keeping the player-side statistics the same for each player, but changing the school to each one that gave the player an offer, I use this model to predict draft probability for each plausible school choice for each player. I then graph and analyze the outputs for each plausible counterfactual. I represent leverage as the interquartile range of modeled draft probabilities for each player, and I examine the relationship between leverage and relevant player- and school-side predictors.

My findings make several contributions to the literature. First, my methods for measuring leverage are novel, and I identify some intriguing patterns in how this measure of agency is distributed and stratified across the population of top-50 high school quarterback recruits from 2011 to 2014. I show that, while there is a loose positive relationship between the raw number of options a player has and how much leverage they have over the probability that they will be drafted, this association is not as linear or as straightforward as might be expected; having more options does not necessarily mean having more leverage. Even with similar choices and similar numbers of choices among schools, different players have different agency over this key life outcome. I find that leverage does not demonstrate a linear positive relationship with status or performance in this case. In fact, it is the "middle class" of semi-talented, semi-high-status players who seem to have the most leverage, suggesting that the leverage associated with this particular decision is highest for players whose status and ability put them on the edge of draftability. It seems more important to go to the best program for players who won't stand out from the crowd than it is for players below the statistical cutoff of draftability or for elite players who would succeed anywhere. I also

make some substantive comments regarding the NFL draft literature. I conclude by discussing the limitations of my analytical methods, and I touch on what future research could help clarify and further explore involving leverage.

THEORY

It is remarkably difficult to articulate a concise, coherent definition of agency that accounts for all, or even most, of what sociologists mean when they use the term. Emirbayer and Mische (1998) arrive at "the temporally constructed engagement by actors of different structural environments — the temporal-relational contexts of action — which, through the interplay of habit, imagination, and judgment, both reproduces and transforms those structures in interactive response to the problems posed by changing historical situations" (970). This labored, elaborate description is the rigorous product of an honest attempt to synthesize decades of sociological thinking and writing about agency, ideas that are so divergent and flexible that they cannot be woven together into a succinct or methodologically applicable definition. Agency has often been defined and redefined in the pursuit of making arguments in the so-called "structure and agency debate," which Fuchs (2001) identifies as a "bad essentialist habit" of assuming that agency and structure are opposites, or that one is actually the other in disguise, or that they can simply be parsed. Indeed, many authors including Sewell (1992), Emirbayer and Goodwin (1994), Archer (2003), and Giddens (1976 and 1984) propose that agency *is* structured, and that understanding *how* it is structured in a particular context should lend insight into understanding the social forces that shape that structure and the effects it has on individual- and group-level life outcomes.

I am far from the first to propose a way to measure agency, and I am indebted to authors like Jejeebhoy et al. (2010) and in particular Franzosi et al. (2012) who have made similar stabs at this persistent problem. A common thread across these projects is that they try to move past rational choice or game-theoretic analysis of choices and strategies to develop a more holistic, though still rigorous, approach to measuring agency. Jejeebhoy et al. (2010) employ a mixed methods approach with detailed survey questions designed to "draw out and measure different dimensions of agency among unmarried young women and men in India" (63), engaging with the tangled landscape of sociological thinking about agency by not articulating a specific definition themselves, but allowing respondents to express their own values and measuring how strongly they hold them and how fulfilled they feel about them. Franzosi et al. (2012) use quantitative narrative analysis to investigate newspaper narratives about lynchings in Georgia between 1875 and 1930, conducting a discourse analysis of these stories to identify specific actors involved, investigate who had agency (under their definition) and who was constrained, and comment on meta-information like the spatial distribution of lynchings and their causes and consequences.

The Jejeebhoy et al. (2010) project and similar surveys are effective means of getting at what Alkire (2005) calls "subjective quantitative measures of human agency at the individual level" (217). However, there is a serious paucity of research on how agency varies between decisions and individuals over specific outcomes in the social world, most notably forms of capital. Yet the foundational literature concerning how actors compete over forms of capital, in particular Bourdieu (1972, 1984, 1993), has clear expectations about how agency *should* vary between individuals (Harvey et al., 2020). Bourdieu and the lush tradition following him posit that agency should be associated strongly with status, with the most privileged actors having the most agency as their strategies dominate those of their peers. Having the most options, or having the largest proportion of one's options be "good" with respect to whatever form of capital a Bourdieusian agent is competing over, should arise from having the most capital to begin with. This process should create a centralizing spiral where those with the most capital have the easiest time getting more capital. Over time, these privileged individuals should entrench themselves in a permanent or semi-permanent state of high agency. A significant incentive for the "measuring agency" literature is to investigate empirically whether these expectations play out in the observed social world.

To that end, I want to build on this burgeoning literature by developing an objective measure of agency that can be estimated for a specific actor in a specific situation. That is, I want to be able to say that someone has a measurable degree of *leverage* over a specific outcome given what choices they have that will influence that outcome. This implies a particular definition of agency, one that draws heavily on Giddens (1976, 1984). Giddens reframes the structure and agency discussion, grappling with questions of action and intention. For Giddens (1984), agency is composed of action (what options are available to an actor) and power (what those actions change in the world), and structure is composed of rules and resources. Agents act with "reflexivity," or "the monitored character of the ongoing flow of social life" (3), when they consciously take what actions they can and expend what resources they have within the rules to affect their social position. This process may have intended and unintended consequences for the social structure, which determines what rules and resources govern the agent the next time they act. To simplify and distinguish this language a bit, my definition of *agency* is *the exercise of actions that make a difference in the world*.

This definition, the difference-making perspective on agency, is incomplete in that it does not cover things like the subjective experience of empowerment or how free someone is from structural constraints in life. Nevertheless, this definition is useful to this study in that it lays out clear terms for what quantities surrounding agency (namely leverage) I can measure and what information I need in order to measure them. Leverage is a contextual measurement. It is a function of the characteristics of an actor, what actions are available to them, and the outcome in the world that a difference is being made in. It is not feasible to ask "how much agency does so-and-so have?" in absolute terms, but it is feasible to ask "how much leverage does so-and-so have over

such-and-such an outcome when making this-or-that decision?" In the case of this study, I now have the methodological framework I need to ask: "How much leverage do recruited high school quarterbacks have over their eventual chance of being drafted into the NFL when choosing which scholarship offer to accept?"

DATA AND METHODS

Case background

The process by which a football player ends up in the NFL begins in childhood. Either out of their own interest or because of the prodding of parents, relatives and coaches, boys get involved in youth football, usually transitioning to full tackle football by the end of middle school. This is the point at which injury concerns, a significant social issue and one for which the institution that might be called the football-industrial complex has provided few satisfactory solutions, intensify (Navia 2012, McCarthy 2015). As early as 8th grade, scouts from football-playing universities and professional firms start visiting practices and games, observing and interviewing players and coaches to evaluate their performance and judge their physical characteristics, readiness, and potential for the next levels of the sport. If their son continues playing in high school, parents shell out serious money in an effort to keep him competitive. In 2019, one Idaho mother reported a cost of over $1000 per player, per season, covering a school athletic fee, a player registration fee, and a summer camp (Flake 2019). This, of course, was for a public school in a moderately low cost-of-living state, and fees for "feeder" programs, private schools, and Catholic schools that promise more attention from division 1 colleges can be many times higher. One football feeder Catholic school in California charges a yearly tuition of $20,150 for Catholic students and $21,675 for non-Catholics, and this is not counting additional football fees[1]. Of course, this cost is paltry compared to the return of an NFL contract, but the NCAA estimates only 1.5% of college players end up in the NFL[2], so parents who can't afford costs like these must hope that their son is both extremely talented and very lucky.

Throughout high school, players (particularly at these higher-end programs) are visited frequently by scouts, and eventually those of them who are perceived to show promise for the next level are given ratings by the three main scouting agencies: ESPN, Rivals.com, and 247Sports.com. These ratings, binned from 1 to 5 stars, have some grounding in players' season stats, but high school football schedules are totally asymmetrical and do not represent a good basis for comparison. Instead, recruit ratings are mostly subjective, determined largely by how physically fit a scout thinks a player is, or whether they perceive him to have "intangible" characteristics that won't show up in the stats, or whether he happens to play well in the one game out of the season that the scout is present for. This means that great players are frequently overlooked and given

[1] https://www.materdei.org/apps/pages/index.jsp?uREC_ID=70022&type=d&pREC_ID=1798540
[2] https://www.ncaa.org/sports/2015/3/6/estimated-probability-of-competing-in-professional-athletics.aspx

1-, 2-, or 3-star ratings, whereas the players with 4- and 5-star ratings — otherwise known as "blue chip" recruits — are usually outstanding performers in high school, but only at high schools that are lucky to get sufficient media attention. There are also strong herding effects in scouting, where a player being highly rated by one firm means that the other firms will give him more attention. Players who fly under the radar may have to turn to social media to get coverage or hope that their coach knows a scout.

College scouts, who have the power to recommend to a coaching staff that a player be given a scholarship, habitually visit high schools in the vicinity of the university to find promising players, but for national recruiting, they rely on professional scout ratings to know what games to visit during their packed schedules. Players are nationally ranked according to their position, so there will be, for example, a top 50 list of quarterbacks across the United States who receive the bulk of attention from division 1 (D1) universities, those that can offer full scholarships for football, each year. Players who impress these scouts and the coaches they represent are extended scholarship offers, one of which they choose to commit to by National Signing Day (usually in early February). Especially for heavily recruited "prospects" — called as such because they are perceived to have a good chance of excelling in college or the NFL eventually — this is a highly ritualized process, with the commitment decision often being made in public in front of the player's teammates, family, and community.

Recruited high school football players are some of the most socially privileged young people in American society, and upon entering a D1 football program, they surrender most of their freedom to their coaches in exchange for the implication that those coaches will do their best to prepare them for the NFL (Logan et al. 2015). They adjust to heavily regimented practice and travel schedules, become local and national celebrities if they succeed on the field, and in recent years see details from their personal lives broadcast on social media. All the while, they compete with their teammates for starting roles, their opponents for wins, and themselves to improve at the game enough that they begin to get noticed by NFL teams. Players who become the elect — that 1.5% who will end up playing in the pros — must excel at the game, get media attention, perform well in interviews with teams, succeed at the NFL combine (an invite-only group workout in front of the media and coaching staffs), and get lucky enough that a team has a weakness or an opening at the position they play. If all those conditions are met, then a player may hear his name selected during the NFL draft, a 3 day process by which teams compete against and bargain with one another for the pick of the litter from that year's college players. The NFL, unlike other American professional sports leagues, requires players to be out of high school for at least 3 years to be eligible for the draft, so most players drafted will have played in college for between 3 and 5 years.

