Economic Inquiry

OWN-NATIONALITY BIAS: EVIDENCE FROM UEFA CHAMPIONS LEAGUE FOOTBALL REFEREES

BRYSON R. POPE and NOLAN G. POPE*

We examine the existence and magnitude of own-nationality bias. Using player-match level data from 12 seasons of the UEFA Champions League (UCL) and referee assignment policies that pair players and referees from the same country, we determine the bias that referees exhibit toward players from their native country. Players officiated by a referee from the same country receive a 10% increase in beneficial foul calls. Referees' own-nationality bias is more pronounced for national team players, players at home, and in later stages of the tournament. Elite referees exhibit as much, or more, own-nationality bias as their less experienced counterparts. (*JEL* L83, J15)

I. INTRODUCTION

Athletes continually search for ways to obtain even the slightest advantage over their opponents. It is often these slim margins that make the difference in competitive games and leagues. League officials have the responsibility to organize and monitor leagues such that no unfair advantage is gained by any particular player or team. Similarly, referees in football, and in most sports, are hired precisely for the purpose of officiating matches with objectivity, and to provide a fair playing environment. However, sometimes referees, whether consciously or subconsciously, exhibit a systematic bias for or against players due to their race, gender, nationality, or other characteristics.

In this paper, we address whether own-nationality bias is exhibited by referees toward professional football (soccer) players in the UEFA Champions League (UCL). We also explore how different types of players and settings affect the magnitude of the own-nationality bias. These players and settings include national team players, players at home, matches officiated by the most elite referees, and in later stages of the tournament. Lastly, we look at how much, if any, of the identified own-nationality bias is due to cultural or linguistic similarities.

The data consist of 12 seasons of player-match specific information from the UCL, which is one of the most prestigious professional club football tournaments in the world and brings in over €1.1 billion in annual revenue. The UCL tournament consists of 10 months of matches with 76 of the best club teams in Europe. It provides a window into high stakes, competitive, and scrutinized interactions between 402 of the most highly trained football referees and 4,294 of the best football players in the world, all from a total of 105 countries.

UEFA assigns referee squads composed of 4–6 referees from the same country to tournament matches such that a referee squad's nationality is different than that of the nationality of either of the two teams being officiated. This method of assignment allows for our identification strategy to detect whether referees exhibit favoritism toward players from their same country, or in other words, own-nationality bias. Although no referee squad and team are from the same country, oftentimes there are

ABBREVIATIONS

CAGE: Cultural, Administrative, Geographic and Economic
GDP: Gross Domestic Product
NBA: National Basketball Association
OLS: Ordinary Least Squares
UCL: UEFA Champions League
UEFA: Union of European Football Associations

*We would like to thank Brent Hickman, Grant Gannaway, Joseph Price, Ryan Hill, Steven Levitt, Nathan Petek, and Mary Li for their helpful comments and discussion. We also thank Henrik Kleven, Camille Ladais, and Emmanuel Saez for use of their European football player database.

Pope, B. R.: Department of Economics, Brigham Young University, Provo, UT 84601. Phone 801-854-8260, Fax 805-893-8830, E-mail brysonrpope@gmail.com

Pope, N. G.: Department of Economics, University of Chicago, Chicago, IL 60615. Phone 801-995-9184, E-mail npope@uchicago.edu

Economic Inquiry
(ISSN 0095-2583)

doi:10.1111/ecin.12180

individual players who are from the same country as the referee squad. These same country pairings of referee squads and individual players allow for a within-player analysis of referees' own-nationality bias. For a specific player across games, we compare the fouls called by a referee squad from his home country to the fouls called by a referee squad not from his home country.

We find that players in a same country referee-player pairing receive more favorable foul calls than when they are not in a same country referee-player pairing. On average, having a referee squad from the same country increases the number of beneficial foul calls received by a player by about 10%. The amount of own-nationality bias exhibited by referees is more pronounced (15–20%) for national team members, when players are playing at their home venue, for the highest qualified referees, and during the later stages of the tournament. We also find that the own-national bias found cannot be solely attributed to similarities in language or culture of referees and players from the same country.

There are two particularly pertinent studies that analyze racial bias in athletics. Price and Wolfers (2010) show that in the National Basketball Association (NBA), more personal fouls are called against players when they are officiated by an opposite-race refereeing crew than when they are officiated by an own-race refereeing crew. Also, Parsons et al. (2011) find that an umpire calls fewer strikes when the pitcher is of a different race than the umpire. They also show that this is particularly true when there is low scrutiny of the officiating. When a computerized camera system is in place, which allows pitches to be shown electronically on television and online, this bias is nearly eliminated. The bias also decreases with high attendance and later in the game. It appears that baseball pitchers attempt to compensate for the bias by throwing less subjective pitches such as fast-balls over home plate in order to give the umpire less of an opportunity to discriminate.

As so much research has been done on racial bias in many different settings, it is notable that little research has been done on own-nationality bias. With the intense patriotism exhibited by passionate fans, football is a likely place for own-nationality bias to be demonstrated. Football hooligans are, in fact, stereotypically known for their ardent support of their team of choice. The frequent violent after-match events such as killings, beatings, and taunting of opposing team supporters indicate that at least the fans, at times, act on these biases. The ubiquitous nature of football riots and on-the-pitch inter-player displays of discrimination, such as players being fined for shouting racially derogatory remarks toward opposing players, raises the question of whether referees also show particular forms of bias toward or against certain players. We hypothesize that just as basketball referees and baseball umpires exhibit own-race bias in their calls, we will observe professional football referees exhibiting an own-nationality bias toward players with the same nationality as themselves by awarding them with more beneficial calls.

Although racial bias has been given much wider attention by the literature than that of own-nationality bias, Zitzewitz's (2006) research on Olympic events is a notable exception. He finds that Olympic figure skating and ski jumping judges display biases in favor of athletes from their same country. He notes that these own-nationality biases can be quite large. Using data from the 2002 Winter Olympics he finds that both figure skating and ski jumping judges score their compatriots about 0.13 standard deviations higher than other judges.

