Referee Bias in South American Football

J. Dean Craig and Anna B. Faria†

† J. Dean Craig is Associate Professor of Economics at the University of Colorado Colorado Springs. Anna B. Faria is Assistant Professor of Economics at the University of Colorado Colorado Springs. Corresponding author: Anna B. Faria, email: afaria@uccs.edu. We thank Philip Heap and Edward Hoang for their helpful comments. The usual disclaimer applies.

Referee Bias in South American Football

Abstract

Researchers have found evidence that football referees have biases against away teams, awarding them more cards and fewer goals than expected. We investigate the extent of home bias in South American football and an alternative source of refereeing bias by exploring a unique feature of elimination football competitions in South America: the language differences between Brazilian teams and Spanish-speaking referees. In South American elimination cups, Libertadores and Sudamericana, a Spanish native speaker always referees matches involving Brazilian teams, which generates complaints from the Portuguese-speaking teams and their supporters. Our results, however, indicate there is no bias against Brazilian clubs and those complaints are unfounded, but we also find that home bias is prevalent and large in South American football. Finally, we conduct early tests on the effectiveness of VAR (Video Assistant Referee) technology in mitigating home bias and find no changes compared to non-VAR matches.

Keywords: soccer, home advantage, referee bias, language

1. Introduction

Nearly every Monday morning, football fans in all corners of the world will debate the results of the weekend's round.[2] While some of the contentious conversations revolve around player quality and manager substitution choices, it is also very common to hear from the round losers that the referee had a strong hand (or whistle) in contributing to their team's demise on the field. "The ref was biased," the loser will say, which the winner will call non-sense. The next Monday morning, the previous round's loser may be praising the same referee for their work that weekend. We would expect a lack of consistent opinions about refereeing from football fans, as they are biased in favor of their club, but do we really observe consistently biased refereeing?

There exists an extensive literature on the issue of refereeing quality and bias in football. There is a consensus that referees care about material and non-material payoffs from officiating. This means they consider not only their monetary compensation but also how their decisions affect their in-game (short run) and overall (long run) reputation. (Dohmen & Sauermann, 2016, pp. 679-80) This may translate into home bias, status bias, and into what Plessner and Betsch (2001) call sequential effects in refereeing, such as the increased likelihood of awarding a penalty kick or red card to a team after having done the same to its opponents and the decreased likelihood of granting a second penalty kick or red card in a given game.

Overall, there appears to be a consistent bias from referees in favor of home teams. This is true in other sports as well and the reasons appear to be similar across modalities: social pressure and other factors affect referees. Despite the material incentives to maintain neutrality, referees tend to favor the home team, leading to outcomes that deviate from unbiased predictions based on skill differentials and other game defining characteristics. Generally, the larger the crowd, the greater the bias for the home team. Dohmen and Sauermann (2016) provide an extensive review of the literature on referee bias, most of which suggests home team bias is present to varying degrees in football and other sports.

In this paper, we test the extent of home bias in South American football and explore an additional potential source of refereeing bias: a language barrier. In South American continental competitions, Brazilian teams, whose players speak Portuguese, face clubs from Spanish South America. Those matches are always refereed by Spanish-speaking referees from a third nationality.[3] This feature of South American continental

[2] We will use the terms soccer and football interchangeably throughout the paper.
[3] It is possible that Brazilian teams have Spanish-speaking South American players or Brazilian players who speak Spanish in their rosters, but they make up a small share of the players. For example, Spanish-speaking players represented between 3.7% and 12.5% of the rosters of Libertadores-playing Brazilian

competitions creates a unique “natural” experiment to test whether a language barrier can negatively affect a team’s expected match performance.

Another feature of our analysis is that the referees’ nationalities differ from the teams’ and, therefore, we do not expect players to adapt their playing based on referee assignment. While it is possible that players and coaches have some knowledge of any given referee’s style, most teams encounter a given referee only on occasion over the years in our sample. This is not the case in national or local competitions, in which players and referees are much more familiar with one another through repeated interaction, thereby possibly influencing the teams’ playing style and aggressiveness level on the pitch. (Hlasny & Sascha, 2015)

The language bias in refereeing has recently been raised as a potential ethical issue in sports. Mike McNamee laid out the ethical implications of an officiating language bias in a “Sports, Ethics and Philosophy” editorial in 2013:

> “unless sports regulatory institutions can find officials who can speak the mother tongue of both teams, some bias is likely to occur when referees are, as part of their officiating duties, required to communicate with players (and coaches) during the game. Can it really be the case that one contestant or team is not privileged by the comprehension of the officials’ communication? And then, of course, there will always be the questions and answers of the opposing players or captains or coaches.” (McNamee, 2013)

It is not enough that the referee can speak or comprehend the language of both teams unless the referee is a native speaker of both. The closer the languages are the easier comprehension is, but in sports like football which have continuous interaction between players, coaches and referees, the lack of fluency between parties cannot be discarded as a source of bias.

