

Labor Market Discrimination in the English Premier League

Nico Rotundo

The University of North Carolina at Chapel Hill

September 2021

I would like to express my sincere thanks to Dr. Luca Maini for advising this project. The contributions of his guidance and mentorship to this study are inestimable and exceedingly appreciated. Additionally, I would like to extend a special thanks to the Matthew Guest Family Fund and the UNC Department of Economics, without whom this research would not have been possible.

Abstract

Discussions of racism in football often focus on incidents involving individuals as opposed to systemic discrimination by clubs. In this study, I assess the existence of labor market discrimination against black players by clubs in the English Premier League over the past 20 years. I use a market test for discrimination adapted from Szymanski (2000). Using a fixed-effects OLS regression, I compare the seasonal performance of teams against the appearances and minutes accounted for by black players for each club relative to the league average of these variables while controlling for wage expenditures. I find that there is not enough evidence to suggest that there exists racial discrimination against black footballers in this labor market for the years considered. This result contrasts that of the earlier Szymanski paper and suggests progress has been made in addressing the issue of discrimination in this space.

Introduction

Conversation in popular media regarding discrimination in football often centers on the actions of fans or referees in the context of matches between teams. Occurrences in this space are closely monitored and captured on video, so it is often that events involving racial discrimination populate the headlines when they transpire. This fan-focused and referee-focused perspective of discrimination has also received a significant degree of attention from academic studies (Pia, 2021; see also Muller, Zoonen, de Roode, 2007; Wagner-Egger, Gyax, Ribordy, 2012; Chu, Nadarajah, Afuecheta, Chan, Xu, 2014; Wynn, 2012). In regard to systemic discrimination against footballers, however, it is likely that such phenomena will also be reflected on the side of club operations in football if it is present.

There has been relatively little empirical analysis thus far that has considered racial discrimination in this space, with an important exception being the work of Stefan Szymanski (2000). In that paper, Szymanski develops a market test for discrimination against black footballers and applies it to the labor market for professional footballers in England for the seasons spanning 1978-1993, finding indications of discrimination. It is essential to revisit this prior study to determine whether or not the result from the previous paper still holds given the time difference and the changes that have occurred in football since.

The goal of this paper is to provide an updated assessment of discrimination in this market. I replicate the methods of Szymanski (2000) in an attempt to determine if statistically significant measures of discrimination are still present in the market for footballers in the English Premier League specifically. While the previous study did determine that there were indications of discrimination when considering the largest 20 clubs of the sample in any given year, it also

considered clubs from the lower divisions of English football¹. For data availability reasons and time constraints, this study only considers clubs in the English Premier League for any given season.

Football is a highly meritocratic space and performance is easily measured, and the labor market for footballers is thus expected to be highly competitive. This conjecture has been empirically investigated, and the results conform to the theoretical expectations (Szymanski, 2000). While this study centers on the labor market for footballers, in particular, the implications of its results may be extrapolatable to other markets that share similar characteristics. Such markets include the labor market for athletes of all sports and generally for occupations where workers' productivity is quantified such as commission-based employment. I find a lack of evidence to suggest racial discrimination in the labor market for professional footballers in English soccer for the seasons spanning 2000-2020, and the discrepancy between this result and that of Szymanski (2000) is promising. Further research is needed to establish what factors contributed to this shift. If causation can be established with the factors that are correlated with this development, we can gain a clearer insight into possible solutions to labor market discrimination.

To start, I introduce the model developed by Szymanski (2000), which is based on the work of Gary Becker (1957) and Kenneth Arrow (1998). The former asserts that since discriminating owners will possess a lesser demand for a particular group of workers, these individuals will command a lower wage in the market. Arrow extends this reasoning, suggesting that this lower wage allows competing firms to hire equally-skilled workers at a lower cost relative to the discriminating firms. In this way, competitive labor markets eliminate

¹ While not specifying the English Premier League specifically, it is assumed that the largest 20 clubs for any given season approximate the modern-era Premier League to an adequate degree for comparison.

discrimination by imposing a type of tax on employers who discriminate based on characteristics that do not affect productivity.