The rest of the paper will proceed as follows. Section II gives a detailed description of the UCL dataset as well as the specifics of the UCL tournament. Section III gives an in-depth discussion of our identification strategy and model. Section IV reports the results along with a discussion of why such own-nationality biases occur. Section V concludes.

II. DATA AND INSTITUTIONAL BACKGROUND

Our main dataset is player-match information drawn from the Union of European Football Associations' (UEFA) official website.[1] The dataset consists of matches from the UCL which is the most lucrative club football competition in the world with an estimated gross commercial revenue of €1.1 billion.[2] Different from international football in which each team consists of players all from the same country, club football teams consist of players from all around the globe. In the UCL, nearly half of all players are not from the country in which the club team they play for is located. In addition to on-site spectators, the most recent 2012–2013 UCL tournament attracted more than 360 million television

1. http://www.uefa.com/uefachampionsleague

2. http://www.uefa.org/management/finance/news/newsid=1840934.html

viewers.[3] The tournament has been held annually since 1955 for the top European football clubs from 54 countries.[4] The UCL lasts the better part of a year (from July to May), and consists of qualification rounds, a 32-team group stage, and a 16-team knockout stage.

The formal tournament begins with 32 teams. Of these teams, 22 earn an "automatic entrant" status by their high ranking in their home country league. The remaining 10 spots in the tournament are selected from 54 teams that participate in three qualifying rounds, with the lower seeded league teams entering in the first round and the higher seeded league teams entering in the later rounds. These 32 teams are divided into eight groups of four to begin the group stage. The group stage proceeds in a round-robin format with every team playing the other three teams twice, both at home and away. Each win is awarded three points, each draw is awarded one point, and each loss is awarded zero points. The two teams with the most points awarded advance to the knockout stage where the top seed from each group is matched against a second-seed team from another group.[5] The round of 16, quarterfinals, and semifinals are identical in setup with two matches played between the two teams, one home game and one away game for each team. The team with the highest aggregated score from the two matches moves to the next round. In contrast, the finals match is a one-match affair played at a predetermined neutral location.

Teams have strong incentives to do well in the UCL tournament. For one, the prestige and even fame that comes from winning this tournament may result in greater ability to recruit the top footballers.[6] The financial incentives are also significant. In 2012, teams in the 32-team group stage were guaranteed at least €7.2 million. For each additional advanced round, teams are guaranteed to earn an additional €3–€9 million with the winner receiving a guaranteed total of €26.7 million. Besides each team's guaranteed allotment, each team earns additional money depending on the proportion of the broadcasting revenue from within the territory of their respective national associations. In 2012, teams in the 32-team group stage earned between €8.6 and €62.9 million with an average of €24.8 million per team.[7] In general, the more games a team plays on television and wins, the more prize money the team will earn throughout the tournament.

The dataset from UEFA's websites provides detailed match and player-specific data for every match in the UEFA Champions League for the 2001–2002 to 2012–2013 seasons. For each match, each team lists 18 players on their roster who are eligible to play in that match. For each 90-minute match, 11 starters begin on the pitch and each team is allowed a maximum of three player substitutions. The dataset contains information for all players on the roster for each match. Player information includes whether the player started or was a substitution and how many minutes the player was on the pitch. These data also include individual statistics on the number of fouls committed, fouls suffered (fouls committed against the player by the opposing team), cautionary cards, offside calls, goals scored, assists, shots on target, and shots off target. Columns 1 and 2 of Table 1 give the individual field players' (excluding goalkeepers) summary statistics for those who played at least 30 minutes. The first variable in Table 1, the difference in fouls suffered and fouls committed, shows the net beneficiary calls obtained by a player. A positive foul difference means the player received more beneficial foul calls than disadvantageous calls. Over the 12-year period of these data, 2,563 matches were played with over 4,000 different football players. All statistics are also available at the team level.

Besides player and team information for each match, the UCL data also include the name and nationality of the center and sideline referees. All referees in a referee squad assigned to a match are from the same country. For example, the center referee, both sideline referees, and a fourth referee would all be from Norway. Figure 1 shows the fraction of matches with referee squads from a given country. German referee squads are the most common with Italian, Spanish, French, and English referee squads following close behind.

An additional dataset provides further player information by combining Kleven, Landais, and Saez's (2013) European football player database

3. Faisal Chishti. "Champions League final at Wembley drew TV audience of 360 million" (May 30, 2013). Sportskeeda. Absolute Sports Private Limited.

4. http://www.uefa.com/memberassociations

5. The one exception is that no team will play a runner up from their own country.

6. Pope and Pope (2009) show that at the collegiate level, universities receive more applications the years immediately following a strong showing in the NCAA March Madness tournament.

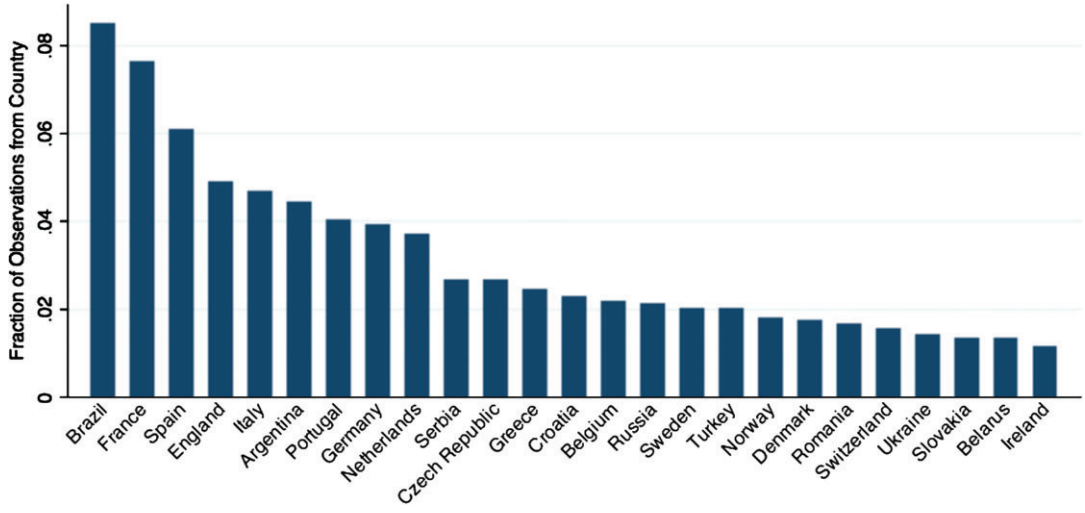
7. http://www.uefa.org/management/finance/news/newsid=1840934.html

