Journal of Economics, Race, and Policy
https://doi.org/10.1007/s41996-023-00118-y

RESEARCH

Discrimination in a Rank Order Contest: Evidence from the NFL Draft

Ian Gregory-Smith[1] · Alex Bryson[2] · Rafael Gomez[3]

Received: 5 August 2022 / Revised: 28 March 2023 / Accepted: 8 April 2023

Abstract
This paper examines discrimination in the NFL draft. The NFL is a favorable empirical setting to examine the role of skin color because franchise selectors are required to make rank-order judgements of players based on noisy signals of future productivity. Since wages are tightly related to the rank-order of the draft for the first four years of a player's career, even if discrimination plays only a marginal role in selection, there could be a large discriminatory impact. We observe racial differences in drafting. However, much of the variation is explained by Black and White players selecting into different playing positions. Conditional upon a large set of control variables, including athletic performance at a marque selection event (the NFL combine), we do not find robust evidence of racial discrimination in NFL drafting between 2000 and 2018. However, we do find some evidence that Black players are disadvantaged relative to White players in later rounds of the draft.

Keywords Racial discrimination · Productivity · NFL draft

Introduction

Entry into professional careers usually requires potential employees being invited to a selection event such as an interview or an assessment day. Performance at this event, combined with other information about the candidate then determines the rank order of candidates and who is offered the job. In some professions such as a law, medicine and academia, graduates will try out for several potential employers and their rank order will be strongly correlated with their starting salary. By this process, top candidates are sorted into the highest paying jobs.

This process is at play in the selection of professional American football players out of college via the NFL draft. College players are invited to a selection event, called the NFL Combine, at which their performance over a series of tasks is observed. Then at the NFL draft, the 32 NFL franchises take turns in picking the candidates with their wages following the order of the pick.[1] After seven rounds of picking players, the remaining college player pool is considered undrafted.

Ideally, gender, age or skin color[2] should play no role in the hiring of candidates, but there is evidence from the economics literature to suggest that this isn't the case (Bertrand and Mullainathan, 2004; Goldin and Rouse, 2000; Neumark, 2020). Either because of prejudice (Becker, 1957), or 'statistical' processes that break along ageist, racial or gendered lines (Arrow, 1972; Phelps, 1972) it is possible that the non-productivity related characteristics of the candidate impacts hiring. We argue that identifying racial bias in the hiring process is important not just because we want hiring processes to be fair. If there is bias in the hiring process, a regression of the observed wage gaps between workers of different races may underestimate the full extent of discrimination in the labor market. However, outside of professional sports, identifying racial bias in hiring is difficult (Parsons et al., 2011). It is usually not possible for researchers to observe an accurate signal

✉ Ian Gregory-Smith
ian.gregory-smith@newcastle.ac.uk

1 Economics, Newcastle University Business School, Newcastle upon Tyne NE1 4SE, UK

2 Social Research Institute, University College London, London, UK

3 Centre for Industrial Relations and Human Resources, University of Toronto, Toronto, Canada

[1] Please refer to the Appendix for details.

[2] We use the term 'skin colour' in this paper to describe differences in NFL rankings and draft position between Black and White players in the NFL. This journal, and the economics literature at large, uses the term 'race' and the variables we define in this paper can be thought of as racial categories. But note that 'race' is to be understood here as a social construction (there are no such things as 'biological races' (Sussman, 2014)). Put simply, we observe a player's skin colour and assign them to one of three categories:'White', 'Black' or 'Non Black Non White'.

Published online: 06 July 2023

of a candidate's potential productivity. As such, it is usually not possible to differentiate between competing interpretations for observed differences between groups. Additionally, acquiring information on unsuccessful candidates is necessary to measure the success rates accurately (Neumark and Rich, 2019).

The contribution of this paper is testing for racial discrimination in the rank-ordering of candidates by building a unique dataset that includes both successful and unsuccessful applicants. This paper merges information from a number of publicly available sources in order to track college players through the NFL Combine and NFL Draft and into their first contract as a professional football player for an NFL franchise. By observing publicly available images of players, we manually assign players to one of three categories: 'Black', 'White' or 'Non Black Non White'. We observe large unconditional differences in drafting based on these groupings. Much of the variation can be explained by Black and White players selecting into different playing positions. Additionally, after controlling for a full set of productivity measures we do not find robust evidence of racial discrimination against Black players in drafting. However, we do find that some aspects of the data that could be a concern for non-White participants.

For Black participants, the first issue is that Black quarterbacks are overlooked in favor of White quarterbacks with the same physical characteristics, albeit this difference is small and only marginally statistically significant. It is only when the measure of cognitive ability, known as the Wonderlic test[3] is included, that Black quarterbacks achieve parity. In other words, it is because Black quarterbacks under perform on this test relative to White quarterbacks, that White quarterbacks have greater success on average in the draft. The average performances on this test mirror those found in the broader US population. Nevertheless, the question remains whether this is a unbiased test of quarterback playing ability (Lyons et al., 2009) and indeed the NFL will no longer administer this test from 2022 onwards (See https://www.nytimes.com/2022/03/02/sports/football/nfl-wonderlic-test.html). A second concern is that we do find evidence that Black players are disadvantaged relative to White players in later rounds of the draft. This is consistent with the notion that it is less costly to discriminate against such players because the forgone talent opportunity is lower at the bottom end of the draft.

Another concern is that players who are neither Black or White, typically players with Hispanic, Pacific Islands, or Asian backgrounds are less successful in the draft than would be expected given their observed characteristics, at all positions, including the quarterback position. We should stress that the statistical power of the result on quarterbacks is low because there are only eleven Non Black Non White quarterbacks in the 18 years of our sample. We can't say if this result would hold had their been more observations. However, the very fact that there are so few quarterbacks from these groups in the sample is perhaps an indication that such players face barriers to entering the NFL.[4]

Background

Economists usually begin from the position that hiring is a hidden type problem. Ex ante, candidates have private information over their own future productivity that is hidden from employers. Even after hiring, since observed productivity typically contains some noise, it may take several time periods before employers are confidently able to identify and remove low productivity workers. This would be costly for the employer. Therefore, it is profitable for employers if they can obtain information to improve their ranking of candidates. Interviews and assessment days therefore represent relatively low cost mechanisms to obtain a signal of future productivity. In our context, the signal is obtained through the physical tests at the NFL Combine along with playing statistics in college. This signal can be combined with other information about the candidate to arrive at a rank order that maximises future productivity.

However, the other information may include demographic averages, as it does in models of 'statistical' discrimination (Arrow, 1972; Phelps, 1972). In such models, it is assumed that the employer knows the distribution of productivity for each group (by age, gender, race, etc) and they can use this information to form a more accurate prediction of expected productivity than would be obtained if they used the individual level productivity signal alone. In short, if the observed average level productivity of one demographic group is higher than the other, candidates from that group will be ranked higher given equal individual productivity signals.[5]

A recent theoretical development in a similar spirit to statistical discrimination are models of biased beliefs in hiring (Bohren et al., 2019; Schwartzstein, 2014). The key difference to the traditional statistical discrimination model is that

[3] Potential NFL players since 1970 have had administered an intelligence measurement called the Wonderlic Personnel Test at the NFL Scouting Combine. The test is used to measure players' aptitude for learning and problem solving. The possible score range is 1 to 50. The average football player scores around 20 points and scores vary by position. See: https://fivethirtyeight.com/features/how-a-multiple-choice-test-became-a-fixture-of-the-nfl-draft/

[4] Low participation from Non Black Non White groups could also reflect different rates of self selection if minority groups have different sporting tastes. Nevertheless, this does not rule out discrimination because self selection rates may be influenced by anticipated discrimination.

[5] Other moments of the productivity distribution might also be relevant. For example, a risk averse employer may prefer candidates from a group with lower variance, all else equal.

the employer's understanding of the productivity distribution by race may be incorrect. Due to stereotyping or historical performances, a franchise may believe, for example, that White players outperform Black players at quarterback and make hiring decisions accordingly. However, this assessment of the productivity distribution may be outdated and should be subject to revision as new information on player performances emerges. Franchise selectors who hold on to outdated stereotypes may fail to update their assessment of the productivity distribution, even when new statistical information becomes available.