With this framework in mind and using the definition of “unequal pay for equal work” to define discrimination, Szymanski defines the following test for discrimination against black footballers: “Discrimination can be said to exist if clubs fielding an above-average proportion of black players systematically outperform clubs with a below-average proportion of black players, after one controls for the wage bill” (pg.1). Since then, a similar type of test has been used to examine ethnicity discrimination in the National Hockey League (Longley & Szymanski, 2001; Mongenon, 2013) and sex discrimination in Japan (Kawaguchi 2007). The studies investigating the National Hockey League find evidence of discrimination against Francophone players in this league and that French Canadian and American players are wage-discriminated against by Canadian teams, respectively. Using Japanese firm-level panel data from the 1990s, the latter study finds evidence of discrimination against females in the workforce.

Using this model, I then review the empirical methodology of this study. The data employed are seasonal club-level aggregations of the following player-level data points: appearances, race, minutes played, and wage. The data come primarily from Worldfootball.net, with wage data being the exception. For club wages, I collect data from English news reports. Additionally, this analysis uses the final positions and point totals of clubs for a given season. These panel data span the seasons from 2000/2001 to 2019/2020 and include 43 clubs. I use python to scrape this data and the machine-learning software Deepface to analyze player races. Information pertaining to this software can be accessed using the associated reference in the reference section of this paper. Through the use of this software, I remove the subjectivity associated with categorizing race and maintain replicability.

The regression models I use are adaptations from Szymanski (2000) and regress individual club performance with the relative wage bills of clubs, relative player turnover of clubs, and the presence of black players at clubs across seasons. Wages and player turnover have been shown to be highly significant in their correlation with performance. Additionally, I use wages as a proxy for the quality of individual club rosters and thus control for variations between clubs in this respect. Furthermore, wages have been shown to explain a high degree of variation in club performance. Thus, by controlling for wages and player turnover, we can observe the contributions of roster demographics to performance. Under the assumptions that player talent does not vary as a factor of race and that the labor market for players is competitive, performance differences between teams using an above-average and below-average proportions of black players will illuminate discrimination if it is present. This market test approach circumvents the issues concerning confounding unobservables normally encountered in traditional studies of labor market discrimination, which often regress individual wages with demographic categories and attribute discrepancies to discrimination. The issues with the latter approach are discussed by James Heckman (1998).

My results suggest that there is not enough evidence to assert that there is discrimination against black players in the labor market for footballers in the English Premier League. This result contrasts that of Szymanski (2000), suggesting that a shift toward nondiscrimination has occurred in this market since this prior study. While not the central goal, this study also finds that club wage bills and player turnover over the course of a season are highly significant in predicting club performance, which is consistent with prior literature. However, club wages explain considerably less of the variation in club performance than when this relationship was considered in the prior study.

I then consider the possible reasons for the shift in the presence of discrimination in this labor market for footballers. One explanation may be that clubs have become profit maximizers to a higher degree. This may result from the significant increases in TV revenues from broadcasting deals that accompanied the formation of the English Premier League or from the increased presence of investors in clubs. Alternatively, simply the increase in the number of black players in the labor market may have resulted in this shift, as this phenomenon theoretically imposes a greater penalty on non-discriminators. However, more research is needed to establish causation between these changes and the alleviation of discrimination in this market.

Literature Review

Studies exploring discrimination in football from the channel of club operations are sparse. As discussed in the introduction, Szymanski (2000) confirmed the presence of discrimination against black footballers in the English Premier League for the years 1974 through 1993. Goddard and Wilson (2009) find that black players tend to be employed by teams in higher divisions of professional English football, suggesting that teams in the lower divisions may be more apt to discriminate. Wilson and Ying (2010) find that certain nationalities are underrepresented in professional football in Europe that parallel the racial biases found by Szymanski (2000). More recently, Principe and Ours (2021) determined that racial wage discrimination in the labor market for players in Italian professional football does not exist. Additionally, Swift (2017) concludes the existence of nationality-derived wage discrimination in the professional football league of the United States. This latter study may present the opportunity for a type of natural experiment, as the structure of professional football in the United States is distinct from that of Europe. In sum, studies of less recent time periods appear to more readily suggest the existence of labor market discrimination in European football than

those of more recent time periods. The results of this study are consistent with this trend, as I do not find sufficient evidence to suggest labor market discrimination.

The literature investigating fan and referee discrimination is comparatively more developed. We will focus primarily on research that investigates whether fan biases account for discrimination in the labor market for footballers. Szymanski (1998, 2000) and Wilson and Ying (2010) conclude that the discrimination found in each study does not originate from fans. However, Principe and Ours (2021) do conclude that black players are systemically rated lower by Italian newspapers, suggesting that unconscious biases still may affect the assessment of individual players. Taken together, these studies suggest that if discrimination exists in the labor market for black footballers, it seems to have a higher probability of originating from the club side of operations.