TABLE 1
Summary Statistics

	All Observations		Paired Players				Difference in Means	
			Paired Matches		Unpaired Matches			
Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Paired - All	Paired - Unpaired
Suffered − committed	−0.06	1.62	0.13	1.80	−0.06	1.80	0.19	0.19
Fouls suffered	1.06	1.36	1.31	1.47	1.29	1.40	0.25	0.02
Fouls committed	1.12	1.34	1.19	1.31	1.34	1.43	0.07	−0.15
Yellow cards	0.16	0.38	0.14	0.37	0.16	0.38	−0.02	−0.02
Red cards	0.01	0.09	0.01	0.08	0.01	0.08	0.00	0.00
Offsides	0.21	0.64	0.26	0.77	0.31	0.79	0.05	−0.05
Goals scored	0.12	0.37	0.12	0.36	0.15	0.42	0.00	−0.03
Assists	0.07	0.27	0.11	0.33	0.10	0.32	0.04	0.01
Shots on target	0.42	0.80	0.49	0.87	0.56	0.95	0.07	−0.07
Shots off target	0.40	0.76	0.52	0.85	0.53	0.87	0.12	−0.01
Minutes played	83.16	13.91	83.34	13.21	83.91	12.93	0.18	−0.57
UCL games played	10.58	12.71	23.20	19.89	23.20	19.89	12.62	0.00
Number of players	4,294		259		259			
Number of matches	2,563		328		1,816			
Player-match obs	45,459		435		5,573			

Notes: All observations include all player-match observations from 2001 to 2013. Paired players are players who have played in at least one match with a referee squad from same country. Paired matches are matches in which the player has a referee squad from the same country. Suffered minus committed are the number of fouls committed against a player by the opposing team minus the number of fouls the player commits against the opposing team.

FIGURE 1
Fraction of Player Observations from Top 25 Countries

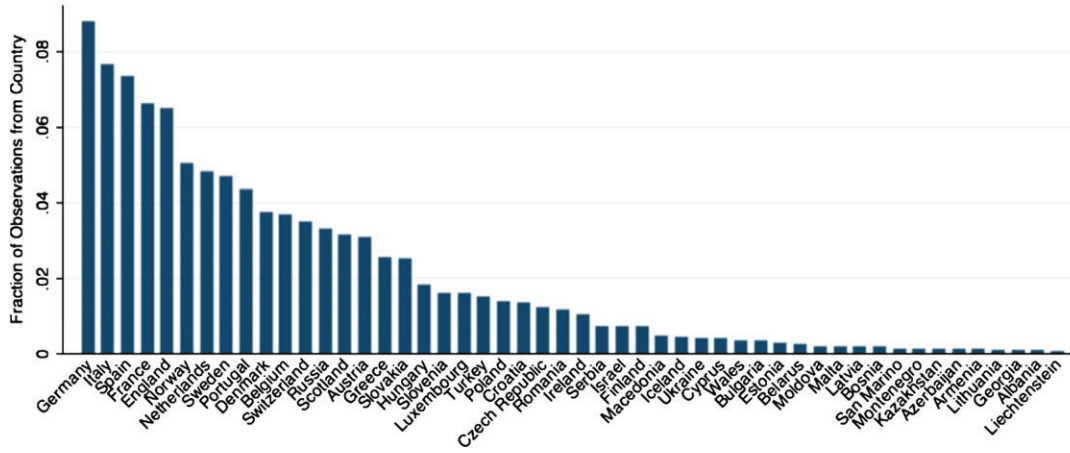


with UEFA's football player database. This combined data provide crucial information on individual player's nationality along with their age and whether they played for their national team or not. Figure 2 shows the fraction of players from a given country for the 25 countries with the most players in the UCL. Of the top 25 countries, all but Brazil and Argentina are European countries.

Lastly, we use the Cultural, Administrative, Geographic and Economic (CAGE) similarity index designed by Pankaj Ghemawat.[8] The

8. http://www.ghemawat.com/cage/

FIGURE 2
Fraction of Referee Observations from Each Country



CAGE index was designed to measure how similar two countries are with the goal of informing business managers how easy or difficult it is to conduct business between the two countries. The CAGE index uses 16 factors such as language, religion, trade, distance, and gross domestic product (GDP) to relate 162 countries with each other. This index gives a measure of more than 99% of the referees' and players' countries in our UCL dataset. The individual factors that the index takes into account to relate referees' and players' countries and the index itself allow us to determine if the identified own-nationality bias is being driven by cultural, linguistic, or other similarities between countries.