There is qualitative evidence consistent with stereotyping. In Dufur and Feinberg (2009) the authors conducted interviews with players invited to the NFL combine. While they found no evidence of overt discrimination, players from ethnic minorities reported practices that were consistent with racial stereotyping. For example, Black and Polynesian players were questioned about family relationships and off-field (potentially criminal) behavior to a greater extent than White players. In Mercurio and Filak (2010) a text analysis was conducted on the media commentary of Black and White college quarterbacks. A large difference was notable in how Black and White quarterbacks were described. Mental strength and decision making were associated with White quarterbacks while Black quarterbacks were more frequently described in terms of their physical attributes.

Alternatively, (or in addition) to statistical discrimination, other forms of discrimination could be at play. The most egregious mechanism is that of simple employer prejudice known as 'taste-based' discrimination (Becker, 1957). In these models, employers are willing to pay for their prejudice in the form of higher wages for White employees. If White and Black employee wages are equalised (for example due to a collective bargaining agreement), then fewer Black employees will be hired. With wages tightly following the draft order, we would expect taste discrimination to be seen in the rank order of the draft, or the selection of marginal White draftees over Black draftees. This was certainly the case in the past, given the known cases of historical racism in American football. For example, from 1869 to 1945, Black players were effectively excluded from playing,[6] with only a handful of exceptions (Levy, 2003; Smith, 1988).

In a perfectly competitive market, such discriminating employers can not survive the forces of competition, but in the NFL, there are significant restrictions on competition that may allow prejudiced employers to exercise their taste for discrimination. In the name of 'competitive balancing', there is revenue sharing between franchises, restrictions on 'rookie' contracts, a hard salary cap, and a reverse order of finish college draft. These competitive balancing mechanisms have been the source of much academic discussion (Késenne, 2014; Szymanski, 2006) but a further unintended consequence of them could be that they permit employer discrimination.

A subtle variation on taste-based discrimination locates the source of the prejudice with the employer's customer base, in this case football fans. Put simply, a majority White fan base may be willing to pay more to see a White quarterback, as the 'face of their franchise'. A profit maximising NFL franchise might be expected to reflect their customers' preferences in the NFL draft, even if those preferences are prejudiced. There is evidence from Kahn (1992) that White and Black player wages vary according to the racial demographic of the metropolitan area from where the fans are drawn. A related mechanism is that fans may reduce the productivity of players from minority groups through racist intimidation. A recent paper by Caselli et al. (2022) uses the absence of fans in association football due to Covid restrictions to estimate the reduction in productivity of players with African heritage caused by racist fans. Again, a profit maximising franchise would be expected to factor in the lower productivity arising from fan abuse, even if the franchise selectors themselves are not prejudiced.

A final concern is that players may face discrimination from other players, known as 'employee discrimination'. Given the potential for productivity spillovers between players (e.g. a good receiver will raise the productivity of a quarterback), the possibility of employee discrimination lowers the expected productivity of a player from a minority group. However, we argue that in our context, while not impossible, this is unlikely to be widespread because by the time players enter the NFL they have been playing in racially diverse teams for several years. Additionally, to our knowledge, we don't have any anecdotal evidence of this having occurred during our sample period.

Empirical Literature

In testing for discrimination in the NFL, the literature has, by and large, taken the traditional regression based approach of Mincer (1970). Essentially, a wage equation is specified with a dummy variable identifying race and other variables controlling for other key determinants of wages such as human capital and prior productivity.[7] When applying this approach to NFL data, it becomes apparent that the race variable is

[6] The first American Football game is said to have occurred on 6th November 1869. The professionalisation of the game begins in 1892 and the American Professional Football Association is founded in 1920 which becomes the NFL in 1922.

[7] As noted by Simmons (2021) in a recent review, the literature has generally not applied the Oaxaca-Blinder decomposition because of the insufficient number of observations of players from the minority group to estimate two separate equations.

typically insignificantly different from zero. That is to say, at the mean, there is no unexplained wage gap between White and Black NFL players (Berri and Simmons, 2009; Burnett and Scyoc, 2015; Ducking et al., 2014).

Nevertheless, the race identifier can be interacted with performance statistics to reveal differences in how productivity between the races is rewarded in the labor market. In particular, Berri and Simmons (2009) find that rushing yards by Black quarterbacks are not compensated in their sample. The authors also find some evidence of wage gaps at points of the distribution away from the mean, particularly at the top end of the wage distribution. This result is also seen in Keefer (2013), who identifies a wage gap for linebackers at all points of the distribution but particularly at the top and bottom deciles. However, the effect does appear to be confined to certain positions. Across a range of different NFL positions, Ducking et al. (2014) do not find wage gaps when adopting a quantile approach.

Our test of discrimination takes a different approach to the literature above. Rather than examine the observed wage distribution, we want to focus on the rank ordering of candidates in the NFL draft, those players having come through college football and the NFL Combine selection event. Since player wages during the first four years of the playing career, (their 'rookie' contract) are very tightly associated with their draft position, we argue that the relevant point of assessment is the draft. Further, a significant number of NFL players do not survive in the league beyond their rookie contract (Volz, 2017).

While we bring the largest sample in the literature to date, our paper is not the first to study the NFL draft. Hendricks et al. (2003) examine statistical discrimination in the draft with respect to players from less visible college programs (non Division 1A). If these players are drafted early, they tend to have better careers than equivalent players from high profile schools. This implies that players from low profile schools have to be better than their equivalents to draft early, reflecting statistical discrimination. However, in later rounds the effect is reversed with Division 1A players ranking lower than equivalent non Division 1A players. The authors interpret this as support for an 'option value' effect. Because a franchise can eliminate poor performers ex post, it becomes profitable for franchises to take a risk on a less visible player in the hope that they uncover a hidden 'star'.

Massey and Thaler (2013) present evidence that franchises overvalue top picks in the draft. Evidence for this claim is based on the difference between the compensation cost of the pick relative to what an equivalent player would cost in free agency. As noted by Berri and Simmons (2011) this approach requires the strong assumption that the draft itself does not impact future pay. Nevertheless, the broader claim that the draft is an imprecise process in terms of correctly identifying future talent is consistent with Berri and Simons (2011), particularly at the quarterback position. While drafting early is predictive of playing time, conditional upon playing time, early picks do not always outperform later ones. Further, the variables that are predictive of where a player will draft do not translate to playing success. Wolfson et al. (2011) revisit this question and although they agree that the process is imprecise and that variables predicting draft and combine success are not strongly predictive of playing performance, they offer a different interpretation to Berri and Simmons (2011). The authors argue that franchises aggregate all the pre-draft information about a player, including non-quantitative information such as scouting reports. This results in the draft position containing significant information about playing suitability, which is then reflected in more game time for the players.

Berri and Simmons (2011) also do not find significant differences in draft rank order between Black and White quarterbacks. Gill and Brajer (2012) continue this investigation in more detail by examining quarterbacks in the draft and focusing on the measure of cognitive performance at the Combine, known as the Wonderlic test. If franchises were treating the Wonderlic signal of cognitive ability as a stronger signal for White quarterbacks than for Black quarterbacks, this could be interpreted as evidence of statistical discrimination. However, the authors find that this is not the case but rather the marginal effect of the Wonderlic for Black and White quarterbacks is equal. In contrast, Conlin and Emerson (2005) do find evidence of discrimination in the draft by separating the drafting decision from the playing decision. While drafting early greatly improves the chances of playing in the NFL, there is no guarantee and players can be dropped before the start of the season. Conditional upon drafting position (and other variables), non-White players are shown to be more likely to play and start games than White players. The authors interpret this result as evidence of discrimination in the draft against non-White players by assuming the decision to play is race neutral. If this assumption holds, coaches who pick the playing team are correcting the drafting bias from the front office, who should have been drafting the playing non-White players earlier in the draft.

Our contribution to this literature is to provide a more complete picture of the role of race in the NFL Draft. We bring together information from several sources to form the largest dataset to date with information on both successful and unsuccessful applicants. In doing so, we emphasise the importance of playing position and the sorting of players by race into different playing positions.

