Referee Bias

Authors: Jonathan Kalodimos*

Affiliations

*Correspondence to: Jonathan Kalodimos at jonathan.kalodimos@gmail.com

†This is the work of Jonathan Kalodimos as an unaffiliated researcher. All opinions and analysis are my own.

Abstract

In my analysis I do not find a statistically significant conditional correlation between skin tone and the issuance of red cards. I caveat this result with the primary intent of this research is to participate in the "crowdstorming exercise" and not to establish the existence or determine the causality of referee bias.

One Sentence Summary

Using a linear probability model I do not find a statistically significant conditional correlation between skin tone and the issuance of red cards.

Results

The central questions in this project are to determine if soccer referees are more likely to give red cards to dark skin toned players than light skin toned players, and if so is that a function of the skin tone prejudice of the referee's home country. The data provided for analysis is *not* sufficient to determine if this bias exists, and if it does what the magnitude is. The root of the problem is confounding variables that bias any potential correlation between referee bias and the likelihood of issuing a red card. As such it is difficult to draw conclusions from the analysis to follow, but my conclusion is that the color of a player's skin does not affect the likelihood of being issued a red card.

The likelihood of a player being issued a red card is a function of the probability of a player committing an infraction[1], the probability of a player being observed committing that infraction conditional on committing an infraction, and the probability of a referee issuing a red card conditional on observing an infraction. Taking the probability of a player committing an infraction as constant, a referee's bias could manifest in a change in the effort the referee devotes to observing a player's actions, or as a change in the leniency of the referee in enforcing the rules conditional on observing an infraction. With the current data I am unable to disentangle these two ways a referee's bias could manifest.

Further complicating the analysis is that the probability of a player committing an infraction is not constant but rather a function of the likelihood of him being caught and subsequently being issued a red card. If referee bias does exist, then this source of endogeneity will temper or even reverse the magnitude of any observable red card bias as players who expect to be caught and issued a red card more often will commit less infractions.

I am taking the stand that a referee views skin color as an informative signal of the likelihood of a player committing a red card worthy infraction. This may be an immutable erroneous prior of the referee i.e. there is no information in this particular signal, or it could be a correct belief in that skin color acts as a visible signal of the style of play of a particular player in a game. Either way the relationship between the probability of a player being issued a red card and skin color will be weakened as the referee observes the player through multiple interactions and updates his priors based on his observations. To mitigate this effect I restrict the sample to player-referee dyads that are the result of one interaction between the player and the referee.

Since the unit of analysis is player-referee observations and not player-game-referee observations I ignore any changes in team strategy due to opponent pairings. This is clearly an unrealistic assumption, but given the data available it is an assumption I have to make.

My primary empirical analysis is based on a linear probability model where the incidence of a red flag is a function of the color of a player's skin, and depending on the specification, player characteristics such as height, weight, age, and club fixed effects. I use an indicator variable for a particular rater's skin tone classification in order to account for any non-linearities. Table 1 presents the results using the classifications of rater 1 in columns 1 and 2, and of rater 2 in

[1] I assume that referees do not give red cards if a player does not commit an infraction.

columns 3 and 4. *t*-statistics are clustered at the referee country level to account for any common training in referees.

The indicator for category 1 (darkest skin tone) is the difference in the probability between the lightest skin player and the darkest skin player being issued a red card. This difference is not statistically significant at conventional levels for any of the specifications. As such I do not have a basis to reject the null that referee's are unbiased i.e. the skin tone of a player does not result in a change in the marginal probability of a player being issued a red card. In this analysis I use a linear probability model because in this context indicator variables for club and position will remove unobserved category-invariant characteristics that will bias my estimates. This is not necessarily true for non-linear models.