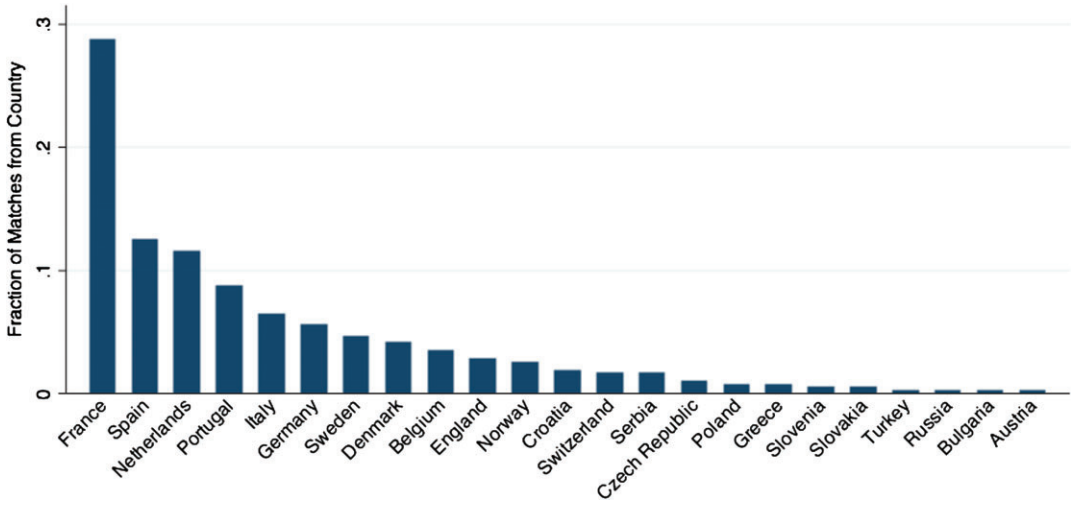
III. METHODOLOGY

Our empirical strategy relies heavily on how referee squads are assigned to matches. For each stage of the tournament an administrative committee known as the UEFA Refereeing Unit creates a proposal of referee assignments for each match. This proposal is then discussed, revised, and finalized by the UEFA Referee Committee. The two committees use observer marks, performance, personality, availability, development, administrative factors, and successful completion of the relevant written and fitness tests to determine referee assignment.[9] In addition to the two committees, there are two guiding policies followed by UEFA when assigning referees. First, to discourage bias, no referee squad can be from the same country as either of the clubs. For example neither an English or Spanish referee squad would be assigned to a Manchester United vs. Barcelona match. Second, referee assignments remain confidential up to 2 days before the match for the purpose of minimizing public influence.

Our identification strategy exploits the fact that a referee squad is never assigned to officiate a club team from the same country as the referee squad. However, while officiating clubs from countries other than their own, referee squads may at times officiate a team with a player from their home country. That is to say there are no same country referee-team pairings but there are frequent same country referee-player pairings. Our basic strategy compares a player who is officiated by a referee squad from his home country to himself when he is officiated by a referee squad not from his home country. For example, if a Portuguese player playing on a team based in Spain had a match against a team in Germany, it is possible that a referee squad from Portugal could be assigned to officiate the match. That specific player would then have a referee squad from his own country, whereas all other non-Portuguese players would not. In this example, we designate this player as a "paired player" since he had at least one match with a referee squad from his home country and designate the match as a

9. http://www.uefa.org/MultimediaFiles/Download/Tech/uefaorg/General/01/89/25/77/1892577_DOWNLOAD.pdf

FIGURE 3
Fraction of Player-Referee Pairings from Each Country



"paired match" since the referee squad is from his home country. Figure 3 gives the proportion of referee-player pairings from each country. The largest portion of referee-player pairings consists of a player and referee squad both of whom are from France.

The ideal identification strategy would be if referees were assigned randomly to matches. If referee assignments were random, then referees' own-nationality bias could easily be identified by taking the difference in beneficial calls between matches when a player is assigned an own-nation referee and matches when the same player is not assigned an own-nation referee. Although referee assignment is not strictly random, the information used by UEFA to make referee assignment is unrelated to pairing players and referees from the same nation but rather uses information on other factors such as referees' availability, fitness tests, and performance level. Although referee assignment is not random, it is likely random with regard to how it assigns referees to players from their same nation. In addition, because we compare players with a same nation referee to themselves without a same nation referee by using individual fixed effects, it would take a special type of selection in the referee assignment process to explain the results of referees exhibiting own-nationality bias. For example, in order for the referee assignment process to explain the results, referees would have to be assigned in a very particular and unlikely manner. For a selection bias to exist referees would have to be assigned to matches with a paired player in which the paired player was well behaved and committed less fouls and players on the opposing team were ill behaved and committed more fouls against the paired player. This type of selection in the referee assignment process would require the UEFA committees to know ex ante in which games a player would be well behaved and commit fewer fouls and the other team would be relatively ill behaved and commit more fouls. In addition to knowing these two things ex ante, the UEFA committees would also need to make referee assignments according to this information. Although this type of referee assignment is possible, it seems unlikely to be the case.

Table 1 shows a simple outline of the identification strategy to be used along with results from differences in group means. Columns 3 and 4 are the player-match summary statistics for players in matches in which the player and referee squad are from the same country (i.e., paired players in paired matches). Columns 5 and 6 are the player-match summary statistics for players who had at least one match with a same country referee squad, but in matches in which the player and referee squad are not from the same country (i.e., paired players in unpaired matches). Column 7 displays the difference between variable means for paired players in paired matches and all player-match observations. The paired players' foul difference (suffered minus committed)

is 0.19 fouls higher than the foul difference for all observations. This corresponds to roughly an 8% increase in beneficial foul calls for players with a referee squad from the same country. There is no such benefit for cautionary cards or offside calls. This difference between paired players in paired matches and all player-match observations in column 7 could be driven by both the beneficial effect of having an own-nation referee and by any underlying differences between paired and unpaired players. Column 8 shows the difference between the mean of each variable for paired players in paired matches and paired players in unpaired matches. The results of this comparison are nearly identical to column 7, with an 8% increase in beneficial foul calls and no change in cautionary cards or offside calls. The difference between paired players in paired matches and unpaired matches is no longer driven by any underlying differences between paired and unpaired players and is just the effect of having an own nation referee. Because the differences in columns 7 and 8 were so similar it appears that the underlying differences between paired and unpaired players are small. However, the difference shown in column 8 is unbalanced and therefore the below analysis will implement a within-player analysis using player fixed effects.