Table 1 Drafting by Position

	All	QB	Oline	Dline	Dcover	LB	RB	WR	Sp
N	6,247	379	1,349	977	1,100	703	658	856	225
N Drafted	3,895	203	834	672	742	489	383	512	60
Black									
N	4,206	90	554	740	1,006	519	565	730	2
N Drafted	2,770	44	367	516	684	365	334	459	1
White									
N	1,669	278	682	150	55	147	60	85	212
N Drafted	962	157	419	109	39	105	35	41	57
Non Black Non White									
N	361	11	109	85	38	36	33	39	10
N Drafted	163	2	48	47	19	19	14	12	2
% of drafted are:									
Black	71.12%	21.67%	44.00%	76.79%	92.18%	74.64%	87.21%	89.65%	1.67%
White	24.70%	77.34%	50.24%	16.22%	5.26%	21.47%	9.14%	8.01%	95.00%
Difference	46.4pp	-55.7pp	-6.2pp	60.6pp	86.9pp	53.2pp	78.1pp	81.6pp	-93.3pp
SE	0.01	0.08	0.03	0.01	0.00	0.01	0.01	0.01	0.18

1. Sample is NFL Draft Combine participants from 2000-2018. Eleven of the 6,247 participants could not be confidently classified into either of the three categories
2. *Difference* is the percentage point difference in the proportion of drafted players who are Black vs the proportion of drafted players who are White. SE is the standard error of this proportion. All differences are statistically significant at conventional levels
3. Column headings denote playing positions. Full description provided in Table 8 in the appendix

Data

Our dataset is an individual level dataset comprising the full population of Combine participants from the year 2000 to 2018. We create three player categories, Black (N=4,206), White (N=1,669) and Non Black Non White (N=361) by observing the images of each player which are in the public domain.[8]

Participation at the Combine is not strictly mandatory and it is possible that some players opt out. Fortunately, these are rare cases. We observe 37/4857 players (0.76%) who played at least one game in the NFL without recording Combine statistics (14 Black, 23 White). 17 of these players were specialist players (kickers or punters) for whom the Combine tests are less relevant. A description of each playing position outlined in Table 1 is provided in the Appendix.

In Table 1, there are 6,247 participants of which 3,895 (62.3%) are successful in the draft (were drafted in one of the seven rounds). Success rates vary between position reflecting the variation in the supply and demand for positions. 54% of quarterbacks draft, while as many as 69% of defensive linesmen draft and as few as 27% of specialists (Kickers and Punters) draft. The majority of draftees are Black but there is large variation between positions. There are more White draftees than Black draftees at the position of quarterback, offensive linesmen (which includes the tight end position) and specialists. In all other positions there are more Black draftees.

However, we can not infer from these numbers alone whether or not Black and White players are more favored in the draft at different positions. This is because the number of Black and White participation also varies dramatically across positions. In all the positions where there are more successful Black (White) draftees, there are more unsuccessful Black (White) participants. For example, although $57/60 = 95\%$ successfully drafted specialist players are White, the vast majority of unsuccessful specialist players are also White ($155/165 = 94\%$). Whatever is causing the vast majority of Kickers and Punters to be White, it is happening prior to participation in the Combine.[9] Conditional upon participation

[8] Classification of players into the three categories was primarily conducted by research assistants Michael Muir and Justin Zelnicker, with oversight from the authors. A random sample from each data inputer was analysed and no statistically significant differences in classification frequency were present. There were 11 players who we could not confidently classify into either of the three categories and these were dropped from the analysis.

[9] This is not to say the draft, or decisions after the draft in the NFL are necessarily colour blind at this position. It is quite possible that anticipation of unequal treatment at this position gives rise to unequal selection into this position. What is being claimed here is that this process, or whatever else drives the unequal selection is taking effect prior to the draft.

Table 2 Black/White Differences in Drafting Success Rates

	All	QB	Oline	Dline	Dcover	LB	RB	WR	Sp
Drafted B vs W									
(a) Difference	46.4pp	-55.7pp	-6.2pp	60.6pp	86.9pp	53.2pp	78.1pp	81.6pp	-93.3pp
SE	0.01	0.08	0.03	0.01	0.00	0.01	0.01	0.01	0.18
Undrafted B vs W									
(b) Difference	31.0pp	-42.6pp	-14.8pp	60.0pp	85.5pp	52.3pp	74.9pp	66.0pp	-93.3pp
SE	0.01	0.08	0.04	0.02	0.01	0.02	0.01	0.01	0.11
(a) - (b)	**15.42pp**	-13.05pp	8.52pp	0.57pp	**1.45pp**	0.83pp	**3.16pp**	**15.65pp**	0.00pp
SE	0.01	0.11	0.04	0.02	0.01	0.03	0.01	0.01	0.21

1. Sample is NFL Draft combine participants from 2000-2018
2. (a) - (b) denotes the difference in the Black-White success-failure rate. (a) - (b) controls for the unequal numbers of Black and White participants at each position. A positive (negative) number indicates Black participants face a higher (lower) drafting success ratio on average than White participants, irrespective of the unequal number of Black and White participants at each position
3. Statistical significance at 5% is highlighted in **bold**

in the Combine, White and Black players are equally likely to be successful at this position.

Nevertheless, the possibility of unequal drafting success rates remain at other positions. This is the focus of Table 2. Overall, of those drafted, the difference in Black-White percentage points (pp) is 46.4pp higher for Black players. The equivalent statistic in the undrafted pool is 31.0 percentage points higher for Black players. Therefore, there is a statistically significant difference in the sucess-failure rate of 15.42pp in favor of Black players. Altogether, Black players are more likely to draft, even controlling for the greater number of Black participants at the Combine.

Turning attention to specific positions (reading along the (a)-(b) row), Black players are favored at defensive line, defensive cover, linebacker, running back and wide receiver. The only position where White players are favored is the position of quarterback, to the tune of 13.05 percentage points. However, this number falls short of statistical significance because of the low number of Black participants at the quarterback position.

The differences outlined in Table 2 are unconditional differences in drafting success rates between Black and White players. In common parlance, these can be thought of as race *gaps* in drafting success, in the same sense that the gender wage gap is the unconditional mean difference in wages between two groups, men and women. As with wage gaps, our race gap in drafting success by position does not imply discrimination in the draft. This is because after conditioning on other variables that determine the drafting success, it may be that there are no outstanding gaps to explain. Equally, the absence of a race gap does not imply the absence of racial discrimination; gaps may emerge after controlling for other determinants of drafting success.

Yet, that is not to say the presence of a race gap in drafting is uninformative on the matter of discrimination in the draft. Rather it provides the context to understanding the avenues through which discrimination may be occurring. In our case, the presence of a positive race gap for Black players overall and at the specific positions of Offensive Line, Defensive Cover, Running Back and Wide Receiver narrows the range of possible inferences. First, prejudice becomes less likely because pure malice should be independent of the position a player occupies. If a taste-based argument is to be maintained, it is more likely to be the stereotyping variant, whereby franchises fail to correctly update their beliefs of a player's expected productivity, given their position and race.

Second, to hold to the position of no discrimination, it must be that Black participants in the Combine at these positions are drawn from a better distribution of potential Combine invitees than White participants. This might be because the distribution of Black players is simply better than White players in the population at these positions, in which case the no discrimination claim may be plausible. Or Black college players at these positions may have faced a tougher selection process to be invited to the Combine, which could imply racial discrimination prior to the Draft. Black players on average may be less well known and / or attend colleges that do not promote as aggressively to Combine selectors. Similarly, the unconditional advantage White quarterbacks have in the draft could be justified if it can be shown that observable characteristics that make a successful quarterback are more frequently observed among White combine participants. Alternatively, it could go the other way. It may be that after observing predictors of quarterback success, Black players are even more disadvantaged than implied by the unconditional differences. What is required is a model of the process by which players are drafted in order to understand why drafting success gaps may be observed in the data. We now turn our attention to this matter.

Empirical Strategy

Probit

The starting point for our empirical specification is to model the extensive margin of being drafted or not drafted. We begin with a binary Probit:

$$p_i = Pr[y_i = 1|X] = \Phi(X_i'\beta) \tag{1}$$

where p_i is the probability that NFL Combine participant i is drafted conditional upon a vector X of explanatory variables, Φ is the usual cumulative function for the standard normal distribution and β is the vector of estimated coefficients. In all tables, we report average marginal effects.