In order to explore the association between club performance and race as necessitated by the market test approach, the variables of club wage bill and player turnover must be controlled. The literature addressing the importance of these variables will now be discussed. Wages have been shown in previous research to be positively correlated with sporting performance in football (Kuper & Szymanski, 2010; see also Gasparetto, 2012, Szymanski, 2000). Furthermore, wages appear to explain a high degree of variability in the final standings of clubs in English football. Szymanski (2000) reports that regressing league position and wages results in an R^2 value of .9. Player turnover has comprehensively been shown in previous research to be negatively correlated with sporting performance across the five major European football leagues (Besson, Poli, & Ravenel, 2018; see also Szymanski 2000).

Economic Model

The economic model I use in this paper is an adaptation of the market test for discrimination developed by Szymanski (2000). The only parameter altered is that clubs seek to maximize their share of non-black footballers as opposed to simply white footballers. I make this change due to the fact that this analysis uses software that employs five separate descriptors of race². By maintaining a high degree of correspondence with the previous paper, the results from this paper can be more readily compared with those of the preceding one.

We begin with the assumption that clubs seek to maximize a weighted average of profits π and non-black footballers s . As such, the utility function U for team i is as follows:

$$U_i = \alpha s_i + (1 - \alpha)\pi. \quad (1)$$

The share of non-black players is expressed as the proportion of non-black talent t_n in the total pool of talent T where $T = t_n + t_b$. Thus,

$$s_i = \frac{t_{in}}{t_{in} + t_{ib}}. \quad (2)$$

The revenue R of team i is determined by sporting success w , which is a function of talent.

Additionally, costs C_i are dependent on the share of talent employed. Thus, profit can be represented by

$$\pi_i = R_i(w_i(T_i)) - C(s_i)T_i. \quad (3)$$

Clubs that impose an additional characteristic on prospective players aside from talent, such as a particular race, will pay a premium based on the degree to which they discriminate. Let s^*

² Deepface classifies an individual's race as being white, black, latino-hispanic, middle-eastern, indian, or asian.

represent the optimal share of non-black talent in the team. This number will be determined by the supply and demand of talent more generally, and having a proportion of non-black players beyond this amount will begin to impose a premium on costs. Costs can thus be represented by

$$c_i = c_{i0}[1 + (s_i - s^*)^2] \text{ if } s_i > s^* \quad (4.1)$$

and

$$c_i = c_{i0} \text{ if } s_i \leq s^*. \quad (4.2)$$

Placing (1) - (4.2) together presents the first-order conditions for firm i :

$$t_{in}: \alpha \frac{\partial s_i}{\partial t_{in}} + (1 - \alpha) \left(\frac{\partial R_i}{\partial w_i} \frac{\partial w_i}{\partial t_{in}} - c_i \frac{\partial s_i}{\partial t_{in}} - c_i \frac{\partial T_i}{\partial t_{in}} \right) = 0. \quad (5)$$

and

$$t_{ib}: \alpha \frac{\partial s_i}{\partial t_{ib}} + (1 - \alpha) \left(\frac{\partial R_i}{\partial w_i} \frac{\partial w_i}{\partial t_{ib}} - c_i \frac{\partial s_i}{\partial t_{ib}} - c_i \frac{\partial T_i}{\partial t_{ib}} \right) = 0. \quad (6)$$

Subtracting (6) from (5) gives us the final result:

$$t_{in} = s^* T + \frac{\alpha}{2(1-\alpha)c_{i0}}. \quad (7)$$

This result characterizes the decision of an owner on how much non-black talent to employ, ranging from a non-discriminating owner ($\alpha = 0$) to a highly discriminatory owner ($\alpha = 1$).

Empirical Methodology

I collect the data in this paper primarily from the football statistics website Worldfootball.net. Before beginning the data collection process, I checked the information from