The facilitate comparison across research teams the organizing authors requested an odds ratio be constructed. I construct an odds ratio based on the odds of a category 1 (darkest skin tone) player receiving a red card relative to the odds of a category 0 (lightest skin tone) player receiving a red card. I do this by first predicting probabilities of each player receiving a red card based on the player's characteristics. I then take the mean probability of a category 1 player receiving a red card to calculate the conditional probability of category 1 player receiving a red card. I perform a similar exercise with a category 0 (the omitted category in the analysis which categorizes the lightest skin toned players) to calculate the conditional probability of category 0 player receiving a red card. With these two conditional probabilities I and their conjugates I calculate an odds ratio of 1.283 using the following formula.[2]

$$\frac{\overline{\hat{P}(red\ card \mid dark)} \ / \ \overline{\hat{P}(no\ red\ card \mid dark)}}{\overline{\hat{P}(red\ card \mid light)} \ / \ \overline{\hat{P}(no\ red\ card \mid light)}}$$

To address if referee bias is only observed in areas with high implicit bias score (high IAT) I interact a dummy variable for a referee from a country with an above mean IAT score with the categorical indicator for a player's skin tone. If referee bias is only prominent in referees from high IAT countries then this interaction term will measure that effect. Table 2 presents the results of this analysis using classifications for rater 1 in columns 1 and 2, and for rater 2 in columns 3 and 4. These results do not support the conclusion that referee bias is more prevalent in referees with high IAT.

I follow the same procedure as to address if referee bias is only observed in areas with high explicit bias score. I create a dummy variable (high Exp) that denotes a referee from a country with an above mean explicit bias score. I then interact that with the categorical indicator for a player's skin tone. These results do not support the conclusion that referee bias is more prevalent in referees with high IAT.

Tables 4 through 6 repeat the analysis of Tables 1 through 3 but do not use "club" as an explanatory variable. This was at the request of the organizers of the project and is the result of

[2] Mechanically this method of calculating the odds ratio is invariant to the model. This is due to using a linear probability model as my method of analysis.

some concern over the construction of the variable. Omitting the club fixed effect does not change the conclusions of this analysis.

Change in Analysis Strategy

Between the first and the second round I did not change my analysis strategy. I have stronger concerns about the endogeneity of player strategy than the authors of the comments I received, and as such I stand by the removal of observations where the player and the referee have met more than once. Additionally while a linear probability model obviously does not meet the constraints of being bounded by 0 and 1, linear probability models do have the nice property of indicators removing category invariant characteristics. This is not necessarily true of non-linear models.

Conclusion

Table 1: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.

Rater 1 = 0.25	0.00047	0.00037	0.00009	0.00000
	(1.17)	(0.95)	(0.17)	(0.00)
Rater 1 = 0.50	0.00119	0.00110	-0.00001	-0.00009
	(1.14)	(1.05)	(-0.01)	(-0.08)
Rater 1 = 0.75	0.00114	0.00096	0.00130	0.00112
	(1.20)	(1.07)	(1.33)	(1.20)
Rater 1 = 1 (Darkest)	0.00072	0.00062	0.00036	0.00026
	(0.72)	(0.61)	(0.42)	(0.30)
Club Fixed Effects	Yes	Yes	Yes	Yes
Player Characteristics	No	Yes	No	Yes
Players	62,310	62,310	62,310	62,310
R^2 (Adjusted)	0.00038	0.00040	0.00037	0.00039

In conclusion, the evidence does not support that there is an economically meaningful bias in a referee issuing red cards based on skin tone.

Tables

Table 2: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone, an indicator for above mean implicit bias score, and their interactions, plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.