Our two main explanatory variables of interest are the indicator variables 'Black' and 'Non Black Non White' which identify the marginal effect of race on drafting success relative to White players. Central to the quality of our estimating Eq. (1) is whether or not the vector X captures the draft selection process. To this end, we split our analysis across three key signals of future playing success. First, the 'physical signal' given by the recorded performances on the physical tests at the NFL Combine. Second, with respect to the quarterback position only, the cognitive signal given by performance on the Wonderlic test. Third, the signal of future productivity, generated by playing statistics during their college playing career years.

We also control for other variables that may be reasonably determining whether or not a Combine participant is successful in the draft.[10] In particular, we have a set of year fixed effects and position fixed effects. We argue controlling for position is particularly important since the descriptive evidence from Tables 1 and 2 could suggest that being Black is an advantage overall, beneficial in some positions but neutral or detrimental in others. Note that by controlling for position, we are limiting the potential mechanisms of racial discrimination to operate only in the draft event, with player position already given. The investigation of discriminatory mechanisms that occur prior to the draft and impact player selection into those positions is beyond the scope of this study. Creating the final dataset for the analysis required merging data from several sources.

Tobit

While the extensive margin of drafting could be considered to be the fundamental discontinuity that will shape the likelihood of a successful career in the NFL, there are also substantial differences between drafting high and low in the draft. The number 1 draft pick in 2021 signed a contract worth a total of $37.7M, whereas a draftee in round seven signed a contract with a total value of approximately $3.5M. Therefore, it is important to consider the determinants of the intensive margin in terms of how high a player drafts.

The drafting position of a player is an ordinal rank, running from first to last, after seven rounds of 32 franchise picks.[11] For the purposes of capturing the intensive margin of this process, we approximate the ordinal rank of the draft with a continuous variable of drafting position. Because this variable is censored, we use a Tobit:

$$y_i^* = (X_i'\beta) + \epsilon_i \tag{2}$$

$$y_i = \begin{cases} y^* \ if \ 1 \ \leq y^* \leq 261 \\ 1 \ if \ y^* \leq 1 \\ 261 \ if \ y^* \geq 261 \end{cases}$$

The variable y_i^* is the unobserved latent variable and we observe y_i as the draft position. This variable is censored from above at 1, because the best a player can draft is in first place, even if this player was many times better than the rest of the field. The variable is also censored from below, because if a player does not draft, we only observe that they are ranked one position worse than the final pick in the draft.

A variant of the Tobit described above is to consider the within position ranking of players. As shown in Table 1, there is variation between positions in drafting success. For example, the quarterback position is the most critical position on the field, contributing the most to the outcome of a game. Highly rated quarterback prospects are likely to outrank highly rated players from other positions. At the other extreme, there have only been three punters who have drafted in the 1st round, the highest being 17th position. Additionally, when making rank choices, franchises do not only consider the quality of the prospect but gaps in their squad that need filling. Therefore, franchises will create their own internal ranking of draft prospects within position.

If candidates with a particular skin color are being favored in the draft, we can expect this to be most acutely felt within position. To allow for this, we can adjust the dependent variable according to the within position rank of each player in each year of the draft. This also allows the explanatory variables measuring the productivity signal received by the franchise to be more accurately tailored to the particular position of interest. For example, measures of passing performance (passing yards, completions, passing touchdowns, interceptions etc) are only relevant for players who have thrown in college.

[10] One limitation is that we do not know the age of each participant. However, the scope for variation by age is small because the majority of participants are final year college students aged 22 or 23.

[11] There is variation between years in the number of franchise picks.

Table 3 Extensive Margin of Drafting Success: Probit

	(1)	(2)	(3)	(4)	(5)
Black	0.0810***	0.0175	−0.0137	0.0697***	−0.0191
	(5.924)	(1.058)	(−0.843)	(3.038)	(−1.193)
Non Black Non White	−0.118***	−0.174***	−0.165***	−0.101***	−0.164***
	(−4.328)	(−6.364)	(−6.174)	(−2.825)	(−6.237)
Physical Signal					
Height			−0.0125	−0.0203	−0.00794
			(−1.275)	(−1.379)	(−0.828)
Weight			0.320***	0.351***	0.328***
			(17.09)	(11.09)	(17.64)
Speed			0.320***	0.224***	0.319***
			(25.73)	(9.88)	(26.00)
Vertical				0.0353**	
				(2.123)	
Bench				0.0385***	
				(3.175)	
Broad jump				−0.00548	
				(−0.283)	
Cone				0.0599***	
				(3.001)	
Shuttle				0.0752***	
				(3.926)	
Productivity Signal					
College Rank					−0.000734***
					(−5.379)
Major School					0.268***
					(10.54)
College Performance					0.115***
					(6.321)
Year FE	No	Yes	Yes	Yes	Yes
Position FE	No	Yes	Yes	Yes	Yes
Pseudo R^2	0.01	0.04	0.11	0.13	0.13
Observations	6,236	6,236	6,069	2,910	6,069

1. The sample is NFL Draft combine participants from 2000-2018. Reported estimates are average marginal effects following a Probit. T-statistics of the average marginal effects in parentheses
2. The variables capturing the physical signal are standardised so that higher values are associated with better performance. College Rank is the within sample historical ranking of the College (lower numbers associated with a stronger college football team). Major School indicates the participant played in a FBS-equivalent conference, or their college played the majority of their games against other rated schools. College Performance is the player's rank in their position according to the playing statistics at College that are relevant to their position

Identification

We are adopting a regression based approach, with the intention that we adequately capture the process by which players are drafted. Regression based approaches are notoriously vulnerable to identification issues. In this case, our main variable of interest is a player's skin colour (which is not a choice), so we are free of many of the usual problems associated with an endogenous explanatory variable.

Nevertheless, we should be clear on the following. First, if there are unobserved variables correlated with skin color that also impact drafting, these will impact the reported marginal effects on race. While we are able to control for a rich set of of signals that likely determine drafting success, it is possible that there is information available to franchise selectors but unavailable to us. For example, off-field disciplinary offences may be unequally distributed between Black and White participants. To the extent that such information is correlated

Table 4 Intensive Margin of Drafting Success: Tobit

	(1)	(2)	(3)	(4)	(5)
Black	−31.73***	−17.26***	−7.207	−26.80***	−5.239
	(−7.885)	(−3.739)	(−1.671)	(−4.594)	(−1.245)
Non Black Non White	26.89***	40.79***	36.16***	16.02*	35.13***
	(2.961)	(4.510)	(4.287)	(1.491)	(4.264)
Physical Signal					
Height			1.922	8.159**	0.448
			(0.703)	(2.071)	(0.168)
Weight			−114.2***	−125.9***	−116.1***
			(−22.03)	(−15.34)	(−22.79)
Speed			−113.5***	−75.34***	−112.1***
			(−30.45)	(−12.40)	(−30.94)
Vertical				−10.02**	
				(−2.310)	
Bench				−10.97***	
				(−3.630)	
Broad Jump				−6.428	
				(−1.286)	
Cone				−22.20***	
				(−4.178)	
Shuttle				−18.16***	
				(−3.559)	
Productivity Signal					
College Rank					0.358***
					(8.959)
Major School					−97.56***
					(−11.57)
College Performance					−37.46***
					(−7.738)
Year FE	No	Yes	Yes	Yes	Yes
Position FE	No	Yes	Yes	Yes	Yes
Pseudo R^2	0.002	0.01	0.02	0.03	0.03
Observations	6,236	6,236	6,069	2,910	6,069

1. The sample is NFL Draft combine participants from 2000-2018. Reported estimates are marginal effects following a Tobit. The dependent runs from 1 through 261, with 1 indicating the highest ranked draftee. Therefore the negative estimates for the variable Black imply an advantage for Black participants. T-statistics in parentheses

with skin color and impacts drafting success, our estimates will be contaminated by this information.

Second, we are making the assumption that the observed signals of future productivity are themselves independent of racial bias. If Franchises are using racially biased signals, the absence of an effect on the skin color identifier itself, would not imply a color blind process. In particular, the Wonderlic test of cognitive ability given to participants at the Combine is controversial, (Gill and Brajer, 2012; Jensen, 1977). Lyons et al. (2009) find that Wonderlic scores do not predict future NFL performance and the NFL has decided no longer to administer the test from 2022 onwards.