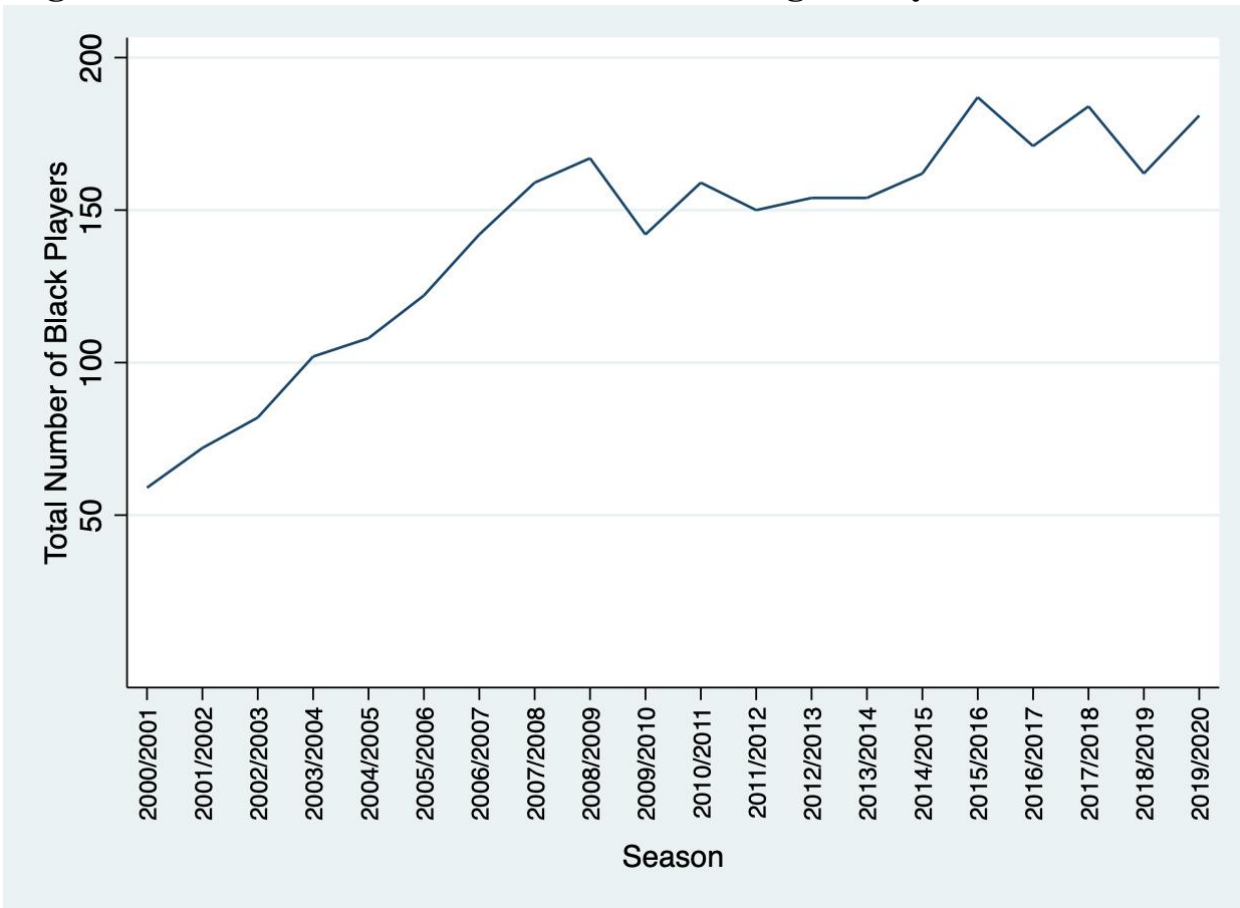
this website with other websites including the official English Premier League website and Transfermarkt to ensure accuracy. The base data points I use are the number of minutes played by players, the appearances of players, and the league points and position of clubs. These data are separated by each Premier League season.

This paper employs the use of the facial analysis software Deepface to determine the race of each player. In doing so, the issue of subjectively determining race is avoided and the option of replicability is maintained. While not infallible, the program follows objective criteria for determining race. Thus, the error margin in determining player race is consistent, so the effects of this error should not bias the results to any significant degree. In other words, the OLS assumption $E[e|X_i] = 0$ is satisfied for all independent variables. I then organize player-level data by club affiliation in order to determine the total appearances and number of minutes for each club and the breakdown of these data points based on race. The only data I do not collect from Worldfootball are club wages. I collect these from the following UK news outlets: The Guardian, The Telegraph, Planetfootball, Talksport, and Totalsport.

There are a few limitations of this data. Some footballers from Worldfootball.net do not have associated pictures. These players are considered missing observations and are removed from the analysis before the aggregation of the data, as their races can not be determined by the software, and constitute roughly 14% of the total player-year level entries. Additionally, particular season wage bill data are not available for 7 club entries out of the 400 total clubs. Thus, these clubs are removed from the data set and this change is reflected in the club-level aggregation of the data.