Rater 1 = 0.25	0.00038	0.00026	-0.00042	-0.00051
	(1.04)	(0.70)	(-0.57)	(-0.64)
Rater 1 = 0.50	0.00004	-0.00009	-0.00099	-0.00108
	(0.05)	(-0.11)	(-0.70)	(-0.80)
Rater 1 = 0.75	0.00208**	0.00188*	0.00313***	0.00291***
	(1.98)	(1.89)	(3.46)	(3.30)
Rater 1 = 1 (Darkest)	0.00024	0.00010	-0.00072	-0.00084
	(0.17)	(0.07)	(-0.81)	(-0.95)
Rater 1 = 0.25 * High IAT	0.00022	0.00029	0.00119	0.00123
	(0.25)	(0.33)	(1.09)	(1.12)
Rater 1 = 0.50 * High IAT	0.00261	0.00270	0.00220	0.00225
	(1.26)	(1.32)	(0.98)	(1.00)
Rater 1 = 0.75 * High IAT	-0.00224	-0.00219	-0.00439***	-0.00428***
	(-1.31)	(-1.28)	(-3.13)	(-3.09)
Rater 1 = 1.00 (Darkest) * High IAT	0.00117	0.00127	0.00276	0.00280
	(0.63)	(0.69)	(1.46)	(1.49)

Table 2: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone, an indicator for above mean implicit bias score, and their interactions, plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.				
High IAT	0.00037	0.00031	0.00001	-0.00004
	(0.59)	(0.49)	(0.01)	(-0.06)
Club Fixed Effects	Yes	Yes	Yes	Yes
Player Characteristics	No	Yes	No	Yes
Players	62,183	62,183	62,183	62,183
R^2 (Adjusted)	0.00039	0.00041	0.00049	0.00050

Table 3: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone, an indicator for above mean explicit bias score, and their interactions, plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.				
Rater 1 = 0.25	0.00017	0.00005	-0.00057	-0.00066
	(0.47)	(0.12)	(-0.76)	(-0.81)
Rater 1 = 0.50	0.00067	0.00054	-0.00055	-0.00063
	(0.55)	(0.44)	(-0.36)	(-0.42)
Rater 1 = 0.75	0.00143	0.00122	0.00291***	0.00269***
	(1.23)	(1.11)	(2.94)	(2.80)
Rater 1 = 1 (Darkest)	0.00092	0.00078	-0.00028	-0.00040
	(0.64)	(0.54)	(-0.31)	(-0.45)
Rater 1 = 0.25 * High Exp	0.00068	0.00075	0.00155	0.00158
	(0.78)	(0.78)	(1.38)	(1.40)
Rater 1 = 0.50 * High Exp	0.00122	0.00130	0.00129	0.00131
	(0.61)	(0.65)	(0.57)	(0.58)
Rater 1 = 0.75 * High Exp	-0.00069	-0.00062	-0.00398***	-0.00385***
	(-0.37)	(-0.34)	(-2.76)	(-2.71)
Rater 1 = 1.00 (Darkest) * High Exp	-0.00047	-0.00036	0.00173	0.00179
	(-0.26)	(-0.20)	(0.91)	(0.94)

Table 3: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone, an indicator for above mean explicit bias score, and their interactions, plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.				
High Exp	-0.00003	-0.00009	-0.00039	-0.00044
	(-0.05)	(-0.15)	(-0.53)	(-0.59)
Club Fixed Effects	Yes	Yes	Yes	Yes
Player Characteristics	No	Yes	No	Yes
Players	62,183	62,183	62,183	62,183
R^2 (Adjusted)	0.00032	0.00034	0.00043	0.00045

Table 4: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.				
Rater 1 = 0.25	0.00071*	0.00063	0.00039	0.00030
	(1.82)	(1.62)	(0.80)	(0.58)
Rater 1 = 0.50	0.00167*	0.00164*	0.00072	0.00070
	(1.76)	(1.72)	(0.70)	(0.68)
Rater 1 = 0.75	0.00131	0.00126	0.00150*	0.00142
	(1.49)	(1.44)	(1.66)	(1.60)
Rater 1 = 1 (Darkest)	0.00097	0.00091	0.00058	0.00055
	(1.00)	(0.95)	(0.64)	(0.61)
Club Fixed Effects	No	No	No	No
Player Characteristics	No	Yes	No	Yes
Players	62,310	62,310	62,310	62,310
R^2 (Adjusted)	0.00002	0.00004	-0.00002	0.00000