Third, in the specifications presented below, we do not allow the signal to vary in its impact by skin color. For example, it is possible that the marginal impact of a quick 40 s dash time is different for Black and White players. This third point is easily addressed with interaction terms. We do this as a robustness exercise and find no differences of note.

Results

Results using the full sample of NFL combine participants from all positions are given in Table 3. Column (1) shows

Table 5 Quarterback Drafting Success

	Drafted: Yes or No			Draft Position		
	(1)	(2)	(3)	(4)	(5)	(6)
Black	−0.0747	−0.109**	0.0286	20.38	44.38**	14.36
	(−1.261)	(−1.985)	(0.464)	(0.933)	(2.294)	(0.736)
Non Black Non White	−0.419***	−0.287**	−0.326**	161.7**	119.5**	140.4*
	(−2.449)	(−2.155)	(−2.473)	(2.354)	(2.065)	(1.915)
Wonderlic			0.00915**			−3.038**
			(2.441)			(−2.393)
Physical Signal						
Height		0.214***	0.206***		−81.77***	−65.64***
		(5.619)	(4.597)		(−5.949)	(−4.595)
Weight		0.189**	0.170		−103.1***	−65.64**
		(1.854)	(1.649)		(−2.969)	(−2.301)
Speed		0.194***	0.182***		−81.74***	−69.93***
		(4.624)	(3.698)		(−5.278)	(−4.313)
Productivity Signal						
College Rank		−0.00095***	−0.00145***		0.519***	0.605***
		(−2.396)	(−3.893)		(3.509)	(3.940)
College Performance		0.460***	0.246**		−171.4***	−79.51**
		(6.872)	(2.438)		(−6.483)	(−2.521)
Heisman Finalist		0.124*	0.150*		−69.06***	−67.17***
		(1.790)	(2.016)		(−3.358)	(−3.407)
Year FE	No	Yes	Yes	No	Yes	Yes
Pseudo R^2	0.01	0.30	0.36	0.002	0.07	0.07
Observations	379	373	243	379	373	243

1. The sample is NFL Draft combine quarterbacks from 2000-2018. Columns (1)-(3) report average marginal effects following a Probit. Columns (4)-(6) report marginal effects following a Tobit. T-statistics in parentheses
2. Wonderlic is a cognitive ability and problem-solving aptitude test. It is administered at the Combine, with the majority of participants being quarterbacks. A Heisman Finalist is an indicator of strong college prospect voted for by sports journalists and previous Heisman trophy winners

the unconditional gap in drafting between Black and White participants, with Black participants being 8.1 percentage points more likely to draft. This mirrors the raw differences in drafting success in Table 1 above. Columns (2) through (5) add controls for position and year fixed effects, the Physical Signal from the Combine tests and the Productivity Signal from the performances at College. As observable controls are added, the estimated average marginal effect declines and is not statistically significant with the full set of controls in column (5). This suggests that the better observable signals (on average) by Black participants explain the majority of the unconditional gap in drafting likelihood.

The estimates reported in column (1) for 'Non Black Non White' indicate an unconditional gap in drafting relative to White participants of 11.8 percentage points. Recall, this category is much smaller including only 361 from 6247 participants and comprises players with Hispanic, Pacific Island or Asian backgrounds. A penalty for these players of 16.4 percentage points emerges with the most complete set of controls included. In other words, White participants with similar observable characteristics are no less likely to draft overall than Black participants but more likely than players who are neither Black or White.

The control variables behave as expected. The set of year and position fixed effects are jointly statistically significant and impact the size of the estimated marginal effect (column (2)). Franchise selectors use the Physical signal from the Combine (columns (3) and (4)) as well as the Productivity signal (column (5)) from College football. A little care is required in reading individual estimates from the table. For example, heavier players are favored in the draft but this holds Height and Speed constant. Together, these variables indicate a more powerful athlete.

Estimates on the intensive margin from Table 4, present a similar qualitative picture to those given on the extensive margin in Table 3. Again, the unconditional gap in drafting in favor of Black participants is large, to the tune of 33 places (a full round earlier). Note here that negative coefficients imply drafting earlier or higher in draft. This is diminished to 5 places with the full set of observable controls and falls short

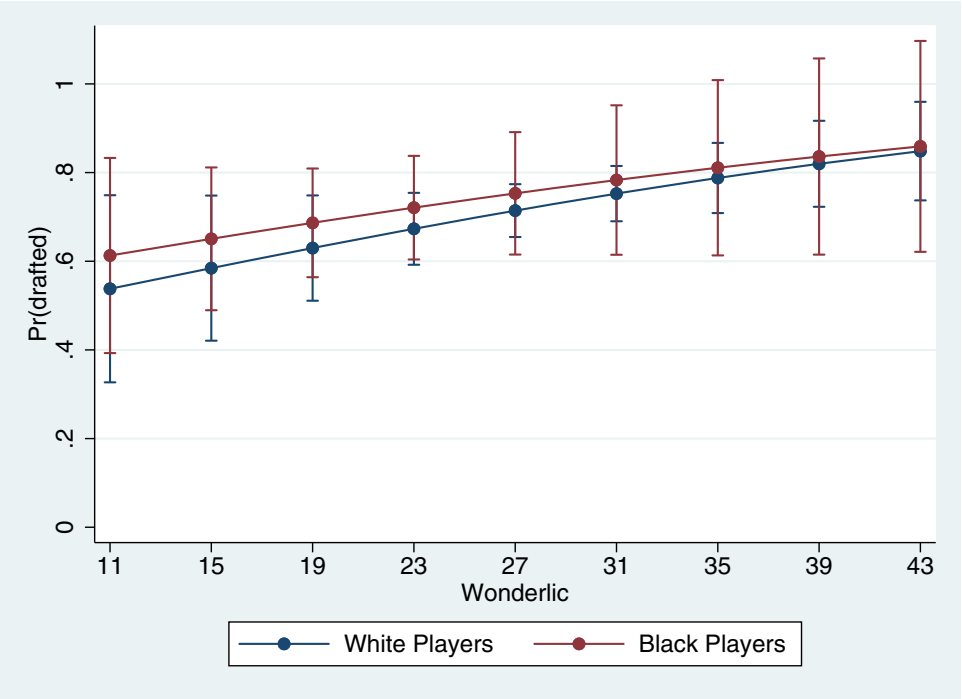

Fig. 1 Impact of Wonderlic on Quarterback Drafting Probability. The Wonderlic scale runs from zero to 50. A score of 11 represents the 1st percentile and 43 represents the 99th percentile. Overlapping confidence intervals for Black and White quarterbacks suggest that there are no statistially significant differences in how Wonderlic impacts drafting success

of statistical significance. The estimates with respect to Non Black and Non White players are consistent with those of Table 3. With the full set of controls, Non Black and Non White players are drafted 37 places later down the order and this result is statistically significant. Altogether, Black players are significantly favored by franchise selectors in the draft but for reasons that are explained by observed characteristics. Whereas, Non Black Non White players draft later for reasons that are *not* explained by observed characteristics.

Quarterbacks

As shown above, Black candidates from the Combine are favored in terms of drafting success and drafting position. However, as shown in Tables 1 and 2, there is significant variation by position. We are particularly interested in the quarterback position for three reasons. First, we have strong historical evidence of racism specific to this position as well as anecdotal contemporary evidence suggesting that stereotypes surrounding Black quarterbacks continue to carry weight. Second, the quarterback is the highest paid and most influential position on the team. Highly rated quarterback prospects typically command the highest positions in the draft. With the distribution of wages being tightly correlated with draft position, but in a non-linear manner (Keefer, 2013), if there is discrimination at the quarterback position against Black candidates it could offset any average wage advantage Black players in other positions experience. Third, we have additional information on cognitive performance through the Wonderlic test that is typically given to quarterback prospects. For these reasons, we focus in on the determinants of draft success for quarterbacks separately.