Data

Figure 1, Total Number of Black Premier League Players



In total, this analysis encompasses 43 clubs that have played in the Premier League from the 2000/2001 season through the 2019/2020 season. There are 393 individual observations used in the regressions. Time-series data detailing the total number of black players in the Premier League since 2001 as measured by Deepface is captured in Figure 1, displaying an increasing trend. Note that in the first year of data used by Szymanski (2000), 1974, only four black players played across the top four English football divisions. This simple fact is already potentially suggestive of a change in attitudes towards black players from the time period observed by Szymanski.

Table 1, Summary Statistics

Number of Teams	393
Number of Players	11375
Average Wage Bill of Clubs in 2000	£28.16m
Average Wage Bill of Clubs in 2020	£160.4m
Average Number of Appearances in 2000 for clubs (all players)	306 ³
Average Number of Appearances in 2000 for clubs (black players)	66
Average Number of Appearances in 2020 for clubs (all players)	530
Average Number of Appearances in 2020 for clubs (black players)	133

Both of the regression models for discrimination that I use in this paper are adaptations from Szymanski (2000). The first one uses the exact same variables as this prior paper for the purposes of direct comparison⁴. The second model is a slight variation of this model, which will be explained below.

The construction of the first model is as follows:

$$\ln(position_{it}) = \alpha_i + \beta_0 + \beta_1 \ln\left(\frac{wage_{it}}{wage_t}\right) + \beta_2 (individual\ appearances_{it} - \overline{individual\ appearances_t}) \quad (1)$$

³ The minimum number of appearances in a given season is 38 games * 11 players = 418. This lower number reflects the removal of players discussed above. However, this number is still useful in a comparative sense when taken together with the number of appearances in 2020 and also the average number of appearances of black players in 2000.

⁴ The dummy variable that documents the different divisions of English Professional Football in Szymanski (2000) is omitted, as this analysis only surveys Premier League teams.

$$+ \beta_3 (\text{appearances of black players}_{it} - \overline{\text{appearances of black players}_{it}}) + \epsilon_{it} .$$

The variable $\ln(\text{position}_{it})$ captures the performance of an individual club in a particular season and is the dependent variable in this analysis. Specifically, this variable is the log of the final position of a team in a particular season. As noted by Szymanski (2000), the log transformation gives higher importance to finishing higher on the table. The variable $\ln(\frac{\text{wage}_{it}}{\text{wage}_t})$ captures the

log difference of the wage bill of an individual club in a season and the average wage bills of all clubs in the same season. The variable

$\text{individual appearances}_{it} - \overline{\text{individual appearances}_t}$ references the difference in the number of players who make at least one appearance for an individual club in a particular season and the average number of players per club who make at least one appearance in the same season. This variable reflects the degree to which teams experience player turnover. The variable $\text{appearances of black players}_{it} - \overline{\text{appearances of black players}_{it}}$ is the appearances of black players for each club in a particular season compared to the average for the same season. If there is discrimination against black players in the labor market, the associated coefficient of this variable should take a positive sign when regressed against club performance. The variable α_i captures the heterogeneity of time-invariant variables specific to clubs such as geographic location, area income, and work culture. This variable is then eliminated by using a fixed-effects regression.

The second model is constructed as follows:

$$\begin{aligned}
 league\ points_{it} = & \alpha_i + \beta_0 + \beta_1 \ln\left(\frac{wage_{it}}{wage_t}\right) \\
 & + \beta_2 (individual\ appearances_{it} - \overline{individual\ appearances_t}) \\
 & + \beta_3 (proportion\ minutes\ black_{it} - \overline{proportion\ minutes\ black_t}) + \epsilon_{it}.
 \end{aligned} \tag{2}$$

Two variables are altered in this model: the measure of performance and the measure of the presence of black players. The number of points that a team accumulates over a season is used as the proxy for performance. This metric provides an additional layer of insight into the performance of a team in a season because it accounts for the relative rank of teams within a given season. For example, using league points captures the difference in winning the league by 1 point and 10 points. The presence of black players on a team is captured by the difference between the proportion of minutes accounted for by black players for each club in a particular season and the average proportion of minutes given to black players by premier league clubs over the course of a season. This metric provides a more precise picture of the relative playing time of black players for each team. The fixed effect constant and the variables measuring relative wages and turnover are unchanged from the first model.