Formally, the model to be used is as follows:

$$(1)\quad Y_i = \alpha + \beta \text{samecountry}_i + \delta X_i + \gamma M_i + \varphi H_i + \pi A_i + \mu_i + \eta_i + \nu_i + \theta_i + \varepsilon_i$$

where Y_i is the outcome of interest (e.g., foul difference, fouls committed, fouls suffered, disciplinary cards, etc.) for a specific match-player observation i. Our main outcome of interest is foul difference, which is calculated as the number of fouls suffered by a player minus the number of fouls committed by that player. A bigger foul difference implies a larger number of beneficial foul calls for that player. The variable samecountry_i is a binary variable equal to one if the player and referee squad are from the same country and zero if not. Additionally, X_i is a vector of player characteristics including age, age squared, minutes played, and whether the player started the match. The vector M_i contains match characteristics including year of match fixed effects, month of match fixed effects, and tournament round fixed effects. The vector H_i contains the player's own-team match statistics of the number of goals scored, shots on target, shots off target, and corner kicks. The vector A_i contains similar opposing-team match statistics. The variables μ_i, η_i, ν_i, and θ_i are player, referee, own-team, and opposing-team fixed effects, respectively. No measures of player aggressiveness, such as tackles, are available in these data to allow us to analyze how playing style similarities between players and referees from the same country might affect the results.

IV. RESULTS

Table 2 shows the estimates of the amount of own-nationality bias in referees' foul calls using various specifications of the model in Equation (1). The outcome variable for Table 2 is the foul difference (fouls suffered minus fouls committed). For all results, unless otherwise specified, estimates are for players who played at least 30 minutes in the match excluding goalkeepers. Column 1 has no controls and is simply the difference between paired players in paired matches to all players in unpaired matches. This is the same value as the "paired minus all" column in Table 1 of 0.187. This implies that players from the same country as the referee squad receive roughly one-fifth more of a beneficial foul call per match. Compared to the base of 2.18 foul calls per player per match this is an 8.5% increase in beneficial foul calls. This is equivalent to a 0.12 standard deviation increase in beneficial foul calls. This estimate, however, could be bias due to systematic differences in paired players and unpaired players. Column 2 adds player fixed effects so the estimate is the difference in beneficial foul calls between a player who is officiated by a referee squad from his/her same home country and himself/herself when he/she is officiated by a referee squad not from his/her home country. With player fixed effects, the estimate of the amount of own-nationality bias in foul calls increases to 0.23 and is statistically significant at the 0.01 level. This is again roughly a 9% increase in beneficial foul calls.

Column 3 adds controls for player characteristics, own team statistics, opposing team statistics, and match information. Adding these controls causes virtually no change in the own-nationality bias in foul calls. Column 4 adds own team, opposing team, and referee fixed effects. Again there is little change in the point estimates. Column 5 includes a match fixed effect. In column 6 the sample is restricted to only those players who started the match. Column 7 uses the

TABLE 2
The Effect of a Same Country Referee on Foul Difference

	Foul Difference (Suffered − Committed)						
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Same country	0.187**	0.230***	0.224**	0.215**	0.220**	0.206**	0.234**
	[0.086]	[0.088]	[0.088]	[0.089]	[0.095]	[0.093]	[0.097]
Started	—	—	−0.05	−0.04	−0.04	—	−0.10
			[0.042]	[0.043]	[0.045]		[0.073]
Minutes played	—	—	0.002***	0.002***	0.002***	0.002***	0.004***
			[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Age	—	—	0.05	0.04	0.04	0.04	0.03
			[0.045]	[0.047]	[0.051]	[0.049]	[0.053]
Age^2	—	—	0.00	0.00	0.00	0.00	0.00
			[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Own team							
Goals	—	—	0.00	0.00	0.108**	0.00	0.00
			[0.008]	[0.008]	[0.052]	[0.008]	[0.009]
Shots on target	—	—	0.00	0.00	0.01	0.00	0.00
			[0.004]	[0.004]	[0.026]	[0.004]	[0.005]
Shots off target	—	—	0.014***	0.014***	0.047**	0.013***	0.016***
			[0.004]	[0.004]	[0.020]	[0.004]	[0.004]
Corners	—	—	0.01	0.00	−0.067**	0.01	0.00
			[0.004]	[0.004]	[0.028]	[0.004]	[0.004]
Opposing team							
Goals	—	—	0.021***	0.026***	0.136***	0.027***	0.028***
			[0.008]	[0.008]	[0.051]	[0.008]	[0.009]
Shots on target	—	—	−0.021***	−0.022***	−0.02	−0.022***	−0.022***
			[0.004]	[0.004]	[0.026]	[0.004]	[0.005]
Shots off target	—	—	0.00	0.00	0.03	0.00	0.00
			[0.004]	[0.004]	[0.020]	[0.004]	[0.004]
Corners	—	—	−0.013***	−0.014***	−0.085***	−0.014***	−0.015***
			[0.004]	[0.004]	[0.028]	[0.004]	[0.004]
Fixed effects							
Player FE		X	X	X	X	X	X
Tournament stage FE			X	X	X	X	X
Year FE			X	X	X	X	X
Month FE			X	X	X	X	X
Own team FE				X	X	X	X
Opposing team FE				X	X	X	X
Referee FE				X	X	X	X
Match FE					X		
Observations	45,455	45,455	45,401	45,401	45,401	43,204	45,401
R^2	0.00	0.18	0.18	0.20	0.21	0.21	0.20

Notes: For all specifications the outcome variable of interest is fouls suffered minus fouls committed (number of beneficial foul calls). The model used is shown in Equation (1). The sample is restricted to players who played at least 30 minutes for columns 1–5 and to starters for column 6. Column 7 weights players' foul difference by the number of minutes played. Standard errors are clustered at the match level and are reported in brackets.

***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

outcomes variable of foul difference weighted by 90 divided by the number of minutes played by the player so that the point estimates are to be interpreted as the number of beneficial foul calls per full match played for players of the same nationality as the referee. For example if a player only played half of the match and has a foul difference of 2, his/her weighted foul difference would be doubled to 4. Regardless of which specification is used, a player being officiated by a corresponding same home country referee squad receives an approximately 0.22 increase in beneficial foul calls during the match. Given that the average number of combined fouls committed and suffered per player per match is 2.2 for all players and 2.5 for paired players, this results in roughly a 9% increase in beneficial foul calls. Owing to the consistent nature of the estimates regardless of the specification chosen, the specification in column 4 will be used for all further analysis performed in this section. We denote this specification as the full model.