I look at only quarterbacks for this study because of certain unique characteristics the quarterback position has. First, each team only needs one starting quarterback, which results in quarterback recruits receiving a relatively smaller number of scholarship offers compared to other positions, and puts the quarterbacks in much more

direct statistical competition with one another than are players at other positions. Second, quarterback is a high-status, strategically critical, cerebral position, the focal point of the offense, the priority target for the defense, and popularly called "the most important position in sports" (Feinstein 2018). These players, more than any others in the sport, are constantly evaluated and compared to one another by scouts, coaches, fans, and analysts, so the data on them is of particularly high quality. Third, because of the importance of the quarterback position to the game of football and the increased status it entails, it has been uniquely racialized. Between 1960 and 2020, the NFL has evolved from a mostly white league to one dominated by Black players at most positions, but a racial hierarchy where white quarterbacks are favored over Black quarterbacks persists. This manifests in several ways, including Black QBs being benched more often (Volz 2015) and discriminated against in media coverage. Despite much discussion about the so-called "rise of the Black quarterback" (Lindsey 1973, Early 2011), quarterbacks at the college and professional levels remain majority white (Marquez-Velarde et al. 2023), though QB is no longer exactly the marquee position reserved for a white player it once was. It has been argued that this disparity persists even as the NFL in general has become overwhelmingly Black because the game revolves around protecting the quarterback, and as early as childhood white players are funneled into the safer, more cerebral quarterback position while Black players are encouraged to play more physically demanding, injury-prone positions like wide receiver, running back, and defensive back. Even those Black players who overcome this racial bias to succeed at the quarterback position are often evaluated based on physical characteristics, as opposed to their white peers, who are usually given more holistic evaluations involving their mental traits (Bigler and Jeffries 2008).

The literature shows, and it must be noted here, that NFL teams are not actually particularly good at evaluating quarterbacks for the draft, at least in the sense that there is only a weak relationship between how teams rate QBs on draft day and how they ultimately perform in the league (Hendricks et al. 2003, Berri and Simmons 2011). Despite being an incredibly moneyed process with high stakes for team success and profit, the draft market is demonstrably inefficient, especially regarding quarterbacks. The performance of quarterbacks after the draft is well outside the scope of this paper, but just because the behavior of the market is inefficient, that does not mean it is random, and I will show in this paper that it is still quite feasible to model draft probability, even if exact draft position is murkier. While it is possible for an undrafted quarterback to succeed in the NFL, including big names like Tony Romo and two Hall of Famers in Warren Moon and Kurt Warner, what is much more likely and much more relevant according to the literature on this subject is for a quarterback who *is* drafted, but in a later round than other quarterbacks, to outperform the players drafted ahead of him. Examples of this include Super Bowl champions like Roger Staubach, Russell Wilson, and Tom Brady. It is partially for this reason that the binary operationalization of drafted versus undrafted works well for my outcome variable — getting drafted at all

gives a quarterback a foot in the door, the opportunity to prove himself in the league, and a lucrative rookie contract even if he does not end up performing. Being drafted is its own reward, *and* it opens the door to even greater rewards down the line. The status, money, and opportunity for more presented here are the reasons why so many sacrifice so much to try to get their son into that lucky 1.5%.

Player-side data

I collected data on the top 50 quarterback recruits out of American high schools each year from 2011 to 2014, a population comprising 200 total individuals. The player-side data I collected on these 200 recruits includes several key variables. The players' recruit class (2011 to 2014), race, home state, recruit rating (3 to 5 stars), which schools gave them scholarship offers, and where they initially committed were collected from 247Sports.com, one of the three major high school sports scouting firms along with ESPN and Rivals.com. I collected data on the players' statistical performance in college football games, and what schools they transferred to if they transferred, from sports-reference.com. This data includes, for the players' final full season in college or the closest to it, passing yards, rushing/receiving yards (coded as 0 if negative due to sacks), total touchdowns, interceptions, and completion percentage. It may be that players are evaluated by NFL teams based on seasons prior to their final college year, but using the final full season puts the players on an even statistical playing field for comparison and represents the most recent and likely the most relevant stats in general for the draft. I collected data on whether, and at what position, players were drafted into the NFL from DraftHistory.com. I collected supplemental data from ESPN and some other sources (see appendix). Some descriptive statistics on this population of players follows here, with associated conditional probabilities for draft outcomes:

Table 1: Recruit Class by Recruit Rating

Recruit rating/ High school class	Class of 2011	Class of 2012	Class of 2013	Class of 2014	Total
5 star	3	2	2	1	8
4 star	12	24	33	25	94
3 star	35	24	15	24	98
Total	50	50	50	50	200

Table 2: NFL Outcome by Recruit Rating

Recruit rating/ NFL outcome	Drafted by NFL Team	Undrafted free agent (UFA)	No NFL career	Total

5 star	4 (50%)	1 (12.5%)	3 (37.5%)	8
4 star	17 (18.1%)	3 (3.2%)	74 (78.7%)	94
3 star	12 (12.2%)	10 (10.2%)	78 (77.6%)	98
Total	33 (16.5%)	14 (7%)	155 (75.5%)	200

I use the "consensus" recruit ratings, basically an unweighted average of the ratings assigned to these players by the three major scouting agencies. As I discussed earlier, these ratings are highly subjective and the difference between a 3-star and a blue chip recruit may come down to individual scouts' opinions or how a player looked at practice one day versus another. Using the consensus ratings does not fix these issues, but it does help make this variable less dependent on the whims of particular scouts. 5-star quarterbacks, the most prestigious and sought-after players at one of the highest-status positions in sports, are rare, but small sample size notwithstanding, the conditional probabilities tell us that they seem to enjoy better NFL outcomes than their 4- and 3-star counterparts. Across my four-year sample, more 4 stars than 3 stars are eventually drafted, but more 3 stars than 4 stars are signed to NFL teams as undrafted free agents (UFAs), which bears considerably fewer social rewards and a smaller rookie contract than does being drafted outright, but is still an NFL contract and should still be considered an NFL outcome.

School-side data

Division 1 college football programs are divided into two categories, the prestigious FBS level and the less prestigious FCS level. Both grant scholarships, but the resources and media attention available for players at FBS schools dwarf those available for those at FCS schools. 130 schools existed at the FBS level during my time period. Of these, by 2017 (when the last player from my dataset was drafted), almost half (64) belonged to so-called "power conferences," whose member schools are generally considered to be in the upper echelon of the FBS. Power conferences, a category that in 2017 included the SEC, the Big Ten, the Big 12, the ACC, and the Pac-12, have the longest football histories, the largest fan bases, and the most media attention. Most players who are drafted come from power conference member institutions.[3] I collected data on the 130 FBS schools during the 2011-2017 period, a timeframe covering when the players in my sample were actively competing in college football. The states these universities are located in and the conferences to which they belonged in 2017 are matters of public record, and I collected information on what offensive scheme each college team utilized during this period from a variety of sources listed in the appendix,

[3] One school, Notre Dame, does not belong to a conference but is believed by the college football world to compete on the level of a strong power conference member, and enjoys media attention, recruiting success, and draft outcomes on the level of the best power conference schools.

along with details about the schemes themselves. It is plausible that certain kinds of offense make a quarterback look more draftable to NFL teams, and it is relatedly likely that in-game statistics are evaluated differently for players in different offenses. Coding this variable is necessary for a plausible counterfactual analysis, not only because of its direct impact on player performance but also because of its effects on how that performance is evaluated for the draft.

The most important school-side variable I need to account for, other than power conference status, is schedule-adjusted team quality. Because college football teams have extremely heterogeneous access to resources, there is a lack of competitive balance between them. Even at the FBS level, the best teams in the best conferences are understood to be far better (at football) than the worst teams. Wins against stronger competition are valued more highly than wins against subpar opponents, and losses against good teams are not as damaging to a team's reputation as are losses against teams viewed as underdogs. Any metric for team quality I use must account for this competitive asymmetry. I use "Dr. Colley's Bias Free Matrix Ranking," otherwise known as the "Colley Matrix," a special case of the generalized row sum method for scoring elements based on internal competition designed by Dr. Wesley Colley. This algorithm is designed to provide an objective score of how good every college football team was in a given year, and it is considered a "major selector" by the NCAA, making it the only computer-based system that the NCAA accepts as a legitimate means of determining a national champion. It was used to schedule bowl games, the college football postseason, for a number of years. Colley (2002) lists the advantages of this ranking system compared to qualitative media polls and "byzantine black box" computer models: "[it] has no bias toward conference, tradition, history, etc., (and, hence, has no pre-season poll); [it] is reproducible; [it] uses a minimum of assumptions; [it] uses no ad hoc adjustments; [it] nonetheless adjusts for strength of schedule; [it] ignores runaway scores; [it] produces common sense results." I collect the end-of-season Colley Matrix ratings for each team from the 2011 season to the 2017 season, and take an unweighted average of these ratings for each team to compute what I call the "Colley average," the average evaluation of each college for the 7-year period of my analysis, which is more stable to breakout or unusually poor performances that a school has for only one season. This number ranges from 0 to 1, with a 0 score representing a team that never wins a game against the worst competition available and a 1 score representing a juggernaut expected to defeat even the strongest opposition. There is a specific Colley rating for each FBS team each season, and for schools that were in the FCS at some point in the 2011-2017 period, I use the "FCS group 1" Colley rating for the relevant seasons.