Results

In both regressions, the variables referencing wages and player turnover possessed the expected sign and were significant. The variable referencing the presence of black players was found to be insignificant in predicting club performance. Thus, the results from both regression models are consistent. Note that I ran different variations of the following regressions, including splitting the data into seasons before and after 2010, omitting the fixed effects dummy, and using

a robust fixed effects regression. There were no significant differences in the outputs in any of these cases. Thus, only the original regressions are reported.

Table 2, Fixed Effects Regression Results for $\ln(\text{position}_{it})$

<u>Independent Variable</u>	<u>Coefficient (standard error)</u>
Constant	-.2895934 (.0357)
$\ln\left(\frac{\text{wage}_{it}}{\text{wage}_t}\right)$	-2.021342 (.2825)*
$\text{individual appearances}_{it} - \overline{\text{individual appearances}_t}$.1753654 (.0146)*
$\text{appearances of black players}_{it} - \overline{\text{appearances of black players}_t}$.001443 (.0010)
$R^2 = 0.6782$	
$n = 393$	

*p < .05

Notes: Standard errors are clustered at the team level

Table 3, Fixed Effects Regression Results for league points_{it}

<u>Independent Variable</u>	<u>Coefficient (standard error)</u>
Constant	55.53128 (.5156)
$\ln\left(\frac{wage_{it}}{wage_t}\right)$	23.89644 (4.076)*
$individual\ appearances_{it} - \overline{individual\ appearances_t}$	-1.694668 (.1525)*
$proportion\ minutes\ black_{it} - \overline{proportion\ minutes\ black_t}$	-8.752926 (5.3175)
$R^2 = 0.7143$	
$n = 393$	

* p < .05

Notes: Standard errors are clustered at the team level

Discussion

These data suggest that there is not enough evidence to conclude there is discrimination against black players in the labor market for professional footballers. This finding parallels those of Principe and van Ours (2021), who found that clubs in the top tier of Italian Football do not participate in wage discrimination against black footballers when analyzing data from the 2009/2010 season to the 2017/2018 season. Instead, their study suggests that there is sufficient competition to remove wage discrimination between races of players. Comparatively, indications of discrimination were found in the top league of English football for the seasons spanning 1978-1993 (Szymanski, 2000).

From these outputs, we also observe that wages and player turnover are significant in explaining the sporting performance clubs. However, note that wages alone appear to explain

less of the variation in club performance than previously thought. While Szymanski (2000) reports an R^2 of around .9 when regressing league rank and wages, I find this value to currently be at around .26. However, this difference may in part derive from the fact my analysis only observes Premier League clubs. As the Szymanski study considers the top four divisions of English Football, it is expected that the regression in their case possesses more explanatory power.

The interpretation of coefficients is slightly different for both models. In regard to the first model, we see that a 10% increase in relative wage expenditure is expected to decrease $\ln(\text{position})$ by 20%. An increase in relative player turnover by 1 is expected to increase $\ln(\text{position})$ by 17%. An increase in relative appearances of black players by 1 is expected to increase $\ln(\text{position})$ by .14%, although note that this result is not statistically significant.

In regard to the second model, we see that a 1% increase in relative wage expenditure is expected to increase a club's season point total by .24. An increase in relative player turnover by 1 is expected to decrease a club's season point total by 1.7. An increase in relative appearances of black players by 10% is expected to decrease a club's season point total by .88, although note that this result is not statistically significant.

Note there are some limitations that must be discussed. First, this data only considers Premier League teams. If the dataset included clubs from lower divisions, we may find that the results of the study shift towards that of more discrimination present. As revenue decreases further down divisions, it is plausible that discrimination may increase as a result of a decreased profit-maximizing incentive. Second, it must be noted that sports markets are unique to some degree, as competitors in this space have certainty in their relative rankings, and performance is

easily measured at a team and individual level. Thus, hesitance should be employed in extrapolating the conclusions of this study to markets that do not share these qualities.

Conclusions

There are some theoretical explanations for this apparent shift in the labor market practices of English clubs since 1993. In order to contextualize the following conjectures, it is necessary to explain the economic implications of the formation of the English Premier League in 1992. The emergence of the English Premier League was accompanied by a then-unprecedented television broadcasting deal worth £191m and a partially merit-based allocation of these funds. Furthermore, this figure would increase significantly in the following years, increasing 64 percent on average every time a new deal has been negotiated (Gazapo, n.d.). In the current cycle (2019-2022), the domestic broadcasting revenue of the Premier League is £4.55b.