The 0.22 foul benefit in the foul difference consists of two parts: fewer fouls committed and more fouls suffered. Table 3 uses the full model to look at other outcomes besides foul difference including fouls committed and fouls suffered. Columns 1 and 2 display the estimates of the coefficient on the same country binary variable with fouls committed and fouls suffered as the outcome variable. The 0.22 foul benefits from a player being from the same country as the referee squad appears roughly equally split between the two parts. In each case the signs are in the expected direction, but are at best marginally significant. Columns 3 and 4 show that there is no own-nationality bias for offside calls or for cautionary (yellow or red) cards given. This could be the case for a number of reasons. One possibility is that there are fewer offside calls and substantially fewer cards given to an individual player each match compared to normal fouls. Another plausible reason is illustrated in the findings of Parsons et al. (2011) with regard to baseball umpires and pitchers. In evaluating same race match ups between pitchers and umpires, they find that when umpires are monitored under a system of computerized cameras used to evaluate the umpires ball/strike calls, evidence of any race or ethnicity preference disappears entirely. Under such explicit evaluation, umpires have strong incentives to suppress any bias that could be detrimental to their own career. With modern day replay it is usually clear whether the linesman made the correct offside call or not. Even with modern day replay, however, fouls and other infractions are far more subjective. Similar to umpires, football referees may be aware that offside calls will place them under explicit evaluation and give them a stronger incentive to suppress any own-nationality bias. It is also interesting to note that typically linesmen are less experienced referees whereas the most elite referees are assigned more frequently to be the center referee position.

In addition to looking at offside calls and cautionary cards, columns 5 and 6 look at whether team managers adjust their coaching decisions due to player-referee pairings and play paired players more. Again, Parsons et al. (2011) find that pitchers, either consciously or subconsciously, recognize the bias being exhibited by umpires and as a result, change their behavior in order to compensate. When bias is being exhibited the pitcher will throw the less subjective fastball over home plate more often than usual. It may be the case that football managers exhibit a similar change in behavior in order to compensate. Column 5 shows that there is no statistical increase in the number of minutes played when a player is paired with the referee squad. Column 6 also shows that players paired with the referee squad are no more likely to enter the match. It appears there is no substantial shift in coaching strategy by allowing a player who creates a player-referee pairing to play more minutes or be on the pitch in that particular match. This finding that managers do not change their coaching behavior due to player-referee pairings has at least two possible explanations. First, managers might be unaware of the beneficial foul calls received by paired players and therefore do not change their behavior. Second, managers may be aware of the beneficial foul calls received by paired players and are in fact behaving optimally, but believe that the benefit of getting more beneficial foul calls from a paired player is outweighed by the cost of playing a less able player. As paired players receive more beneficial foul calls but do not receive more beneficial calls on the more important card decisions, it is likely that having a more able player on the field is more influential to the outcome of the match than having a less able paired player. Later in Table 6, we consider the impact of having a paired player on match outcomes.

Table 4 shows estimates of own-nationality bias in beneficial foul calls using the full model for different types of players, matches, and referees. Column 1 restricts the sample to only those players who play for their home country's national team.[10] The national team is the top squad for any given country which usually consists of 18–22 players. These players are typically widely recognized by citizens of their country and are high-profile figures. When restricted to only national players, the own-nationality bias increases to 0.28 and is highly significant. Column 2 shows that the own-nationality bias completely disappears when the sample is restricted to players who are not national team members. This difference could be due to referees recognizing that national team players are from the same country as themselves, whereas they fail to recognize the players not on the national team to be from the same country as themselves. Alternatively, referees might recognize that all same

10. The national team consists of only those players who played on the adult national team. For example, the USA's U-21 team is technically a national team, however it would not classify as the national team in this context.

TABLE 3
The Effect of a Same Country Referee on Other Variables

Variable	Fouls Committed	Fouls Suffered	Offsides	Card	Minutes Played	Enters Match
Same country	−0.122*	0.093	0.001	−0.004	0.853	0.005
	[0.065]	[0.064]	[0.035]	[0.018]	[1.129]	[0.015]
Observations	45,401	45,401	45,401	45,401	53,469	67,132
R^2	0.374	0.403	0.378	0.155	0.359	0.328

Notes: The top of each column indicates the outcome variable of interest. The model used is the same as column 4 of Table 2 and is shown in Equation (1). The sample is restricted to players who played at least 30 minutes in the match. Standard errors are clustered at the match level and are reported in brackets.
***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

country players are from their home country, but there exists a closer nationality association between the referees and high-profile national team players than with nonnational team players.

A number of studies have shown that referees can be pressured by crowds to favor the home team (Buraimo, Forrest, and Simmons 2010; Garicano, Palacios-Huerta, and Prendergast 2005). Scoppa (2008) found that professional Italian referees added more injury time if home teams were losing. Sutter and Kocher (2004) and Dohmen (2008) found that, for the German Bundesliga, referees added more time in matches where the home team was behind. Columns 3 and 4 of Table 4 restrict the sample to players in matches at home and matches away from home. We find that the own-nationality bias is also larger for players at home than for players away from home. It appears that referees are more willing to act on their own-nationality favoritism when the home crowd is supportive of their decision.