Model specification

I specify a logistic regression model for draft probability based on various player-side and school-side covariates (denoted by the subscript p and s respectively):

$$D = R_p + \sigma_p + Ypass_p + Yrush_p + TD_p + Int_p + C_s + Power_s + Scheme_s$$

where D represents the probability of being drafted, R_p is player race, σ_p is recruit rating in stars, $Ypass_p$ is passing yards, $Yrush_p$ rushing and receiving yards, TD_p total touchdowns, Int_p interceptions, C_s Colley average, $Power_s$ power conference status, and $Scheme_s$ offensive scheme. I operationalize race with a white/Black/other categorization, since 197 of the 200 players are either white or Black. I use "other" as the reference category for race. The reference category for recruit rating is 3 star status, and the reference category for scheme is air raid. The reference category for power status is the lack of that status.

I pare down my population of 200 quarterback recruits to isolate those of them who have valid data for this model specification. Of these players, 196 went on to play college football, with the other 4 being multi-sport athletes who chose baseball over football (all 4 ended up with major league careers). Of these 196 quarterbacks, 20 made a move to a different position at some point, making them statistically incomparable to the rest of the population. Of the 176 remaining quarterbacks, 41 did not put up any passing yards at any point in college at an FBS program, which indicates that their college football careers did not pan out. Some quarterbacks are pure "pocket passers" who tallied zero rushing or receiving yards in college, but recording zero *passing* yards is tantamount to not being a quarterback in a meaningful sense, and I exclude these players from the model.[4]

After this process, I am left with 135 college quarterbacks with valid data. Whether a unit has invalid data is not random in this case, but selective of several features that are relevant to this study design. Whether a player ends up at another sport or another position over playing football as a quarterback is not entirely up to them; they are constrained by the willingness of their coaches to accommodate these choices, and indeed it is often the coaches who are making choices like these for them (Logan et al. 2015). The majority of invalid quarterbacks — 41 people, just over one fifth of the original population — did not give up on football, but they did not achieve the goal of playing in a college football game in the FBS. This could be for a variety of reasons. Perhaps they lacked the talent or physical gifts required to play quarterback at the college level, or perhaps they were outshined and outcompeted by another QB at their school, or perhaps they had off-field issues that prevented them from playing, or perhaps they were injured, and so on. My study design does not give me the ability to speculate on what if any of these things might have been different had players attended

[4] One player, Chad Kanoff, transferred from Vanderbilt to Princeton. The Ivy League schools are all FCS football programs and thus included in the Colley Matrix, but they play fewer games in a normal season than other division 1 programs and Kanoff's stats are not directly comparable to those of his peers, so I exclude him from the model. Kanoff was ultimately signed as an undrafted free agent.

other schools — although in many cases there surely would have been such differences — but generally speaking, if someone never played in a college football game at the FBS level, there is no plausible counterfactual based simply on choice of school that would see them getting playing time elsewhere. Much more fundamental structural, agentive, and stochastic factors would have to be different for it to make sense to include these players in the model. In that sense, there is a consistent pattern with the players who *do* end up included in my model: they are (at least) competent, dedicated players whose coaches have some level of faith in them, and who got their chance, at least relatively speaking, at the dream of parlaying a college career into an NFL contract.

Transfers

The biggest difficulty when working with this data is determining how to deal with players who transferred between colleges during the course of their college career. In 2018, the NCAA introduced the so-called "transfer portal," a new system for managing college athletes' transfer decisions, which along with the loosening of regulations on transferring for college athletes contributed to an enormous increase in the number of college athletes who transfer each year, especially in revenue-generating sports like football, women's volleyball, and men's and women's basketball. My data concerns only players whose entire college careers predated the portal, when transfer agency was much more tightly restricted (Logan et al. 2015). However, a large proportion of players in my dataset still transferred, including 58 of the 135 quarterbacks with valid data for inference. I handle this simply by treating the school where the player ended up as though the player committed there originally, and I treat the school where they were originally as a counterfactual.

This is not a perfect solution for a few reasons. First, if a player "transfers down" from a big-name program to a less prestigious one, draft analysts may still associate the player with the big-name school, which will be hidden to the model. Second, if a player does "transfer down," it likely indicates that they would not have earned the starting quarterback job at their original school, making that counterfactual less plausible. Of the players who transferred, 30 of the 58 transferred to a school that had extended them a scholarship offer out of high school anyway, essentially getting a do-over of their commitment decision and choosing differently. The question of who gets a do-over in this way is fascinating, and I hope to explore it further in future research.

Despite these concerns, I have several reasons for handling transfers in the way I do. First, the transfer decision is an important part of player agency in college football, and it is part of a broader decision complex that includes the player's original commitment decision. Players may well commit to one program knowing that they intend to transfer somewhere else, even before the era of the transfer portal. Second, because the proportion of transfer players who transferred to a school that originally offered them is so high, it indicates that most top 50 high school quarterback recruits who transfer look to schools that originally offered them first, meaning the transfer

decision is colored by the commitment decision such that one cannot be extracted from the other. Third, the aim of this study is not to determine why a player chooses one school over another; it is to determine what effect on the desired outcome of being drafted into the NFL a player's choice has. I cannot opine on whether a player is "transferring down" or otherwise, or why they would transfer versus stick around. What I can examine is the effect of college choice on NFL draft probability, and handling transfers in this way is a plausible, coherent means of measuring that effect. I include a robustness check in the appendix, showing that using players' original school instead of their final school for transfers somewhat degrades model performance and may mask some of the impact of attending a power conference school.

RESULTS AND ANALYSIS

Logistic regression model for draft probability

I specify a logistic regression model based on the parameters laid out in the preceding section, and train it on the factual outcomes for the 135 valid quarterbacks only. UFAs are considered not drafted. The covariate table follows:

```
                  Estimate  Std. Error z value Pr(>|z|)
(Intercept)      -15.978390    6.772880  -2.359   0.0183 *
RaceBlack          1.275651    5.041695   0.253   0.8003
RaceWhite          1.841427    5.039918   0.365   0.7148
RecruitStars4      0.392310    0.935174   0.420   0.6748
RecruitStars5      2.590796    1.583069   1.637   0.1017
FSPassingYds       0.002655    0.001205   2.203   0.0276 *
FSRushRcvYds       0.001820    0.002280   0.798   0.4247
FSTD               0.023696    0.105577   0.224   0.8224
FSInt             -0.172233    0.142201  -1.211   0.2258
ColleyAvg          5.132354    3.365243   1.525   0.1272
Power1             4.111060    2.067633   1.988   0.0468 *
SchemeTripleOpt  -12.208796 1747.919088  -0.007   0.9944
SchemeProStyle    -0.674781    2.872129  -0.235   0.8143
SchemeSpread      -1.965336    2.838629  -0.692   0.4887
SchemeMultiple    -0.248600    2.874212  -0.086   0.9311
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 147.863  on 134  degrees of freedom
Residual deviance:  51.655  on 120  degrees of freedom
AIC: 81.655

Number of Fisher Scoring iterations: 16
```

There are statistically significant effects at the 95% level of confidence for passing yards, the primary quarterback performance metric, and power status, the primary indicator of program prestige. The scales of these variables are very different, but it is clear that passing for more yards in a player's final season and coming from a power conference school (or Notre Dame) are associated with statistically significant positive effects on a player's draft odds. In fact, power conference status increases the odds of being drafted by about a factor of 4, nearly as much as would going to a school with a Colley average of 1 (the maximum possible score, the best program in the country) over a school with a Colley average of 0 (the minimum possible score, the worst program in the country). As for race, white quarterbacks are slightly advantaged over Black quarterbacks according to the model, but these effects are not statistically significant. Although "other" status seems to be a negative for draft odds, controlling for the other covariates the sample size of non-white, non-Black quarterbacks is so small that this effect size is unreliable, which says something about the racial makeup of the quarterback position on its own. Once Black quarterbacks have overcome the forces pushing them to move to other positions, the evidence that they are seriously unfavored compared to white QBs in the draft is not strong, but the practical nonexistence of "other" quarterbacks supports findings in the literature about how Black quarterbacks are stereotyped as their own category, albeit one that is flourishing in the modern NFL, but non-white, non-Black quarterbacks are nearly unimaginable (Niven 2005, Bigler and Jeffries 2008, Early 2011). Rushing/receiving yards and touchdowns have positive effects on draft odds, but these effects are not considered statistically significant. Interceptions, predictably, have a negative effect, but it is also nonsignificant. Colley average has a positive effect on draft odds, which is not quite significant at the 90% level of confidence but close to it. Recruit rating is also interesting: 4-star status does not bear a large effect on draft odds relative to 3-star status, but 5-star status does (though just shy of significance at 90% confidence), reflecting the similar pattern from the descriptive statistics. Controlling for the other independent variables, 5-star status seems to increase the odds of being drafted, suggesting that being so highly rated out of high school is a lingering source of status for NFL-hopeful quarterbacks years later (NFL teams might think that the scouts identified some characteristic about a recruit that might be useful in the league if they take a chance on him). A Bourdieusian approach might hold that status-advantaged players have a competitive advantage in the draft over their lower-status peers, even controlling for their actual on-field performance and the quality of the programs they play for (Bourdieu 1972, 1984, Harvey et al. 2020).

I test the model for predictive performance using k-fold cross-validation with 5 groups and ridge regression (L2 regularization). The outputs from this testing can be found in the appendix, and the model performs quite well under both tests.