The first explanation of this shift is that the increase in potential revenue from on-field success has incentivized clubs to put more emphasis on revenue and less on outside considerations. In the context of the model, the increase in potential revenue is expressed as an increase in the derivative of revenue with respect to sporting success. As revenues increase in this space, discriminating clubs are forced to pay a progressively higher premium for their bias against black footballers, as the forgone profits of winning are larger. In sum, clubs have more to lose by discriminating than was the case previously.

It is important to note that clubs in European football are generally understood to operate as win-maximizers as opposed to profit-maximizers (Garcia-del-Barrio & Szymanski, 2009; Sloane, 2015). However, Goddard and Sloane (2015) suggest that win-maximizing and profit-maximizing models become increasingly synonymous as broadcasting revenue increases.

Thus, while not necessarily by choice, the formation of the Premier League and the subsequent increases in broadcasting income have resulted in clubs becoming profit maximizers to a higher degree. Theoretically, this shift should prompt clubs to exploit any inefficiencies in the labor market, including those caused by discrimination, to a higher degree than previously. In this way, discrimination is competed out of the market. The key observation from this conjecture is that individual preferences do not necessarily need to change for discrimination to be alleviated at a systemic level. This explanation is in line with the theoretical model proposed by Becker (1957).

This shift toward profit maximization may also be driven by the increased presence of professional investors and profit-focused businessmen in the club-side of football operations. Likewise, it is now common for clubs to raise money in some form on the stock market. Szymanski (2000) conjectured that this shift from clubs being a “hobby” for wealthy businessmen may be a possible conduit for addressing the presence of discrimination in the market. Indeed, Arrow (1998) suggests that “competition in the capital market should be effective in eliminating discrimination.” Theoretically, a profit-maximizing investor would be able to exploit the inefficiencies faced by a discriminating owner in the case that they take over a club. However, more empirical research is needed to determine the true effects of such events.

An additional explanation of this shift is the overall increase in the number of black footballers in the labor market documented in Figure 1 since the first paper. This increase may be driven by better scouting techniques and expansions of scouting networks. As noted by Szymanski (2000), a small relative number of black players in the market reduces the tradeoff of a discriminating club, as the supply of talent for non-discriminating clubs and discriminating clubs is similar. However, as the number of black footballers increases and discriminators thus

continually restrict their relative market for players, they are forced to pay a higher premium for non-black footballers.

In sum, the results of my study display that progress has been made regarding the issue of discrimination against black players in the labor market for footballers. Contrasting the results of Szymanski (2000), I find that there is not enough evidence to suggest racial discrimination in this form is present in this market. While we consider conjectures as to why this shift has occurred, it is clear that more research is needed to establish causality. Further work on this topic is valuable, as it may illuminate possible methods of addressing discrimination in other markets or other settings.

References

- Arrow, K. (1998). What Has Economics to Say About Racial Discrimination?. *Journal Of Economic Perspectives*, 12(2), 91-100. <https://doi.org/10.1257/jep.12.2.91>
- Becker, G. (1957). *The Economics of Discrimination*. University of Chicago Press.
- Chu, J., Nadarajah, S., Afuecheta, E., Chan, S., & Xu, Y. (2013). A statistical study of racism in English football. *Quality & Quantity*, 48(5), 2915-2937. doi: 10.1007/s11135-013-9932-3
- deepface. (2021). Retrieved 11 September 2021, from <https://pypi.org/project/deepface/>
- Garcia-del-Barrio, P., & Szymanski, S. (2009). Goal! Profit Maximization Versus Win Maximization in Soccer. *Review Of Industrial Organization*, 34(1), 45-68. <https://doi.org/10.1007/s11151-009-9203-6>
- Gasparetto, T. (2012). Relationship between Wages and Sports Performance. *The Empirical Economics Letters*.

Goddard, J., & Sloane, P. (2015). *Handbook on the Economics of Professional Football*. Edward Elgar Publishing.

Goddard, J., & Wilson, J. (2008). Racial discrimination in English professional football: evidence from an empirical analysis of players' career progression. *Cambridge Journal Of Economics*, 33(2), 295-316. <https://doi.org/10.1093/cje/ben038>

Heckman, J. (1998). Detecting Discrimination. *Journal Of Economic Perspectives*, 12(2), 101-116. doi: 10.1257/jep.12.2.101

Kawaguchi, D. (2007). A market test for sex discrimination: Evidence from Japanese firm-level panel data. *International Journal Of Industrial Organization*, 25(3), 441-460. <https://doi.org/10.1016/j.ijindorg.2006.05.006>

Mongeon, K. (2013). A Market Test for Ethnic Discrimination in the National Hockey League. *Journal Of Sports Economics*, 16(5), 460-481. <https://doi.org/10.1177/1527002513502490>

Müller, F., van Zoonen, L., & de Roode, L. (2007). Accidental Racists: Experiences and Contradictions of Racism in local Amsterdam Soccer Fan Culture. *Soccer & Society*, 8(2-3), 335-350. <https://doi.org/10.1080/14660970701224608>

Poli, R., Ravenel, L., & Besson, R. (2018). *Monthly Report 38*. Football-observatory.com. Retrieved 11 September 2021, from <https://football-observatory.com/IMG/sites/mr/mr38/en/>.