Similarly, nationality-based differences in refereeing have been explored as a potential source of bias. In 2010, Peter Dawson and Stephen Dobson found that referees of different nationalities punished home and away teams at different rates in UEFA continental competitions, but they do not identify the exact mechanism that leads to this variability. Dawson and Dobson found, for example, that Portuguese referees punish away teams at a higher rate than other nationalities, whereas Greeks punish teams at

---

teams in 2016. Their average roster size was 46 and the average number of Spanish-speaking players 3.8. (WorldFootball.Net, 2020) The South American cups we analyze do not impose limits on the use of foreign players in rosters, but the Brazilian Championship, the Brasileirão, imposes a match roster limit of five foreign players. (CBF, 2021) This limits the Brazilian teams’ incentives to hire foreign players in large numbers.

higher rates overall. They suggest that “a referee is likely to be influenced by his (national) identity and the nationality of the team.” (Dawson & Dobson, 2010, p. 189)

Our paper provides an improved strategy to further assess the mechanism Dawson and Dobson (2010) outline in their work, as we isolate the dichotomous language difference as an observable characteristic. This is the main contribution of our paper. While our main purpose is to explore this exogeneity, we also add to the literature by revisiting the home advantage question, by providing early estimates of the effects of VAR (Video Assistant Referee) technology on referee decision-making. With VAR, goal and card calls are reviewed by one or more remote assistant referees with access to video replays and offside line detection software. VAR may recommend invalidating a goal or switching card awards and therefore affect match results.

To the best of our knowledge, there are no other published studies that empirically assess language differences as a source of referee bias. At the time of this writing, we were able to find a single working paper on the topic, exploring language variations in Swiss football. Richard Faltings, Alex Krumer, Michael Lechner (2019) find that referees favor teams from their same linguistic region when playing against “outsiders.”

Language differences can influence referee decision-making in two main ways. First, referees, players and coaches are more likely to misunderstand one another when they speak different languages than when they share one. These misunderstandings may lead referee to dismiss players’ or coaches’ requests more often than they would or should were they to share a common language. Second, a language difference can lower the trust between players and referees. Players and coaches already tend to have low or negative perception of referees’ abilities. (Balch & Scott, 2007) A language barrier can only work to increase on-field tension between them. While referees are unlikely to report that stress or disturbance can cause them to affect their decision-making on the pitch, the fact that home bias exists suggests a language bias is also plausible. (Di Corrado, Pellarin, & Agostini, 2011; Johansen & Haugen, 2013)

We use data from the CONMEBOL Libertadores and Sudamericana cups between 2016 and 2019. We use match-level data to estimate differences in Brazilian and non-Brazilian match outcomes. Our main match outcome variables are CardRate (total number of cards over fouls) and NetGoals (difference between goals scored for and against). We follow Boyko, Boyko and Boyko (2007) in their choice of goal differential as a main way to assess match outcome differences.

Our main results come from two fixed-effects models. In the first model, we ask if there is bias for (or against) a Brazilian (B) team playing a non-Brazilian (N) team relative to two non-Brazilian (N vs. N) teams playing a similar match. We use this model to ascertain the possibility of a bias for (or against) Brazilian teams. In the second model, we ask if there is a bias for (or against) a non-Brazilian (N) team playing a Brazilian (B) team relative to two Brazilian (B vs. B) teams playing a similar match. We use this second model to check for an alternative source of bias for (or against) non-Brazilian teams.

Our main results indicate referees do not appear to have an anti-Brazilian bias. While this may appear to the reader as a null result, we suggest instead that the absence of bias is relevant in that it stymies common complaints of said bias. Moreover, we find economically significant home bias that cannot be explained away with controls, but may be referee-contingent. Initial results suggest the effect of VAR on our measure of goal differential and card rate is small and that VAR matches do not appear to have smaller home bias. Since there are not many years of data available on VAR and the implementation of VAR technology is not random between matches in our sample, we hesitate to derive strong conclusions about its effectiveness in correcting referee error and minimizing bias.