Draft probability analysis and plots

I match players to the schools that gave them scholarship offers (or where they transferred), then use the model to estimate a draft probability for each valid player-school combination, a total of 1224 combinations across the 135 players (~9.07 average plausible schools per player). I do not report confidence intervals around these points because the important thing for analysis is not the specific predictions, but the variation between the point estimates. I plot these modeled draft probabilities for both the factual and counterfactual cases below, broken down by recruiting class:

Figure 1:

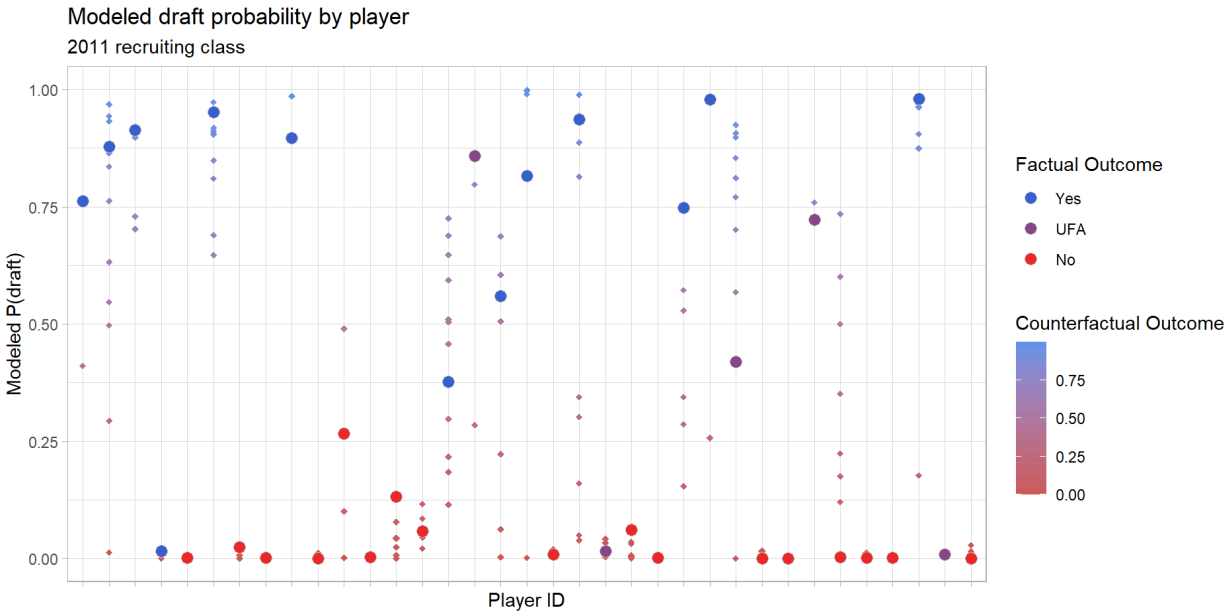


Figure 2:

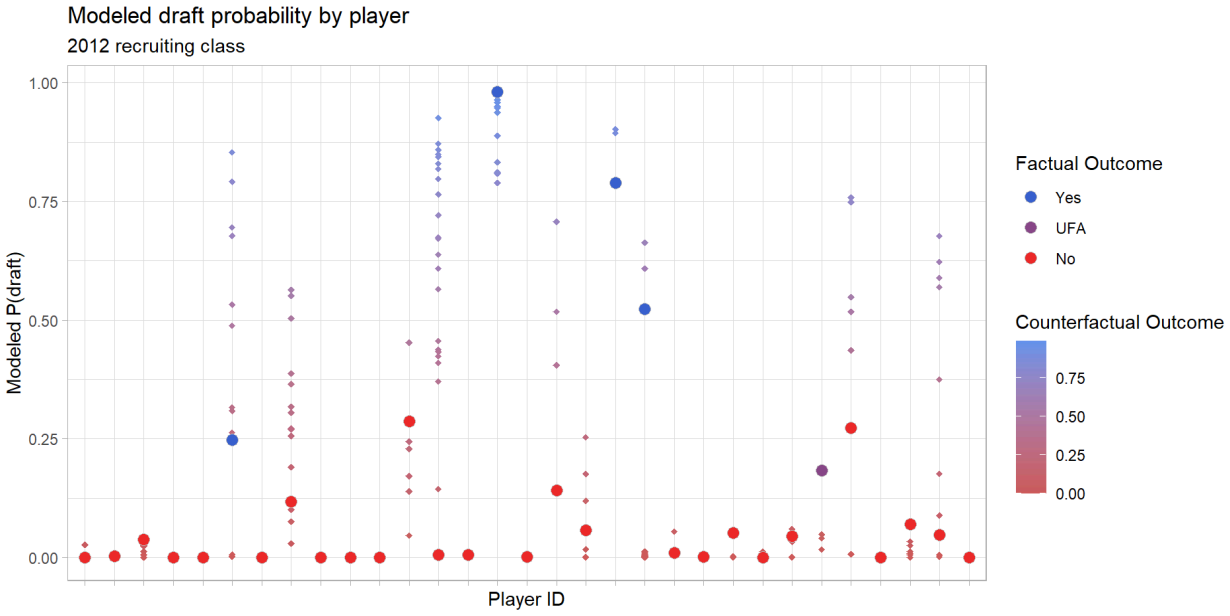


Figure 3:

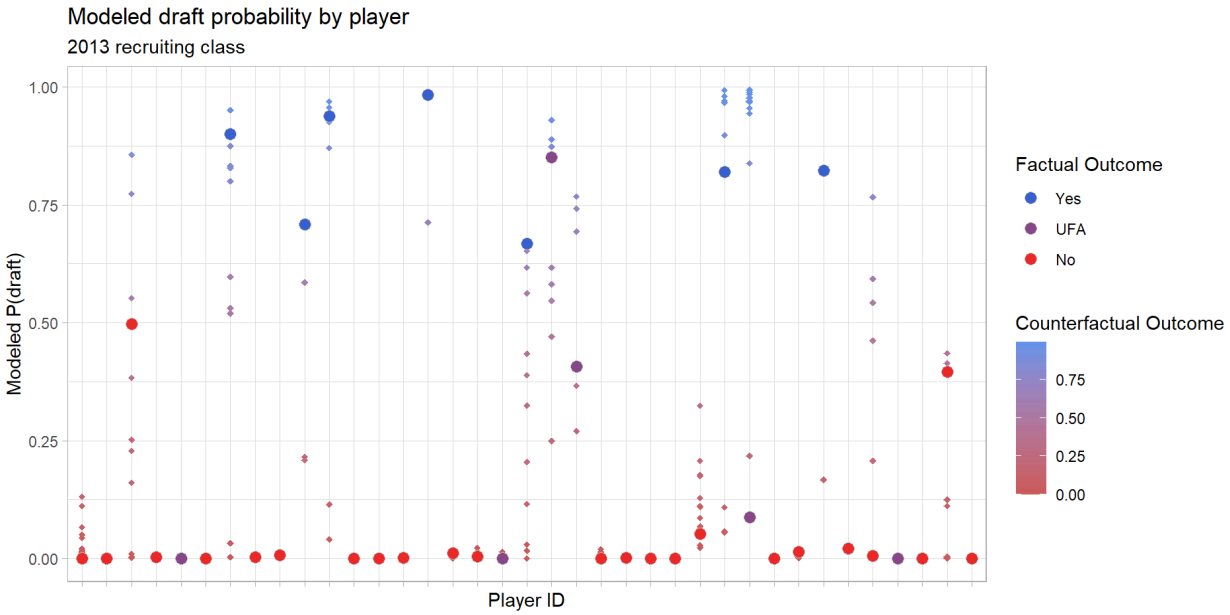
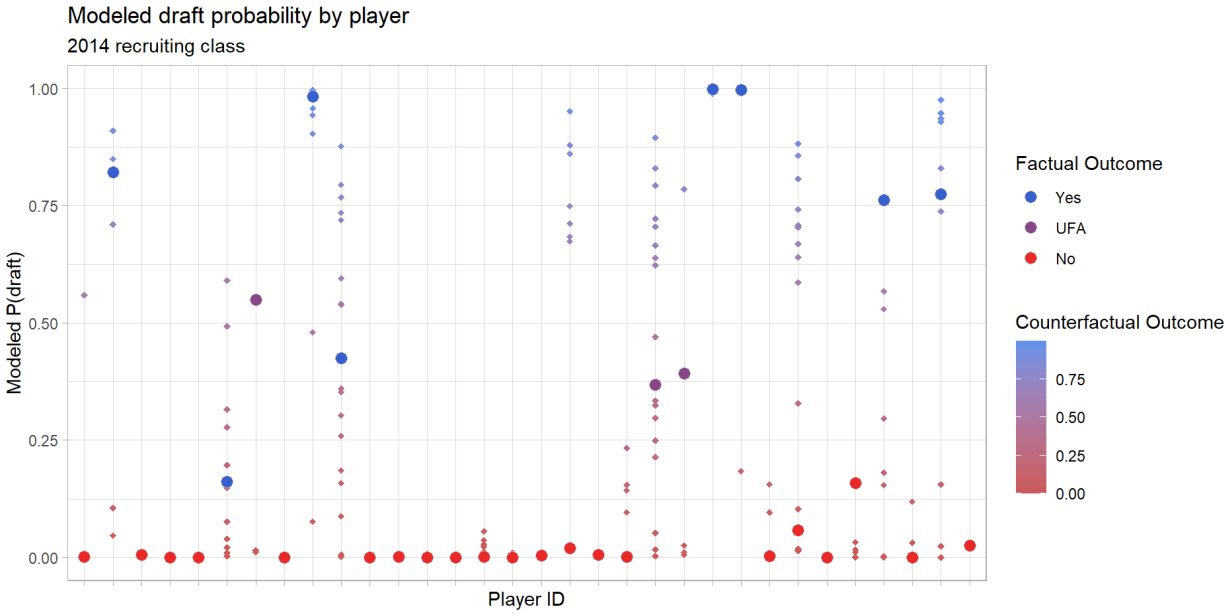
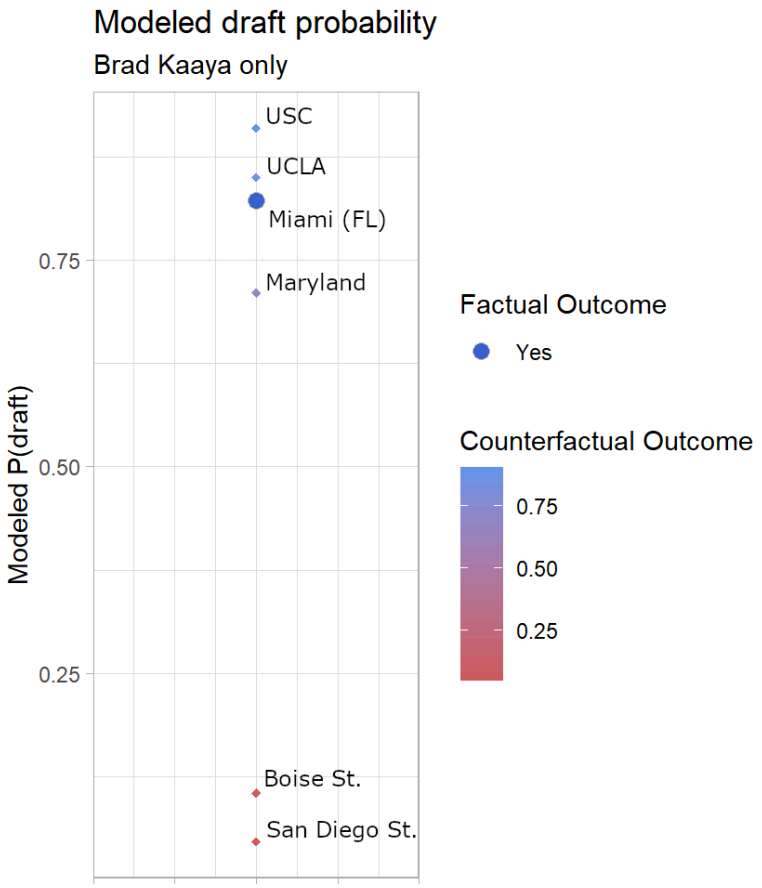


Figure 4:



To help read these plots, we can take a closer look at one recruit as an example: Brad Kaaya, who attended Miami (Florida) but also received scholarship offers from USC, UCLA, Maryland, Boise State, and San Diego State.