Principe, F., & van Ours, J. (2021). Racial Bias in Newspaper Ratings of Professional Football Players. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3897655>

Siering, P. (2021). KICKING DISCRIMINATION: EXPLORING THE IMPACT OF CSR INITIATIVES REGARDING DISCRIMINATION IN SOCCER ON FAN ATTITUDES.

Sloane, P. (2015). The Economics of Professional Football Revisited. *Scottish Journal Of Political Economy*, 62(1), 1-7. <https://doi.org/10.1111/sjpe.12063>

Swift, M. (2017). Testing for Nationality Discrimination in Major League Soccer.

Szymański, S., & Kuper, S. (2015). *Money and football*. Nation Books.

Szymanski, S., & Longley, N. (2001).

Szymanski, S. (2000). A Market Test for Discrimination in the English Professional Soccer Leagues. *Journal Of Political Economy*, 108(3), 590-603.
<https://doi.org/10.1086/262130>

Wagner-Egger, P., Gygax, P., & Ribordy, F. (2012). Racism in Soccer? Perception of Challenges of Black and White Players by White Referees, Soccer Players, and Fans. *Perceptual And Motor Skills*, 114(1), 275-289.
<https://doi.org/10.2466/05.07.17.pms.114.1.275-289>

Wilson, D., & Ying, Y. (2010). Nationality preferences for labour in the international football industry. *Applied Economics*, 35(14), 1551-1559.
<https://doi.org/10.1080/000368403200010048>

Wynn, A. (2012). Red Card Racism: Using the Court of Arbitration for Sport (CAS) to Prevent and Punish Racist Conduct Perpetrated by Fans Attending European Soccer Games.

Appendix A

Definitions of Variables

- $league\ position_{it}$: the position that a club finished in a particular season
- $\ln (league\ position_{it}) = (\ln (league\ position_{it}) / 21 - league\ position_{it})^5$
- $league\ points_{it}$: the points accumulated by a club in a particular season
- $wage\ bill_{it}$: the wage bill of a club for a particular season
- $\overline{wage\ bill}_t$: the average wage bills across all clubs for a particular season
- $individual\ appearances_{it}$: the number of players who made at least one appearance for each club in a particular season
- $\overline{individual\ appearances}_t$: the average number of players that made at least one appearance for each club in a particular season
- $black\ appearances_{it}$: the total number of appearances made by black players for each club in a particular season
- $total\ black\ appearances_t$: the total number of appearances made by black players by premier league clubs in a particular season
- $diff\ average\ black\ appearances_{it}$:
$$appearances\ of\ black\ players_{it} - \overline{appearances\ of\ black\ players}_t$$

⁵ This transformation derives from Szymanski (2000) and is the log odds of position. The denominator reflects the fact that there are 20 teams that participate in any given season of the Premier League.

- *diff average wage*_{it}: $wage_{it} - \overline{wage_t}$
- *diff average individual appearances*_{it}:
 $individual\ appearances_{it} - \overline{individual\ appearances_t}$
- *proportion minutes black*_{it}: the proportion of minutes accounted for by black players for each club in a particular season
- $\overline{proportion\ minutes\ black_t}$: the average proportion of minutes given to black players by premier league clubs over the course of a season

Appendix B

Data Sources

Player Data

1. worldfootball.net. (2021). from <https://www.worldfootball.net>

Wage Data⁶

1. telegraph.co.uk. (2021) from <https://www.telegraph.co.uk>
2. theguardian.com. (2021) from <https://www.theguardian.com/uk>
3. talksport.com. (2021) from <https://talksport.com>
4. planetfootball.com. (2021) from <https://www.planetfootball.com>
5. totalsportek.com. (2021) from <https://www.totalsportek.com>

⁶ All wage data originates from Companies House and can be accessed from the website <https://www.gov.uk>