Columns (1) through (3) of Table 5 report the average marginal effects following a binary probit of whether the player drafted or not, while columns (4) through (6) report the results of a Tobit for drafting position. There is no robust evidence that Black players are disadvantaged. A small gap emerges when controlling for the Physical and Productivity signals pertinent to quarterback drafting. However, when additionally controlling for the Wonderlic signal of cognitive ability available to Franchise selectors, there is no unexplained gap for Black quarterbacks, either in drafting or in terms of where a quarterback drafts. This implies, and indeed it is the case, that Black quarterbacks under perform White quarterbacks in the Wonderlic (by approximately 1 standard deviation in our sample).[12]

This raises the question whether or not the Wonderlic is racially biased against black participants. Although, the answer to that question is beyond the scope of the study here, we can answer the claim as to whether or not Franchise selectors use the Wonderlic differently for Black and White quarterbacks. Figure 1 presents the distribution from the 1st percentile to the 99th percentile of Wonderlic performance and its impact on drafting sucess. While higher scores on the Wonderlic increase drafting likelihood, and as stated White

[12] This is a consistent finding seen as early as 1977 in the literature (Jensen, 1977; Gill and Brajer, 2012). Wonderlic information is not universally available. In total, we have Wonderlic information for 247/379 = 65% of QB observations and 340/5,868 = 5.8% for non-QB observations. For this reason, we only use the Wonderlic in the case of QBs.

Table 6 Within Position-year Rank

	All Positions		QB
	(1)	(2)	(3)
Black	−2.377***	−1.032*	−0.133
	(−3.885)	(−1.781)	(−0.156)
Non Black Non White	6.522***	3.994***	5.554*
	(4.887)	(3.579)	(1.907)
Wonderlic			−0.130***
			(−2.389)
Physical Signal			
Height		−0.003	−3.103***
		(−0.009)	(−4.954)
Weight		−15.51***	−3.002
		(−21.50)	(−1.932)
Speed		−14.90***	−3.008***
		(−29.27)	(−4.371)
Productivity Signal			
College Rank		0.0436***	0.0259***
		(8.104)	(3.831)
Majorschool		−12.12***	
		(−10.66)	
College Performance		−4.275***	−3.285**
		(−6.584)	(−2.339)
Heisman Finalist		−12.28***	−2.904**
		(−11.31)	(−3.480)
Year FE	No	Yes	Yes
Position FE	No	Yes	N/A
Pseudo R^2	0.002	0.05	0.12
Observations	6,236	6,069	243

1. The sample is NFL Draft combine participants from 2000-2018. Columns (1) and (2) report marginal effects within position-year for all positions. Column (3) reports marginal effects for quarterbacks only. Negative numbers imply a positive impact on drafting success, since the number 1 represents the highest pick within position-year. T-statistics in parentheses

quarterbacks on average score higher, there is no statistically significant difference in terms of the marginal impact of a higher Wonderlic score.[13] Altogether, these results do not suggest that being Black is a disadvantage in the draft at the quarterback position.

A different picture emerges for players who are neither Black nor White. There is an large unconditional gap in drafting success and drafting position. This gap is diminished but remains large and statistically significant after controlling for the Physical, Productivity and cognitive signals known to Franchise selectors. However, it should be noted that there are only eleven Non Black Non White quarterbacks in our sample. The fact that participation is low reduces the confidence that the results are non-random, while at the same time gives pause for thought as to why the participation of Non Black Non White players is so low. One possibility is that these players understand the low likelihood of drafting success at the quarterback position and select out of the competition, perhaps mirroring the experience of Black players 50 years ago.

Within Position Rank

Table 6 shows the impact of our variables on the within position-year rank of the player. That is, where, relative to the competition in their chosen position and year cohort, did the player rank. This is arguably a more direct comparison over player ability that eliminates differences associated with the value of different playing positions. Columns (1) and (2) return the raw and conditional results respectively for all playing positions, while column (3) returns the estimates for the quarterback position. The qualitative picture from the previous analysis is unchanged by focusing on the within position competition. Black players are preferred overall and experience no disadvantage at the quarterback position, conditional upon observabales. In the case of the eleven Non Black Non White quarterbacks, these players are ranking lower than expected, given their observed characteristics.

Cost of Discrimination

In Becker (1957), a taste-based discriminating employer weighs the costs of discrimination. In the draft, the cost of discrimination is the forgone player talent. However, talent is not uniformly distributed in the draft. The draft is a rank-order contest where talent diminishes from the first picks in the first round to final picks in the last round. Therefore, since the cost of discrimination is lower for later round picks we might expect to find more evidence for racial discrimination there.[14] If some franchises are on the margin of discriminating or not, the cost of discrimination might be particularly relevant.

Table 7 splits the sample into three groups: players drafting in rounds 1 and 2; players drafting in rounds 3 through 5; and players who went undrafted along with those drafting in rounds 6 and 7. The later group comprises players who could

[13] This seems to cohere with what is quantitatively understood to be the use of the Wonderlic score in practice. As noted in a recent article "Wonderlic gives you an area to investigate", the late New York Giants general manager George Young told the Philadelphia Daily News in 1997. "If a guy doesn't have a good score on the test, you don't say he's not smart. But you go in and investigate and find out [why he scored low]. You go in and talk to his coach. You find out how he did in school. You find out how he retains. If you think he's a poor reader and did poorly because it was a verbal test, you give him a non-verbal test." (Siegal, 2015)

[14] We thank our anonymous referees for this suggestion.

be broadly similar in terms of talent yet experienced a sharp difference in drafting success.

Quite a striking pattern results. When all playing positions are considered, the estimated coefficients identifying a Black player changes significantly between the rounds. From the Tobit, in rounds 1-5, Black players do not draft later than White players (marginally earlier in the case of rounds 3-5). In rounds +6, including undrafted players, Black players are more likely to draft later by 10 places, or in the case of the Probit, not at all (6.9 percentage points less likely to draft). A similar pattern is shown in respect Non Black Non White players with even larger differences.

These results are consistent with a cost of discrimination model. However, they are also consistent with a statistical discrimination model in which greater weight is placed on the group average as the individual signals of expected productivity become less informative. This is arguably the case for later round picks whose Combine and College performances are less exceptional.

It is also interesting that this result does not hold in the case of quarterbacks in Table 7. For QBs the estimates are short of statistical significance, albeit the low numbers of observations that results from the sample split could prevent effective inference (Non Black Non White QBs are dropped

Table 7 Drafting sucess by draft round

	Tobit			Probit
Rounds	1-2	3-5	+6	+6
Black	−1.620	−3.973*	10.58***	−0.0694***
	(−1.005)	(−1.775)	(3.016)	(−3.295)
Non Black Non White	1.675	−5.752	34.69***	−0.205***
	(0.573)	(−1.415)	(5.301)	(−5.627)
Physical Signal				
Height	−1.664*	−0.468	0.913	0.0021
	(−1.731)	(−0.343)	(0.429)	(0.162)
Weight	−13.02***	−9.492***	−28.34***	0.163***
	(−7.116)	(−3.458)	(−6.893)	(6.675)
Speed	−10.92***	−10.91***	−27.16***	0.161***
	(−7.976)	(−5.355)	(−8.893)	(9.017)
Productivity Signal				
College Rank	0.0765***	0.0576***	0.00732	0.00002
	(4.561)	(2.632)	(0.240)	(−0.118)
Major School	−14.57***	−6.671	−27.69***	0.167***
	(−3.699)	(−1.447)	(−4.268)	(4.797)
College Performance	−4.636***	−4.098	−9.992**	0.0732***
	(−2.828)	(−1.645)	(−2.337)	(2.882)
Year FE	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.02	0.01	0.02	0.05
Observations	1,156	1,747	3,166	3,166
Quarterbacks only				
Black	6.999	19.75*	4.779	−0.0507
	(0.935)	(1.884)	(0.371)	(−0.748)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.077	0.086	0.092	0.229
Observations	73	72	228	222

1. The sample is NFL Draft combine participants from 2000-2018. The table reports estimates from splitting the sample into three groups, draftees in rounds 1-2, draftees in rounds 3-5 and draftees in rounds 6 and 7 plus undrafted players. The top panel (all playing positions) shows an unexplained disadvantage for Black and Non Black Non White players emerges in the later stages of the draft. T-statistics in parentheses
2. The result does not hold in the case of quarterbacks only. Albeit the number of observations is low even after interpolating Wonderlic scores where missing (114 cases)

due to insufficient variation). If one takes the QB estimates at face value then a cost of discrimination model can still be maintained. It is possible that late round QB picks remain too costly to discriminate against. The top end potential of a late round QB pick is very high. There are some high profile examples of late round QB picks going on to have outstanding NFL careers (most famously Tom Brady). The absence of a QB result is harder to reconcile with statistical discrimination.