Figure 5:



The factual outcome is represented by the large circle, which is colored blue because Brad Kaaya was indeed drafted in the factual case, an outcome with a modeled probability of ~0.822. Brad had two scholarship offers the model believes were even likelier to get him drafted, USC (~0.910) and UCLA (~0.849). At Maryland, the model gives Brad a draft probability of ~0.710. Brad's non-power scholarship options, Boise State and San Diego State, are associated with much lower modeled draft probabilities of ~0.106 and ~0.047 respectively.

Modeled draft probabilities are distributed quite heterogeneously across the players in the sample. In Table 3 below, I lay out six ideal types to which the players seem to converge. Most players, represented by the sedimentary layer of red circles across the bottom of the plots, have low enough in-game stats that the model predicts they will likely never be drafted, regardless of what school they choose. These players converge to two ideal types with low leverage from college choice: the "happy to be here" players, who received few FBS offers, were not regarded as elite recruits, and did not end up impressing in college, and the "glory days" players who were regarded highly enough in high school to get many offers, but who did not perform well enough in college to get drafted, no matter where they went. Some players, the "lunch pail"[5] ideal

[5] This term is used to refer to college football players who play well and are dedicated to their team, but do not really stand out on the field. They are generally regarded as good locker room presences with strong work ethics, not stars but key contributors.

type, have average or above-average statistical performance and a high diversity of scholarship offers, and these players have very large ranges of modeled draft probabilities across the schools they could plausibly have attended. This reflects a high level of leverage. The "cookie crumblers" perform similarly to the "lunch pail" players, but they have a lower diversity of options, which results in moderate leverage. Their likelihood of being drafted depends mostly on how the cookie crumbles, so to speak, in terms of which offers they get. A few players have exceptional or very good in-game stats coupled with offers from a few solid programs, perhaps due to being overlooked in high school. These players, the "diamonds in the rough," have lower leverage; that is, the only choices before them eventually lead to a high probability of being drafted into the NFL. Last, the "elite" quarterbacks are highly touted out of high school and make good on their potential in college, but have moderate leverage because they have some plausible bad options, whereas "diamonds in the rough" don't have many bad options to choose. Brad Kaaya is somewhere between a "cookie crumbler" and an "elite": a medium number of options, very solid but not exceptional in-game performance, and moderate leverage with mostly good outcomes but a risk of choosing a non-power school associated with a risk of a worse outcome.

Table 3: Ideal Types

Characteristic/ Ideal type	Leverage	Performance	Diversity of options
Diamond in the Rough	Low	High	Low
Happy to Be Here	Low	Low	Low
Glory Days	Low	Low	High
Elite	Moderate	High	High
Cookie Crumbler	Moderate	Moderate	Low
Lunch Pail	High	Moderate	High

Leverage analysis and plots

As a means of measuring leverage over the NFL draft outcome for my population of interest, I compute the interquartile range (IQR) of the modeled draft probabilities for each player. I use IQR instead of another measure of variability like range or standard deviation so as to be more robust to outliers, which are visibly not uncommon in my data (see figures 1-4). These outliers may represent the non-power schools that occasionally extend scholarship offers to highly touted recruits otherwise deciding between prestigious power conference programs, who are very unlikely to choose the

non-power programs, or they may correspond to the power conference schools who occasionally take a shot on a recruit who is otherwise flying under the radar. Below, I plot the IQR for each player against three key metrics. First, I visualize the relationship between this measure of leverage and the raw number of scholarship offers received, which is a cruder but more straightforward representation of agency than leverage is. This raw count is highest for five stars and the other most pursued recruits, meeting Bourdieu's (1993) expectation that the highest-status actors should have the most *options*, but they do not necessarily have the most *leverage*. Next, I plot leverage against the Colley average of the school each player attended in the factual case. Then, I plot it against passing yards, the most statistically significant continuous metric for player-level performance. These plots include non-parametric smoothing lines, calculated with the Locally Estimated Scatterplot Smoothing (LOESS) method, to visualize the trends I will be discussing regarding these relationships. In the appendix, I include similar plots for leverage calculated with standard deviation instead of IQR. I also expand on my reasoning for using IQR and discuss in what contexts other measures of variability might be more appropriate.

Figure 6:

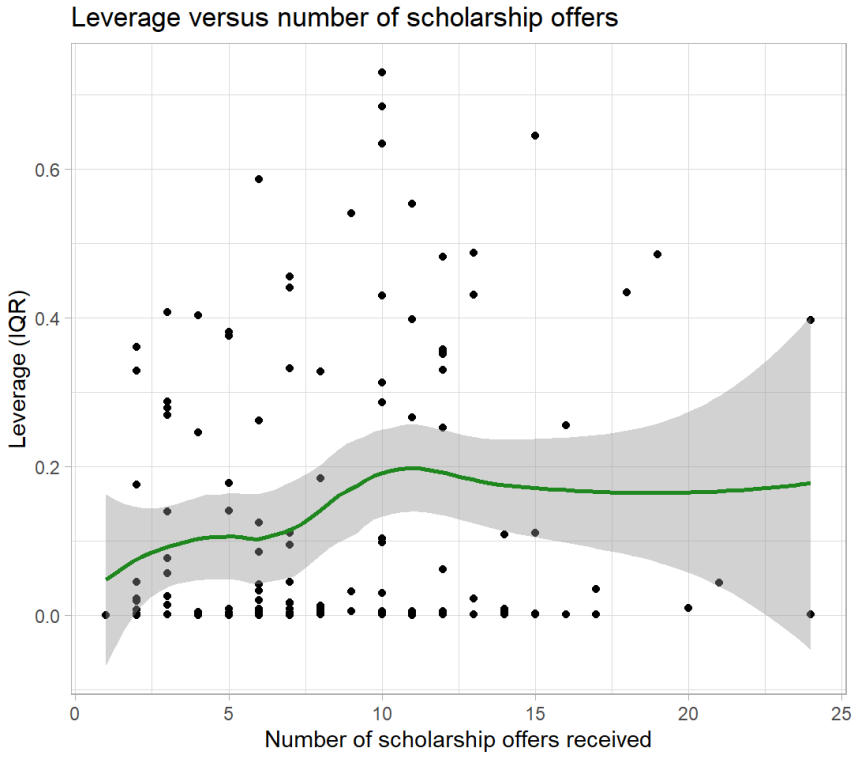


Figure 7:

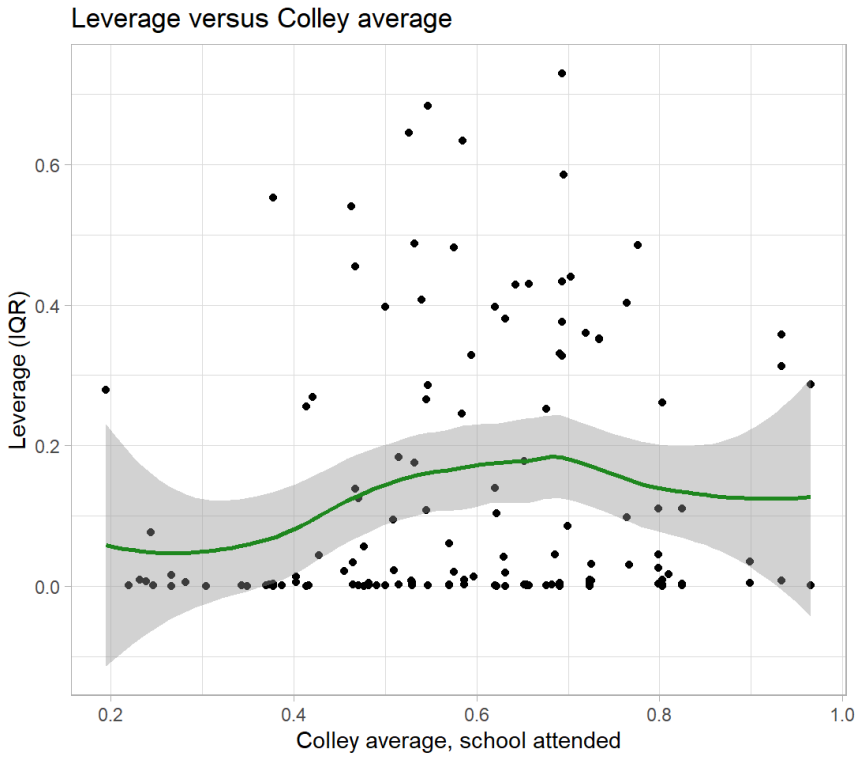
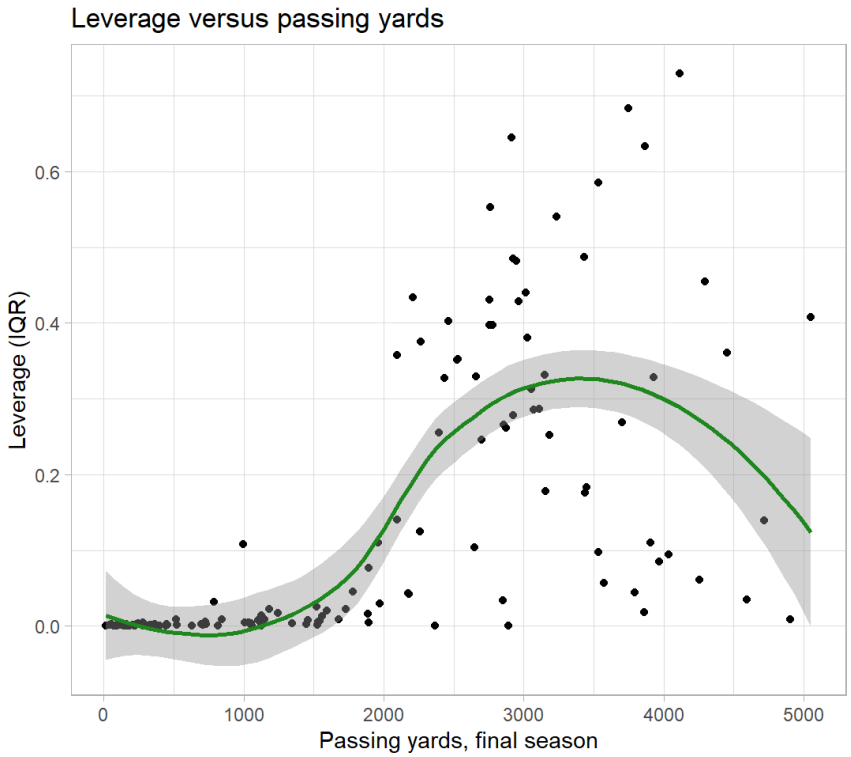
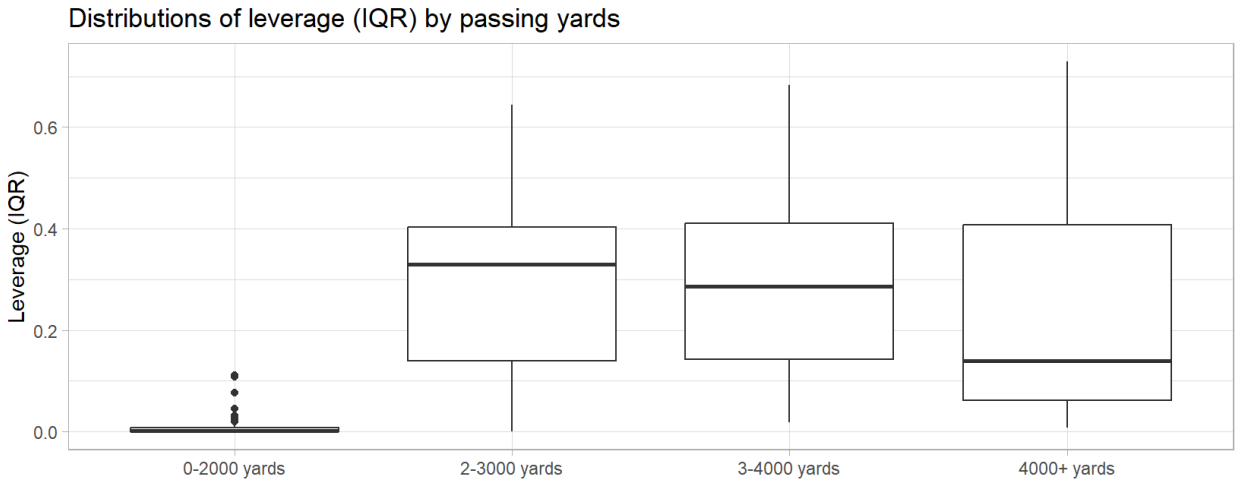


Figure 8:



These results are interesting because they contradict some traditional sociological thinking about agency, broadly defined. There may be a loose positive association between leverage and the number of scholarship offers, but there is not a clear linear relationship, and certainly not after 10 offers, which is about the average number. Receiving more offers — that is, having more options — does not necessarily mean a player has more leverage, once a certain number of options are available. It is not the number of options, but the diversity of options that make a decision important for the world. The findings regarding Colley average support a clear pattern where the "middle class" of players, corresponding to the "cookie crumbler" and especially the "lunch pail" ideal types, have the most leverage agency. Players who in the factual case attended schools with about average Colley ratings having the highest leverage when measured in this way. As for passing yards, the same relationship emerges. Below about 2000 passing yards, players (the "happy to be here" and "glory days" ideal types) all have low leverage because their statistical performance is not good enough for them to be considered draftable. The 2000-yard cutoff is in fact quite stark — it seems that NFL teams are extremely uncomfortable drafting players whose passing yard total starts with "1," but once that largely arbitrary bar is cleared, a player becomes draftable. Above the 2000-yard margin, leverage rapidly increases, but the best performing players with over about 4000 passing yards (the "diamond in the rough" and "elite" ideal types) do not display more leverage than their solid but not exceptionally well performing counterparts with between 2000 and 4000 passing yards. The median leverage for the 9 players with at least 4000 passing yards is ~0.139, and the median leverage for the 29 players with between 2000 and 3000 passing yards is considerably larger, ~0.329. For the 23 players between 3000 and 4000 passing yards, the median leverage is ~0.286. The 74 players with below 2000 passing yards are, on the whole, undraftable, and a median leverage of ~0.002, essentially zero, reflects that.

Figure 9:



DISCUSSION

What causes leverage to be distributed like this in this case? My finding that the "middle class" of players have the most leverage seems to run counter to, for example, the Bourdieusian approach identified by Harvey et al. (2020) where the most privileged, most powerful actors have access to a broader and more effective set of strategies that can be applied to "nullify the moves" of their lower-status competitors (Bourdieu 1993). It is therefore important to clarify that having more *leverage* does not necessarily equate to having more *privilege*. In the competitive field of American college football, where the source of capital being fought over is the NFL draft, the most elite players likely have less leverage than the "middle class" because leverage is a double-edged sword. I mean by this that more leverage means a wider range of outcomes, not just better outcomes, because actors in this case are free to make a "bad" decision, at least in terms of their draft odds. The most leverage over draft chances belongs to those players who are good enough, and favored enough by status factors like recruit rating, to be considered draftable — but not so good or so favored that their being drafted is a foregone conclusion. Those exceptional players, the "elite" and "diamond in the rough" ideal types, either don't have very many offers out of high school or receive mostly offers from the highest-tier programs, and go on to shine statistically in college, so they have a safety net of sorts where there are no or very few options they can choose that will result in anything other than a high chance of being drafted. This itself is a form of privilege: the best and most favored players are not even faced with the "bad" choices that their competitors must avoid, whether by luck or by strategy.

For the high-leverage players, who have certain options that are likely to lead to their desired outcome and others that are not, Bourdieu again becomes quite relevant. Under the difference-making perspective on agency, the difference that these actors want to make is between getting what they want and not getting it, and they are incentivized to practice behaviors that help them decline the options that are less likely to lead to the outcome they are trying to get. Over time, we learn from Bourdieu (1972), as more and more actors follow these behaviors, they crystallize into strategies, central features of *habitus* that reward those who faithfully abide by them. Some fundamental work on career decision-making, for example Hodkinson and Sparkes (1997), builds on this idea by investigating which strategies in making a career form along "the partly unpredictable pattern of turning-points and routines that make up the life course" (29). In the case of college football, one such strategy is the simple heuristic, "never go to a non-power conference school when a power conference school is available." For example, consider again Brad Kaaya (Figure 5). According to my model, his leverage (IQR) over the probability of being drafted is ~0.586, and his median draft probability is ~0.766. By adopting the "never go to a non-power conference school" heuristic, he can reduce his leverage to ~0.071, even though his median draft probability only increases modestly, to ~0.836. Thanks to this strategy, Brad can considerably decrease his

leverage to protect himself from the risk of going undrafted, even though it does not actually increase his median chance of being drafted very much. It is of course far from impossible to get drafted from a non-power program, but this strategy seems like a good one. Indeed, 188 of the 200 players in my data followed it, initially committing to power conference schools (or Notre Dame), even though at least 100 of them received offers from both power and non-power programs.

One limitation of this study, of my analytical methods, and really of the social world, is that leverage can only be estimated retrospectively. Without knowing what the outcomes were in the factual case, it is highly speculative to model them for the counterfactual cases; that is, it is much easier to say what *would* have happened had an actor done one thing or the other in the past than it is to say what *will* happen in the future if an actor chooses an option available to them in the present. This temporal element is central to the Emirbayer and Mische (1998) definition of agency, and strategies are so desirable for agents in competitive fields because they help reduce the uncertainty it introduces. One exercises leverage over countless outcomes in the world whenever one acts, or refrains from acting (Giddens 1984)… but it is not possible to know what leverage one had until the outcomes one has influenced are observed.

Another limitation, or rather research challenge, worth noting is that estimating plausible counterfactuals for continuous outcomes (for instance, draft position) is a considerably more difficult task than estimating them, as I do in this paper, for binary or odds outcomes (like draft probability). Questions about which loss function to use, whether to estimate leverage by distance between counterfactual outcomes or variability among them, and how to compare leverage for probability (measured from 0 to 1) to leverage for continuous outcomes (measured in whatever units the outcome is) abound. There is also the problem of temporality. The literature is very concerned with the temporal properties of agency — the Emirbayer and Mische (1998) definition contains the term "temporal" twice, and everyone from Parsons (1937) to Archer (2003) is concerned with how agency transforms social structures over time — and this study has side-stepped that problem because everyone makes the decision of where to go to college at a similar time (during their final year of high school) and ultimately declares for the draft 3-5 years later, if they ever do so. For decision complexes where actors can choose *when* to act and *whether* to act in addition to *how* to act, leverage becomes much harder to estimate. Future research involving leverage will need to grapple with these methodological concerns.