Table 5: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone, an indicator for above mean implicit bias score, and their interactions, plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.				
Rater 1 = 0.25	0.00066*	0.00056	-0.00009	-0.00017
	(1.70)	(1.49)	(-0.13)	(-0.24)
Rater 1 = 0.50	0.00044	0.00039	-0.00037	-0.0040
	(0.65)	(0.57)	(-0.34)	(-0.37)
Rater 1 = 0.75	0.00226**	0.00220**	0.00336***	0.00325***
	(2.39)	(2.33)	(4.14)	(4.04)
Rater 1 = 1 (Darkest)	0.00048	0.00040	-0.00055	-0.00058
	(0.35)	(0.29)	(-0.56)	(-0.61)
Rater 1 = 0.25 * High IAT	0.00008	0.00013	0.00105	0.00107
	(0.10)	(0.16)	(0.99)	(1.01)
Rater 1 = 0.50 * High IAT	0.00274	0.00279	0.00233	0.00235
	(1.40)	(1.43)	(1.12)	(1.13)
Rater 1 = 0.75 * High IAT	-0.00231	-0.00230	-0.00456***	-0.00449***
	(-1.36)	(-1.35)	(-3.24)	(-3.24)
Rater 1 = 1.00 (Darkest) * High IAT	0.00119	0.00127	0.00278	0.00280
	(0.63)	(0.67)	(1.46)	(1.48)
High IAT	0.00037	0.00035	0.00006	0.00003
	(0.60)	(0.58)	(0.09)	(0.05)

Club Fixed Effects	No	No	No	No
Player Characteristics	No	Yes	No	Yes
Players	62,183	62,183	62,183	62,183
R^2 (Adjusted)	0.00004	0.00006	0.00010	0.00013

Table 6: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone, an indicator for above mean explicit bias score, and their interactions, plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.				
Rater 1 = 0.25	0.00045	0.00035	-0.00021	-0.00030
	(1.13)	(0.87)	(-0.31)	(-0.40)
Rater 1 = 0.50	0.00106	0.00101	0.00006	0.00004
	(0.94)	(0.89)	(0.05)	(0.04)
Rater 1 = 0.75	0.00158	0.00152	0.00310***	0.00299***
	(1.52)	(1.46)	(3.46)	(3.38)
Rater 1 = 1 (Darkest)	0.00107	0.00098	-0.00013	-0.00017
	(0.75)	(0.70)	(-0.14)	(-0.18)
Rater 1 = 0.25 * High Exp	0.00059	0.00064	0.00138	0.00139
	(0.71)	(0.75)	(1.25)	(1.26)
Rater 1 = 0.50 * High Exp	0.00141	0.00145	0.00147	0.00146
	(0.73)	(0.76)	(0.70)	(0.70)
Rater 1 = 0.75 * High Exp	-0.00066	-0.00064	-0.00403***	-0.00394***
	(-0.36)	(-0.35)	(-2.78)	(-2.76)
Rater 1 = 1.00 (Darkest) * High Exp	-0.00024	-0.00016	0.00185	0.00188
	(-0.13)	(-0.09)	(0.96)	(0.99)
High Exp	-0.00003	-0.00006	-0.00031	-0.00033
	(-0.05)	(-0.09)	(-0.43)	(-0.47)

Table 6: This table presents the results of regressing an indicator for a player being issued a red card on indicators for the category of skin tone, an indicator for above mean explicit bias score, and their interactions, plus control variables. *t*-statistic are in parenthesis and clustered at the referee's country level.

Club Fixed Effects	No	No	No	No
Player Characteristics	No	Yes	No	Yes
Players	62,183	62,183	62,183	62,183
R^2 (Adjusted)	-0.00004	-0.00002	0.00004	0.00006