Conclusion

Many field experiments and regression-based studies have considered the traditional dimensions of discrimination in labor markets, such as discrimination based on skin color or gender. These studies nearly always find evidence of discrimination against minorities (for gender, see for example, Heilman and Caleo (2018); Wicker et al. (2021)). The estimates of discrimination in these studies can be biased, however, if there is differential variation in the unobservable determinants of productivity or in the quality of majority and minority groups. Moreover, there is also a potential for bias that runs the other way, in the form of selection effects that may preferentially rank some candidates by tint of their skin color more highly. This higher ranking may itself produce higher returns later on in careers irrespective of quality or productivity differences.

In this paper by using observable NFL player draft and performance data, we have a study of discrimination that has sufficient information to correct for these two biases. We have data on how highly workers are ranked prior to hiring and have extensive information on player ability and expected productivity.

Altogether, our results do not find robust evidence of hiring discrimination against Black players in the NFL. However, we note four important caveats. First, the absence of evidence of discrimination at the point of hire does not preclude discrimination at other moments in a player's career, either prior to entry or later on as the player enters free agency. Second, the remarkable selection into playing positions prior to hiring may itself reflect racial stereotyping or the anticipation of discrimination. Third, we *do* find some evidence of hiring discrimination against players who are neither Black or White. Recall, this category comprises players with Hispanic, Pacific Island or Asian backgrounds, approximately 5% of the drafting population. While the low numbers of players from these backgrounds reduces the confidence that the finding is robust, it also provides pause for thought as to why the participation of Non Black Non White players is so low. One possibility is that these players understand the low likelihood of drafting success and select out of the competition, perhaps mirroring the experience of Black players 50 years ago. Fourth, we find some evidence of discrimination against Non White later round picks, consistent with the idea that it is less costly for the franchise to exercise a taste for discrimination with respect to these players.

Nevertheless, the absence of widespread discrimination between Black and White professional sportsmen at the point of hire might not be a surprise. Goff et al. (2002) show that the productivity differential by race in professional basketball and baseball had effectively disappeared by the 1980 s, implying the absence of hiring discrimination. Yet in the labor market more generally, correspondence studies suggest substantial and persistent racial discrimination in call backs to interview (Bertrand and Mullainathan, 2004; Heath and Di Stasio, 2019). One rationalisation for the difference in findings could be that the observed call back discrimination is predominately statistical and that the high quality of individual level signals in the NFL and other professional sports settings is sufficient to render the group level signal of limited value.

A key takeaway from our analysis is that the main descriptive differences in the data are a result of sorting into different playing positions by White and Black players. This is happening prior to being assessed for entry into the NFL, most likely in High School. However, why this happens, what the mechanisms are, the extent to which anticipated discrimination is a factor and whether sorting by position will continue to the same extent into the future are open questions for future research. Our work here only suggests that the policy response should be focused earlier than the point of hire at the NFL Draft.

Author Contributions All authors contributed to the generation of ideas and preparation of the manuscript. Gregory-Smith led the empirical analysis and drafting of the manuscript. All authors are equally responsible for the final product. The author(s) read and approved the final manuscript.

Availability of data and materials Files for replication can be requested from the corresponding author.

Declarations

Conflict of interest The authors declare no competing interests.

A Appendix

The NFL Draft: Details

The NFL draft is an annual event whereby NFL Franchises pick from a pool of eligible college players. Each Franchise is awarded one pick for each of the seven rounds. It is a 'reverse order of finish' draft which means the worst performing team in the prior year is awarded the first pick in each round, the second worst the second pick and so on until the Superbowl champions pick last (32nd). Each franchise will have their own internal rank order informed by the players' performances at College and at the NFL Combine event which occurs earlier in the year. Franchises will also use written opinions from their professional scouts. Franchises can select a player from any playing position.

Each pick in the draft is a asset that can be traded. Franchises may swap picks (e.g. their first round pick for two later round picks) or trade their picks for veteran players, who are already established in the NFL. Trades may cross years (a first round pick today, might be traded for two picks in later years). Trading can occur during the off-season from March to November (the season runs September to February). NFL franchises use draft trade charts which guide how much a draft pick is worth.

Draftee wages are largely determined by the league wide Collective Bargaining Agreement (CBA) which has been in place since 2011. The CBA assigns a maximum and minimum value for each pick position in the Draft. The exact formula for draftee wages is not disclosed but negotiations over wages typically conclude quickly after drafting upon which the agreed wage package is observed. Top draft picks may be able to command amounts towards the maximum for the pick position and may be able to negotiate more guaranteed amounts in their contract. However, the scope for negotiation is small because drafted players are unable to refuse the contract and play for a rival franchise. Additionally, the collective bargaining agreement imposes further restrictions on spending on rookie wages in addition to the total franchise salary cap. There is also a minimum spend for the Franchise; at least 90% of the salary cap. The result is a very tight correlation between the draft pick rank order and player wages for the first four years of the draft class.

Players who are unsuccessful in the NFL draft, can still be picked up as undrafted free agents. Such players are paid significantly less than the last picks in the draft. In 2022, 7th round draftees earned a minimum of $705,000, while the minimum for an undrafted player is $160,000.

Neither drafted players or undrafted players are guaranteed to play. However, our data suggest drafted players are very likely to play at least one NFL game (>90%). Once a draftee has signed to a Franchise they will join the preseason training camps along with any undrafted players that the Franchise has picked up. During preseason, the 90 man squad (including last season's players) is reduced to the playing roster of 53 players by 30th August (the season begins in September). Players who do not make the roster are either released or kept by the Franchise and assigned to their 'Practice Squad'. Cut players keep their signing bonus (which is guaranteed) but do not get their salary which is usually conditional upon making the roster. Practice squad players will get a fraction of the roster salary. Players can be brought up from the Practice Squad during the season to cover injuries, or promoted from the Practice Squad to the playing roster for the following season.

Variable Descriptions

Table 8 Position description

Position Name	Description
Quarterback (QB)	The QB is the leader of the Offense. They relay the play design to the Offense prior to the snap
	They may change the play at the line of scrimmage
	They will either attempt a forward pass, a hand off to their running back or run with the ball themselves
	They are usually the highest paid player in the franchise
Offensive Linesman (Oline)	The Oline engages the opposition's defensive linesman in order to protect the QB or make room for a running back
	A special position on the Oline is the Tight End, who may also attempt to receive a forward pass
Defensive Linesman (Dline)	The Dline engages the opposition's Oline after the snap. They are trying to stop the running back or rush the QB
Defensive Cover (Dcover)	The Dcover protects against the forward pass down the field. They try to block or intercept the pass
Linebackers (LB)	LBs try to anticipate whether the play is a run or pass and defend accordingly
	They may either drop into cover, tackle the running back or rush the QB
Running back (RB)	RBs carry the ball forward on a running play
	They may also block pass rushers from the defense on a passing play or make themselves available for a catch
Wide Receiver (WR)	Wide receivers run down field and attempt to catch the ball on a passing play
Specialists (Sp)	Specialists are Kickers and Punters. A Kicker tries to kick the ball through the sticks for points
	They also start the game with a kick and restart the game after a score
	Punters punt the ball down field giving up possession for field position. They are otherwise not involved

Table 9 Variable description

Variable	Description
Physical Signal	
Height	Height measured in inches
Weight	Weight measured in pounds
Speed	Time over 40yds in seconds. This is the marque event at the Combine
Vertical	The vertical jump. The differential in inches between a player's reach and the top of his standing jump
Bench	The bench press. Total number of repetitions of 225 pounds
Broad jump	Horizontal distance from a standing jump measured in inches
Cone	Time in seconds for the 3 cone drill set 5 yards apart in an L-shape
Shuttle	Time in seconds for the short shuttle cone drill variant
Productivity Signal	
College Rank	College Rank is the football performance ranking of the College from 1 to 298
	(lower numbers associated with a stronger college football team)
Major School	Dummy variable identifying a college that played the majority of their games against teams in Division 1
	Football Sub Division (FBS)
College Performance	The player's rank in their position according to their playing statistics at College relevant to their position as follows
	QB: passing rating; Oline: Average offensive yards per play; Dline: total tackles; Dcover: incomplete passes
	LB: Tackles and incomplete passes weighted equally; RB: rushing yards per play; WR: receiving yards per play
	Sp: total points
Heisman Finalist	Dummy variable as voted for by sports journalists and previous Heisman trophy winners
Wonderlic	A problem solving aptitude score measured from 1 to 50 with higher numbers indicating better performance