CONCLUSION

So what can leverage, and this study specifically, actually teach us about agency? First, sociologists mean many things when they say "agency," so many that there is no operational definition for the term that can simultaneously satisfy every theoretical concern and present a useful framework for measurement and analysis. The difference-making perspective, where agency is defined as the exercise of actions that make a difference in the world, is incomplete but operationally useful, providing an opportunity

to measure agency, a problem with which the literature has long struggled. Leverage, the measure of how much difference a given decision induces in a specific outcome in the world, is an elegant and potentially rigorous way to achieve this task.

Under this perspective, decisions are more important than others if an actors' choices induce a large difference in the odds of the outcome occurring for binary outcomes, or in the value of the outcome for numerical outcomes. A key insight here is that the diversity of options matters much more for leverage than the number of options (Figure 6). Having many options that are not meaningfully different in terms of their contribution to the outcome an actor is influencing does not increase leverage any more than having one comparable outcome would. This has implications beyond college choice among recruited football players. Consider a family choosing between neighborhoods when moving to a new city. Just a few neighborhoods that differ a lot in terms of amenities, transit, schools, local ordinances, demographics, and so forth would give this family more leverage over their life chances than would many homogeneous neighborhoods. Or consider a young person choosing between expensive hobbies when their family can only afford one. Perhaps choosing between many musical instruments bears less leverage over what social milieu the young person will belong to in high school than choosing between one instrument, one sport, and one visual art. The potential applications of this way of thinking about agency are plentiful.

This project is only possible because of how unusually granular and detailed the sports data I use are. Future research into leverage, how to measure it, and how it is distributed in social arenas other than sports may benefit if targeted at contexts with similarly high-quality data. Dale and Krueger (2002) estimate "returns," a very similar concept to leverage, for future income based on how selective of a university a student attends, using excellent data from the College and Beyond database. Usage data from dating apps could be examined to investigate leverage in relationships, social media data could help explain leverage in network formation and evolution, data from historical correspondence could be used to determine the significance of a person or family moving or migrating, and so on, across all the social worlds where we want to know more about how important decisions structure patterns of agency in our lives.

APPENDIX A: FOOTBALL GLOSSARY

Institutional terms

College football: The social institution surrounding the playing of football by students at American colleges and universities, with team symbols and history associated with those schools. College football is well over 100 years old, and has become an extraordinarily popular, moneyed institution and way of life for many Americans.

Conference: A group of football-playing schools that have agreed to share resources and play football against each other. Some conferences are much more prestigious than others, and the most prestigious conferences are called power conferences.

Division 1 (D1): The highest-status level of college athletics, and football in particular. D1 teams issue athletic scholarships, build stadiums, and devote a disproportionate level of resources to sports, especially "revenue-generating sports" like football.

Football Bowl Subdivision (FBS): Formerly known as Division 1-A, the FBS is the upper echelon of division 1 football. FBS schools are allowed to offer the most scholarships, and participate in inter-conference postseason "bowl games." The FBS historically lacked a means of determining a national champion, and the BCS (Bowl Championship Series) and CFP (College Football Playoff) were developed to meet that need.

Football Championship Subdivision (FCS): Formerly known as Division 1-AA, the FCS is the lower level of division 1 football. The FCS determines a national champion through a postseason tournament, similar to college basketball.

National Collegiate Athletic Association (NCAA): The governing body of most American intercollegiate athletics. The NCAA regulates college sports, and has significantly restricted certain rights and privileges for players, including transferring and profiting from their own name, image, and likeness (NIL).

Professional football: The institution surrounding the playing of football by professional athletes. Professional football was once much less important than college football, but in the decades following World War II, it rose to prominence as one of the four major American sports alongside baseball, basketball, and hockey.

National Football League (NFL): The oldest and by far the most significant professional football league in the United States. Nearly all NFL players were once college players.

NFL draft: The highly publicized, televised yearly event where NFL teams select college players to sign to their rosters. Being selected in the NFL draft is one of the highest honors in college football.

Recruit: A high school football player being actively recruited by college teams. Coaches, assistant coaches, and staff members representing these teams will court players starting as early as middle school, with the intention of ultimately extending scholarship offers to players they want for their team.

Recruit rating: The rating (from 1 to 5 stars) assigned by professional scouts to a recruit. 1-2 stars means the scouts believe the recruit is not likely to make an impact at the highest level, 3 stars means the scouts believe the recruit may have success in college, 4

stars means the scouts believe the player is likely to be a college starter, and 5 stars means the scouts are very confident that the player is of an elite caliber. These ratings are extremely subjective, and players under- or outperform their ratings all the time.
Scout: A professional evaluator of football talent, employed either by a college to evaluate high school and transfer players, by an NFL team to evaluate college players for the draft, or by a media source to produce external evaluations.

Gameplay terms

Line of scrimmage: Football is a game based on control of territory. The offense tries to move the ball up the field to the end zone, measured in yards, and the defense tries to prevent them from doing so, or cause a loss of yards. The line of scrimmage marks where the ball starts each play.
Quarterback (QB): The quarterback is the most important player in the football offense. It is the job of this player to get the ball after the snap (the beginning of play), then either pass the ball to a receiver, hand the ball off to another player who will run with it, or run with the ball themself. As football has evolved, the passing game has become more and more important, and as the team's primary passer and signal caller (who communicates what play will be attempted to the rest of the offense), the quarterback must make countless high-pressure decisions and also perform as an athlete.
Passing yards: When the quarterback passes the ball upfield to a receiver, the number of yards gained on the play (or the distance from the line of scrimmage to where the ball is after the play) is recorded as "passing yards" for the quarterback and "receiving yards" for the receiver. This includes both the yards traveled by the ball in the air, and any yards gained by the receiver running with the ball after the catch. The quarterback may only throw a forward pass from behind the line of scrimmage.
Rushing yards: When a player runs from behind the line of scrimmage with the ball, the number of yards gained on the play is recorded as "rushing yards." If the player is tackled behind the line of scrimmage, they are assigned negative rushing yards.
Touchdown (TD): If an offensive player has possession of the ball in the opposing team's end zone, it is recorded as a touchdown (worth 6 points, the most that can be gained on a single play, with an opportunity for an extra 1 or 2 points afterward). If the touchdown comes from a completed pass, it is recorded as a passing touchdown for the passer and a receiving touchdown for the receiver. If a player runs into the end zone with the ball, it is recorded as a rushing touchdown for that player. Touchdowns are almost always the best possible outcome for passing plays.
Interception: If the quarterback throws the ball and it is caught by a defensive player who retains possession, it is recorded as an interception. Possession switches and the opposing team's offense comes onto the field to begin at the point of the interception. Interceptions are almost always the worst possible outcome for passing plays.

Offensive schemes

Scheme: The general strategy employed by a football offense or defense. Schemes are remarkably diverse, and I am not trying to capture the whole world of football offense in these few broad categories. They are simply intended to reflect general patterns in the systems college quarterbacks were part of from 2011-2017.
Air raid: A scheme most often seen in college football which heavily emphasizes passing plays and de-emphasizing running plays. Quarterbacks in this scheme should have the most passing yards, all else being equal.
Spread: A scheme where the offense comes to the line of scrimmage with receivers set up far away from the offensive line and the quarterback, to "spread out" the field. Spread offenses generally feature a lot of passing, though not as much as air raid.
Multiple: A scheme involving plays from multiple other schemes. Multiple offenses typically incorporate both passing and running plays to keep defenders guessing.
Pro-style: A scheme meant to resemble offense as seen in the NFL, with a mix of passing and running plays and a heavier emphasis on misdirection through play action (pass plays meant to look like run plays) and RPOs (run/pass option plays).
Triple option: A unique scheme traditionally used by the service academies, where nearly every play is a run and the main task of the quarterback is to determine which runner should get the ball, not to make a forward pass. The triple option resembles rugby, and antiquated football strategy before the development of the forward pass. I also coded other unconventional single-play offenses like Wake Forest's "slow mesh" and Florida International's "pro set" under this category.

APPENDIX B: DATA SOURCES

Data on player-side statistics was sourced from sports-reference.com wherever possible and ESPN otherwise. Data on schools was sourced from sports-reference.com and colleyrankings.com, with additional information on offensive schemes obtained from media coverage, team playbooks where available, the *NCAA Football 14* video game (itself researched from actual playbooks), and Wikipedia. Recruiting data was obtained from 247Sports.com whenever possible, with supplemental information from On3.com, media coverage, and ESPN.

APPENDIX C: MODEL TESTS AND ROBUSTNESS CHECKS

I test the model's performance using k-fold cross-validation with 5 groups. The ROC score from resampling is ~0.904, representing about 90% of the values falling under the ROC curve. The sensitivity is ~0.923, and the specificity is ~0.719. This performance is not perfect, but it's quite good, and I regard this model as plausible for conducting counterfactual analysis. The extremely large standard error for the triple option offensive scheme results from the triple option being a perfect predictor that a quarterback will not be drafted, which is expected because the triple option is a highly

unconventional strategy with essentially no relationship to how NFL football is played. A variance inflation factor (VIF) test associates the scheme variable with a score of only ~1.82, with no variable associated with a score above ~7.08 (touchdowns). These scores are not high enough to suggest serious issues with collinearity.

```
Generalized Linear Model

135 samples
  9 predictor
  2 classes: 'No', 'Yes'

No pre-processing
Resampling: Cross-Validated (5 fold)
Summary of sample sizes: 109, 107, 109, 107, 108
Resampling results:

  ROC        Sens       Spec
  0.9040249  0.9228571  0.7190476
```

I also test the model using ridge regression (L2 regularization). I find that the value of the λ parameter that results in the least error is very small (0.0631), which indicates that regularization may help slightly with model stability, but does not change the model coefficients very much; that is, the model is predictively accurate without being significantly simplified. Ridge regression is my preferred method of regularization because it does not drop variables that might contribute less to the outcome but are theoretically important (for example, race or offensive scheme) like LASSO (L1 regularization) does.

```
predictors <- data.matrix(filtered[, c("Race", "RecruitStars", "FSPassingYds",
"FSRushRcvYds", "FSTD", "FSInt", "ColleyAvg", "Power", "Scheme")])
outcome <- as.numeric(filtered$dummyDrafted)
Call:  cv.glmnet(x = predictors, y = outcome, lambda = lambda_seq, alpha = 0)

Measure: Mean-Squared Error

    Lambda Index Measure       SE Nonzero
min 0.0631    33  0.1067 0.012155       9
1se 0.6310    23  0.1181 0.008187       9
```