Biases are often thought to attenuate or disappear in high stakes, competitive, and scrutinized situations or with highly trained individuals (Parsons et al. 2011). The UCL setting is a high stakes, competitive, and scrutinized situation. To assure the employment of the highest quality referees, UEFA systematically reviews how each referee performs. For every UEFA match played in the tournament, an observer is assigned to evaluate the performance of the referee. All observers are experienced former European referees. According to UEFA guidelines they must conduct an oral debriefing with the referee after the match and then produce a written report along with a score or mark between 0 and 10 to rate their performance. These referee observers' reports are kept confidential within the referee committee. The UCL referees are also highly trained and experienced. UEFA has four main categories of referees ranging from third class to Elite. Each increase in level is accompanied by a wage increase per match ranging from €100 to €3,000. To achieve Elite status a referee must be very experienced and have consistently performed well in the past. For example, for the 2012/2013 season there were a total of 36 elite referees. Among these referees, they averaged roughly 13 UEFA champion's league appearances each. They also averaged over 22 years as a professional referee and over 7 years with the FIFA badge which is required to be eligible to referee UEFA championship games. Using UEFA's referee categories, we are able to vertically differentiate referees. Typically it is assumed that better and more experienced referees show fewer biases. The results in column 5 and 6 of Table 4 show that the elite referees in fact exhibit as much or more own-nationality bias as non-elite referees. The own-nationality bias in foul calls appears to persist despite, or perhaps because of, the high stakes, competitive, and scrutinized situation of the UCL even among the most highly trained and experienced referees.

Club teams have large incentives to do well in the UCL. Teams that made it to the 32-team group stage of the tournament in 2012 on average earned €24.8 million and gained the prestige and recruiting power that comes with a strong showing in the tournament. Each club is rewarded with more and more money as they progress through the tournament. The benefit of performing well in each match for a team and player increases as the tournament progresses. The monetary benefit is nonlinear and the benefit in terms of prestige and recruiting power is also likely highly nonlinear with a much larger benefit toward the end of the tournament than the beginning. With a good performance later in the tournament being much more valuable than at the beginning, referees might demonstrate more own-nationality bias as the benefit of each advantageous call

TABLE 4
The Effect of a Same Country Referee by National Team, Home, and Elite Referee

Variable	National Team	Not National Team	Home	Away	Elite Ref	Not Elite Ref
Same country	0.285***	−0.005	0.326**	0.092	0.464***	0.183
	[0.102]	[0.204]	[0.131]	[0.135]	[0.165]	[0.112]
Observations	32,478	5,511	22,150	23,251	11,683	33,638
R^2	0.194	0.318	0.275	0.267	0.314	0.224

Notes: For all specifications the outcome variable of interest is fouls suffered minus fouls committed (number of beneficial foul calls). The model used is the same as column 4 of Table 2 and is shown in Equation (1). The sample is restricted to players who played at least 30 minutes in the match. The title at the top of each column indicates the subsample of the data on which the analysis is performed. Standard errors are clustered at the match level and are reported in brackets.

***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

TABLE 5
The Effect of a Same Country Referee by Stage in Tournament

Variable	Qualification Matches	Group Stage Matches	Knockout Matches
Same country	−0.008	0.177	0.428**
	[0.131]	[0.122]	[0.189]
Observations	13,914	23,140	8,347
R^2	0.347	0.25	0.289

Notes: For all specifications the outcome variable of interest is fouls suffered minus fouls committed (number of beneficial foul calls). The model used is the same as column 4 of Table 2 and is shown in Equation (1). The sample is restricted to players who played at least 30 minutes in the match. The title at the top of each column indicates the stage of the tournament on which the sub-analysis is performed. Standard errors are clustered at the match level and are reported in brackets.

***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

TABLE 6
The Effect of a Same Country Referee on Not Losing

	Not Lose			
Variable	(1)	(2)	(3)	(4)
Number of more pairs	0.094***	0.042**	0.026	0.042
	[0.016]	[0.018]	[0.027]	[0.066]
Own team FE		X		
Opposing team FE		X		
Own team-year FE			X	
Opposing team-year FE			X	
Matchup FE				X
Observations	2,254	2,254	2,254	2,254
R^2	0.011	0.341	0.767	0.855

Notes: The observation level is at the team-match level for all away teams. The number of more pairs indicates how many more same country player-referee parings the away team has compared to the home team. Not lose is a binary variable equal to one if the away team wins or ties. Standard errors are clustered at the match level and are reported in brackets.

***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

increases. In Table 5, the sample is restricted to three different stages of the tournament: qualification matches, the 32-team group stage, and the 16-team knockout stage. As the tournament progresses we observe increasingly strong evidence for the existence of own-nationality bias. This bias goes from nonexistent in the qualification matches to almost one-half of a beneficial foul call per match for paired players in the later knockout matches. This is roughly a 20% increase in beneficial foul calls.

Although we find significant evidence of referees exhibiting own-nationality bias, it is unclear whether this favoritism translates into more victories for the team with more paired players. It is also unclear if team managers should be adjusting their strategy owing to the own-nationality bias exhibited by referees. In order to determine if own-nationality bias influences match outcomes, we analyze match level data for all away teams for each match in the dataset. In Table 6 the outcome variable of interest is a binary variable equal to one if the away team wins or draws. The independent variable of interest is the number of paired players that the away team has more than the home team. It is important to note that in some games both teams have one or more referee-player pairs which could potentially give them an advantage. Of the 2,564 matches in our dataset, 250 matches have only one referee-player pair whereas 56 matches have two pairs, 1 match has three pairs, 5 matches have four pairs, and 1 match has five pairs.

Column 1 of Table 6 is a simple ordinary least squares (OLS) regression of whether the away team did not lose (win or tie) on the number of more paired players that the away team had compared to the home team. Column 1 implies a team

with one more paired player is 9.4 percentage points more likely to win or draw. Besides the estimate being implausibly high, it is clear that this estimate is biased upward due to the fact that the number of paired players on a team and team quality is likely positively correlated. The reason for this positive correlation is because only foreign players on a team can possibly be a paired player and richer and higher quality teams buy more foreign players. Specifications used in columns 2–4 help to mitigate some of this bias, but may or may not fully control for it. Column 2 includes own and opposing-team fixed effects and the estimate is cut in half. The correlation between the number of foreign players on a team and how rich and high quality a team is, is also likely to exist within a team over different years. Column 3 includes own and opposing-team-year fixed effects. The estimate is reduced to a 2.6 percentage point increase in the likelihood of not losing, but is statistically insignificant. Column 4 includes matchup fixed effects, which includes a fixed effect for each pair of two teams that play each other. Again the estimate is positive but not statistically significant.