Table 10 Descriptive statistics by demographic group

	Black			White			Non Black Non White		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Physical Signal									
Height	4,206	-.220	.975	1,669	.531	.867	361	.117	.904
Weight	4,206	-.149	.987	1,669	.286	.926	361	.409	1.07
Speed	4,124	.236	.944	1,591	-.514	.893	354	-.433	1.06
Vertical	3,355	.168	.978	1,197	-.395	.9 2	268	-.344	1.05
Bench	2,975	-.152	.966	998	.359	.955	258	.364	1.09
Broad jump	3,305	.208	.954	1,187	-.475	.908	262	-.469	1.067
Cone	2,660	.077	.989	1,104	-.121	.970	236	-.303	1.132
Shuttle	2,753	.067	.996	1,133	-.112	.959	240	-.240	1.135
Productivity Signal									
College Rank	4,206	40.0	42.8	1,672	43.4	41.0	361	43.9	42.1
Major School	4,206	.953	.210	1,672	.935	.245	361	.939	.239
College Performance	4,206	.301	.347	1,669	.228	.344	361	.248	.334
Heisman Finalist	90	.156	.363	278	.151	.358	11	.181	.404
Wonderlic	57	21.9	6.97	183	28.5	5.94	7	23.3	2.87

Notes:
1. Low N on Wonderlic and Heisman Finalist due to limited data at non-QB positions
2. All Physical signals standardised to a mean of zero and a standard deviation of 1

References

Arrow K. In: Pascal AH, editor. Models of Job Discrimination. D. C. Heath; 1972.

Becker GS. The Economics of Discrimination. University of Chicago Press; 1957.

Berri DJ, Simmons R. Race and the Evaluation of Signal Callers in the National Football League. J Sports Econ. 2009;10(1):23–43. https://doi.org/10.1177/1527002508327383.

Berri DJ, Simmons R. Catching a draft: On the process of selecting quarterbacks in the National Football League amateur draft. J Prod Anal. 2011;35(1):37–49.

Bertrand M, Mullainathan S. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. Am Econ Rev. 2004;94(4):991–1013.

Bohren JA, Imas A, Rosenberg M. The dynamics of discrimination: Theory and evidence. Am Econ Rev. 2019;109(10):3395–436.

Burnett NJ, Scyoc LJV. Compensation Discrimination for Defensive Players: Applying Quantile Regression to the National Football League Market for Linebackers and Offensive Linemen. J Sports Econ. 2015;16(4):375–89. https://doi.org/10.1177/1527002513492772.

Caselli M, Falco P, Mattera G. When the Stadium Goes Silent: How Crowds Affect the Performance of Discriminated Groups. J Labor Econ Forthcom. 2022.

Conlin M, Emerson PM. Discrimination in Hiring Versus Retention and Promotion: An Empirical Analysis of Within-Firm Treatment of Players in the NFL. J Law Econ Organ. 2005;22(1):115–36. https://doi.org/10.1093/jleo/ewj005.

Ducking J, Groothuis PA, Hill JR. Compensation discrimination in the NFL: an analysis of career earnings. Appl Econ Lett. 2014;21(10):679–82. https://doi.org/10.1080/13504851.2014.884688.

Dufur MJ, Feinberg SL. Race and the NFL draft: Views from the auction block. Qual Sociol. 2009;32(1):53–73.

Gill A, Brajer V. Wonderlic, Race, and the NFL Draft. J Sports Econ. 2012;13(6):642–53. https://doi.org/10.1177/1527002511429575.

Goff BL, McCormick RE, Tollison RD. Racial Integration as an Innovation: Empirical Evidence from Sports Leagues. Am Econ Rev. 2002;92(1):16–26. https://doi.org/10.1257/000282802760015586.

Goldin C, Rouse C. Orchestrating impartiality: The impact of "blind" auditions on female musicians. Am Econ Rev. 2000;90(4):715–41.

Heath AF, Di Stasio V. Racial discrimination in Britain, 1969–2017: a meta-analysis of field experiments on racial discrimination in the British labour market. Br J Sociol. 2019;70(5):1774–98. https://doi.org/10.1111/1468-4446.12676.

Heilman ME, Caleo S. Combatting gender discrimination: A lack of fit framework. Group Process Intergroup Relat. 2018;21(5):725–44.

Hendricks W, DeBrock L, Koenker R. Uncertainty, Hiring, and Subsequent Performance: The NFL Draft. J Labor Econ. 2003;21(4):857–86. https://doi.org/10.1086/377025.

Jensen AR. An examination of culture bias in the wonderlic personnel test. Intelligence. 1977;1(1):51–64. https://doi.org/10.1016/0160-2896(77)90026-5.

Kahn LM. The Effects of Race on Professional Football Players' Compensation. ILR Rev. 1992;45(2):295–310.

Keefer QAW. Compensation Discrimination for Defensive Players: Applying Quantile Regression to the National Football League Market for Linebackers. J Sports Econ. 2013;14(1):23–44. https://doi.org/10.1177/1527002511413288.

Késenne S. The Economic Theory of Professional Team Sports: An Analytical Treatment. Edward Elgar Publishing; 2014.

Levy AH. Tackling Jim Crow, Racial Segregation in Professional Football. McFarland and Co., Inc; 2003.

Lyons B, Hoffman B, Michel J. Not Much More Than g? An Examination of the Impact of Intelligence on NFL Performance. Hum Perform. 2009;22:225–45.

Massey C, Thaler RH. The loser's curse: Decision making and market efficiency in the National Football League draft. Manag Sci. 2013;59(7):1479–95.

Mercurio E, Filak VF. Roughing the passer: The framing of black and white quarterbacks prior to the NFL draft. Howard J Commun. 2010;21(1):56–71.

Mincer J. The Distribution of Labor Incomes: A Survey With Special Reference to the Human Capital Approach. J Econ Lit. 1970;8(1):1–26.

Neumark D. Age Discrimination in Hiring: Evidence from Age-Blind vs. Non-Age-Blind Hiring Procedures. NBER Working Paper 26623. 2020.

Neumark D, Rich J. Do Field Experiments on Labor and Housing Markets Overstate Discrimination? A Re-examination of the Evidence. ILR Rev. 2019;72(1):223–52.

Parsons CA, Sulaeman J, Yates MC, Hamermesh DS. Strike Three: Discrimination, Incentives, and Evaluation. Am Econ Rev. 2011;101(4):1410–35.

Phelps ES. The statistical theory of racism and sexism. Am Econ Rev. 1972;62(4):659–61.

Schwartzstein J. Selective attention and learning. J Eur Econ Assoc. 2014;12(6):1423–52.

Siegal A. How A Multiple-Choice Test Became A Fixture Of The NFL Draft. FiveThirtyEight. 2015.

Simmons R. Professional Labor Markets in the Journal of Sports Economics. J Sports Econ Forthcom. 2021.

Smith TG. Outside the pale: The exclusion of Blacks from the National Football League, 1934–1946. J Sport Hist. 1988;15(3):255–81.

Sussman RW. The Myth of Race: The Troubling and Persistence of an Unscientific Idea. Cambridge: Harvard University Press; 2014.

Szymanski S. Chapter 34 The theory of tournaments. In: Andreff W, Szymanski S, editors. Handbook on the economics of sport. Edward Elgar Publishing Ltd; 2006. p. 337–41.

Volz BD. Race and quarterback survival in the National Football League. J Sports Econ. 2017;18(8):850–66.

Wicker P, Breuer C, Dallmeyer S. The gender earnings gap among elite athletes in semi-professional sports. Manag Sport Leis. 2021:1–18.

Wolfson J, Addona V, Schmicker RH. The quarterback prediction problem: Forecasting the performance of college quarterbacks selected in the NFL draft. J Quant Anal Sports. 2011;7(3)