I build the model again under the same specifications, but with players' originally committed school, regardless of if they transferred. The outputs are below — effect sizes are similar across the board, but the AIC is worse, suggesting that players' final schools seem to have a cleaner statistical relationship to their eventual draft odds than their original schools (should those schools differ). An odd quirk of this model is that, controlling for other player-side stats, touchdowns for some reason have a small,

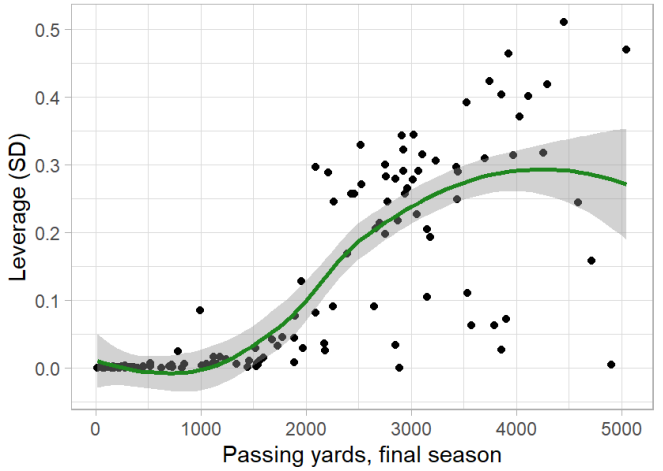
nonsignificant negative effect on draft probability. This may be related to the modest collinearity concerns associated with this variable by the VIF test. The impact of attending a power conference school is also smaller than in the original model, likely because of players who "transfer down" and ultimately go undrafted. The other effects are all as expected from the original model, and this check does not alter my analytical conclusions or lessen my confidence in my model's ability to produce plausible counterfactual outcomes.

```
                  Estimate  Std. Error z value Pr(>|z|)
(Intercept)      -11.190421    4.214588  -2.655  0.00793 **
RaceOther         -1.215368    4.131451  -0.294  0.76862
RaceWhite          0.061676    0.910317   0.068  0.94598
RecruitStars4      0.265355    0.801702   0.331  0.74065
RecruitStars5      2.014801    1.417930   1.421  0.15533
FSPassingYds       0.002889    0.001169   2.473  0.01341 *
FSRushRcvYds       0.002223    0.002090   1.064  0.28744
FSTD              -0.010373    0.093382  -0.111  0.91155
FSInt             -0.219241    0.122065  -1.796  0.07248 .
ColleyAvg          5.161987    2.981869   1.731  0.08343 .
Power1             1.519292    1.725505   0.880  0.37859
SchemeMultiple    -0.474153    2.356877  -0.201  0.84056
SchemeProStyle    -1.070655    2.386614  -0.449  0.65371
SchemeSpread      -2.592386    2.419155  -1.072  0.28390
SchemeTripleOpt  -13.304879 2161.635904  -0.006  0.99509
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 147.320  on 133  degrees of freedom
Residual deviance:  62.548  on 119  degrees of freedom
  (1 observation deleted due to missingness)
AIC: 92.548

Number of Fisher Scoring iterations: 16
```

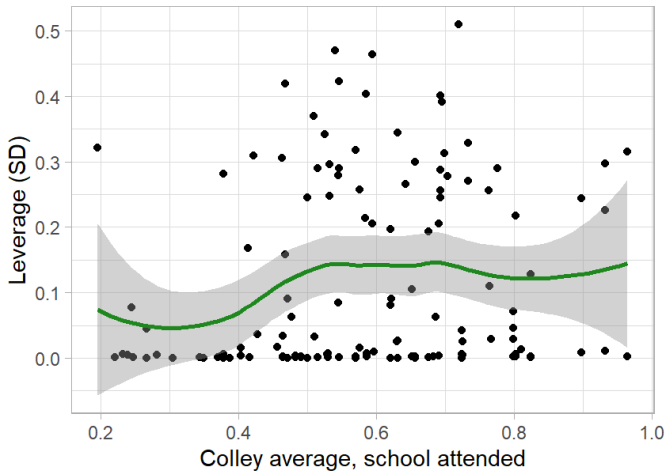
I plot leverage, calculated with standard deviation (SD) instead of IQR, against the same three predictors I use in the results section. The results are mostly similar, but with some apparent inflation of leverage values for players with few offers and players with high passing yards, which slightly reduces the visual impact of the "middle class" effect I discuss in the analysis. These are the players for whom SD should be larger than IQR, since the players with few offers have more sparse distributions of offers and since elite quarterbacks have skewed draft probability distributions from non-power programs taking a chance by offering them a scholarship. This does not meaningfully change my findings, but it is an interesting point of comparison. It is also worth noting that 5 players have non-finite standard deviations due to receiving only one offer.

Leverage versus number of scholarship offers

Leverage versus passing yards



Leverage versus Colley average



WORKS CITED

Alkire, Sabina. 2005. "Subjective Quantitative Studies of Human Agency." *Social Indicators Research* 74 (1): 217-260.

Archer, Margaret S. 2003. *Structure, Agency, and the Internal Conversation*. Cambridge University Press.

Berri, David J. and Rob **Simmons**. 2011. "Catching a draft: on the process of selecting quarterbacks in the National Football League amateur draft." *Journal of Productivity Analysis* 35 (1): 37-49.

Bigler, Matthew and Judson L. **Jeffries**. 2008. ""An Amazing Specimen": NFL Draft Experts' Evaluations of Black Quarterbacks." *Journal of African American Studies* 12 (2): 120-141.

Bourdieu, Pierre (trans. Richard Nice). 1972. *Outline of a Theory of Practice*. Cambridge University Press.

Bourdieu, Pierre (trans. Richard Nice). 1984. *Distinction*. Harvard University Press.

Bourdieu, Pierre (trans. Richard Nice). 1993. *Sociology in Question*. Sage Publications.

Colley, Wesley. 2002. "Colley's Bias Free College Football Ranking Method: The Colley Matrix Explained." *colleymatrix.org*.

Crabtree, Jeremy. 2022. "Recruits Say NFL Development, Coaches Matter More than NIL." *On3.com*.

Dale, Stacy Berg, and Alan **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.

Early, Gerald L. 2011. *A Level Playing Field*, ch. 3: "Donovan McNabb, Rush Limbaugh, and the Making of the Black Quarterback." Harvard University Press.

Emirbayer, Mustafa and Jeff **Goodwin**. 1994. "Network Analysis, Culture, and the Problem of Agency." *American Journal of Sociology* 99 (6): 1411–1454.

Emirbayer, Mustafa and Ann **Mische**. 1998. "What Is Agency?" *American Journal of Sociology* 103 (4): 962–1023.

Ewick, Patricia and Susan S. **Silbey**. 1995. "Subversive Stories and Hegemonic Tales: Toward a Sociology of Narrative". *Law and Society Review* 29 (2): 197–226.

Feinstein, Josh. 2018. *Quarterback: Inside the Most Important Position in Professional Sports*. New York: Doubleday.

Flake, Melanie. 2019. "I was shocked at the cost to play high school football." *Idaho Ed News*, via *idahoednews.com*.

Franzosi, Roberto, Gianluca **De Fazio**, and Stefania **Vicari**. 2012. "Ways of Measuring Agency: An Application of Quantitative Narrative Analysis to Lynchings in Georgia (1875—1930)." *Sociological Methodology* 42: 1–42.

Fuchs, Stephan (2001). "Beyond Agency." *Sociological Theory* 19 (1): 24–40.

Giddens, Anthony. 1976. *New Rules of Sociological Method*. London: Hutchinson.

Giddens, Anthony. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Polity Press, in association with Basil Blackwell, Oxford.
Harvey, Charles et al. 2020. "Bourdieu, strategy and the field of power." *Critical Perspectives on Accounting* 73.
Hendricks, Wallace, Lawrence **DeBrock**, and Roger **Koenker**. 2003. "Uncertainty, Hiring, and Subsequent Performance: The NFL Draft." *Journal of Labor Economics* 21 (4): 857–886.
Hodkinson, Phil and Andrew C. **Sparkes**. (1997). "Careership: A Sociological Theory of Career Decision Making." *British Journal of Sociology of Education* 18 (1): 29–44.
Jejeebhoy, Shireen et al. 2010. "Measuring Agency among Unmarried Young Women and Men." *Economic and Political Weekly* 45 (30): 56–64.
Lindsay, Andrew. 1973. "Emergence of the Black Quarterback." *Black View* 1 (4): 6, 23.
Logan, Alvin D., Louis **Harrison**, Jr., and Alex **Logan**. 2015. "4 Years a Football Player: The Social Reproduction of Restricted Agency." *Race, Gender, and Class* 22 (1-2): 36–53.
Marquez-Velarde, Guadalupe et al. 2023. "The Paradox of Integration: Racial Composition of NFL Positions from 1960 to 2020." *Sociology of Race and Ethnicity* 9 (4): 1–19.
McCarthy, Michael. 2015. "US pediatricians call for tougher rule enforcement to reduce football injuries." *BMJ: British Medical Journal* 351.
Navia, Jason. 2012. "Sitting on the Bench: the Failure of Youth Football Helmet Regulation and the Necessity of Government Intervention." *Administrative Law Review* 64 (1): 265–293.
Niven, David. 2005. "Race, Quarterbacks, and the Media: Testing the Rush Limbaugh Hypothesis." *Journal of Black Studies* 35 (5): 684–694.
Parsons, Talcott. 1937. *The Structure of Social Action*. Glencoe, IL: The Free Press.
Sewell, Jr., William H. 1992. "A Theory of Structure: Duality, Agency, and Transformation." *American Journal of Sociology* 98 (1): 1–29.
Volz, Brian. 2015. "Race and Quarterback Survival in the National Football League." *Journal of Sports Economics* 18.