In addition to looking at the own-nationality bias of referees toward players, in Table 7 we look at how cultural and language similarities between referees and players, and referees and teams, affect beneficial foul calls. For example, referees could be beneficially biased toward players who speak the same language as themselves or perhaps toward players from countries with similar cultures as themselves. For example, a referee from Germany may favor Swiss players or teams, or a referee from Norway perhaps might favor Swedish players or teams given the similarities in language and culture. Alternatively, commonality in language in referee-player pairings plausibly could be the driving force behind the bias that referees exhibit. Perhaps players without a language barrier can more effectively dispute calls or persuade the referee to be more lenient. For the majority of referees the data contain the native language and second languages spoken by each referee.

With a large enough dataset, and by using a similar methodology as above, one could potentially create an index of how much bias referees from any given country have toward players from any other specific country. In addition, one could then start to determine what characteristics are driving the bias between the two countries involved. Owing to limited statistical power this is not feasible with the currently available data. However, Table 7 provides some information regarding which characteristics affect the amount of bias between two countries.

TABLE 7
The Effect of Linguistic and Cultural Similarities of Referees and Players or Teams

	Foul Difference	
Variable	**Player Level**	**Team Level**
Same language	0.037	0.534
	[0.047]	[0.626]
Same native language	0.105	−0.138
	[0.065]	[0.659]
Cultural distance	−0.054*	0.26
	[0.031]	[0.975]
Geographic distance	−0.018	0.144
	[0.015]	[0.285]
Border	−0.047	−0.432
	[0.033]	[0.461]

Notes: Each cell reports the coefficient on the indicated variable when estimating the full model version of Equation (1) and replacing the same country binary variable with the indicated variable. For the team level, Equation (1) is used, but with observations at the team-match level instead of the player-match level. Standard errors are clustered at the match level and are reported in brackets.

***Significant at the 1% level; **significant at the 5% level; *significant at the 10% level.

Column 1 of Table 7 uses the full model specification of Equation (1) and replaces the same country variable with one of the five language or cultural variables listed in Table 7. The first two rows show that there is a positive, but not statistically significant, increase in beneficial foul calls for players who speak the same language or the same native language as the referee. Note that the most commonly spoken national language in the players' home country determines the players' spoken language. Row 3 uses the standardized CAGE index in which higher scores represent more cultural differences between the two countries. Row 3 shows that there is a small and marginally significant decrease in beneficial foul calls for every standard deviation increase in the CAGE index. This implies that there may be some effect of differing cultures between referees and players on foul calls; however, when using two alternative measures of cultural distance between referees and players no such relationship exists. Row 4 uses geographical distance measured in thousands of kilometers between the countries' largest cities to calculate cultural distance measure between countries. Row 5 uses whether or not the two countries being compared border each other as a cultural distance measure.

It does not appear that having a referee squad that speaks the same language or that has a similar cultural background significantly affects beneficial foul calls.

Column 2 of Table 7 reports a similar analysis to column 1 except at the team level. Column 2 shows that a team playing a match with a referee squad that speaks the same language or has a similar cultural background as the club's country of origin has no effect on the number of team beneficial foul calls.

V. CONCLUSION

Professional UEFA Champions League football matches provide a competitive, scrutinized, and high-stakes situation to analyze the existence and amount of own-nationality bias in highly trained referees. Taking advantage of the pairings of referees and players from the same country as set up by UEFA referee assignments allows us to identify an own-nationality bias exhibited by referees. The own-nationality bias found among professional UCL referees increases the number of beneficial foul calls for players from the referee's home country by 8–10%. This bias is concentrated on national team members especially while playing at home or in later stages of the tournament. When vertically differentiating referees, own-nationality bias is as strong or stronger for higher qualified referees. We also find that language and cultural similarities have little impact on referees' foul calls. Although direct application of these results to other markets may be tenuous at present, further research may continue to give insights and a better understanding of own-nationality bias in other contexts involving subjective assessments whether on the football pitch, in the classroom, or in a job interview.

REFERENCES

Buraimo, B., D. Forrest, and R. Simmons. "The 12th Man? Refereeing Bias in English and German Soccer." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173(2), 2010, 431–49.

Dohmen, T. J. "The Influence of Social Forces: Evidence from the Behavior of Football Referees." *Economic Inquiry*, 46(3), 2008, 411–24.

Garicano, L., I. Palacios-Huerta, and C. Prendergast. "Favoritism under Social Pressure." *Review of Economics and Statistics*, 87(2), 2005, 208–16.

Kleven, H. J., C. Landais, and E. Saez. "Taxation and International Migration of Superstars: Evidence from the European Football Market." *American Economic Review*, 103(5), 2013, 1892–924.

Parsons, C. A., J. Sulaeman, M. C. Yates, and D. S. Hamermesh. "Strike Three: Discrimination, Incentives, and Evaluation." *American Economic Review*, 101(4), 2011, 1410–35.

Pope, D. G., and J. C. Pope. "The Impact of College Sports Success on the Quantity and Quality of Student Applications." *Southern Economic Journal*, 75(3), 2009, 750–80.

Price, J., and J. Wolfers. "Racial Discrimination among NBA Referees." *Quarterly Journal of Economics*, 125(4), 2010, 1859–87.

Scoppa, V. "Are Subjective Evaluations Biased by Social Factors or Connections? An Econometric Analysis of Soccer Referee Decisions." *Empirical Economics*, 35(1), 2008, 123–40.

Sutter, M., and M. G. Kocher. "Favoritism of Agents—The Case of Referees' Home Bias." *Journal of Economic Psychology*, 25(4), 2004, 461–69.

Zitzewitz, E. "Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making." *Journal of Economics & Management Strategy*, 15(1), 2006, 67–99.