2. Biases in refereeing

Most studies regarding referee bias in football address the issue of home bias or advantage. That teams have better outcomes at home than away has been shown to be prevalent if not ubiquitous in soccer and other sports. The reason for this home advantage, however, is disputed. It is possible that teams simply play better at home than away. The alternative reason is that referees exhibit an anti-away or pro-home bias in their decision. There exists an extensive literature on referee home bias in football, mostly dealing with European competitions. Our purpose here is not to provide a comprehensive review of this body of work, but instead to highlight the proposed mechanisms that may partially explain this bias and how this bias presents itself in football and other sports.

There exists some evidence that players perform better at home than away (home advantage), in part due to crowd support. Research by Damien Poulter (2009) and Carmichael and Thomas (2005) shows home performance advantages for teams and players. Michela Ponzo and Vicenzo Scoppa (2018) have shown that crowd support appears to affect player performance in Italian Serie A same-stadium derbies. By looking only at same-stadium derbies, i.e., matches between teams that share a home stadium, they can control for travel fatigue and stadium familiarity (or lack thereof), potential confounding factors. Ponzo and Scoppa find that playing at “home”, i.e., being assigned home status for a particular match and therefore enjoying a larger share of supporters, increases a team’s probability of winning by 15 percentage points in same-stadium derbies. (Ponzo & Scoppa, 2018, p. 570)

None of these studies can fundamentally establish if performance differences and not referee decision-making are responsible for better at-home outcomes. If teams play better at home, for whatever reason, why do they not choose the same playing style when away? If performance at home cannot fully explain home advantage, then that leaves room for potential referee bias as an explanation. Referees are people too, after all, and are motivated by material and non-material payoffs.

Assuming away bribery and corruption, material payoffs do not change based on a referee's on-pitch decision, i.e., there is no additional compensation for high-scoring matches, or more or fewer cards awarded, although there is evidence that professionalization of referees and the introduction of annual salaried contracts reduces bias. (Rickman & Witt, 2008) Non-material payoffs change from match to match and this variation can shed light into the mechanisms that explain referee bias. The mechanisms we discuss are crowd pressure, match stakes and player or team reputation, with referee- and team-specific characteristic also playing a role. Most of the literature has focused on goal differentials, yellow and red card awards, stoppage or injury time and penalty-kick awards.

Despite the athletes' and fan's negative perceptions of referees, their personalities do not differ significantly from the population at large. (Balch & Scott, 2007) Referees are not wired to explicitly harm a team or player, nor do they portray or see themselves as such. (Johansen & Haugen, 2013) Still, evidence from football and other sports has demonstrated that referees make biased decisions that can affect match outcomes, suggesting most instances of bias are implicit. Bias is, therefore, created subconsciously but determined based on external factors. Crowd pressure in favor of home teams has been hypothesized and empirically checked as the main generator of referee bias.

Most of the literature on crowd pressure points to positive link between attendance levels or crowd density and referee bias, as most recently summarized in Dohmen and Sauermann (2016). The more home supporters are in attendance or the larger the occupation of the stadium, the more likely referees are to favor the home team, as referees are subjected to strong social influence from the one-sided supporters. (Boyko, Boyko, & Boyko, 2007; Buraimo, Forrest, & Simmons, 2010; Buraimo, Simmons, & Maciaszczyk, 2012) However, whenever an away team plays close to home, referee bias is less prevalent as more away supporters attend the match. (Dohmen, 2008; Garicano, Palacios-Huerta, & Prendergast, 2005)

Even in same-stadium derbies, as Ponzo and Scoppa's (2018) work has shown, referees will bias their decisions in favor of the home team, contingent on crowd size. This is particularly relevant for tourneys like the Libertadores and Sudamericana cups, since each additional stage moves a team directly closer to victory, whereas the stakes in league matches vary far less, even at the close of the tourney.[4] While teams and players know the prize and rewards structure they face at the season's start, as Garicano et al. (2005) put it," games at the end of the season may have different importance than those at the beginning, both because the end of the season is more imminent, and because teams have a better idea of their likely finishing position."

In parallel, as the season progresses, supporters tend to become more vocal during matches. Not only does attendance tend to change quantitatively, but also qualitatively. As a result, when the stakes are high, referee bias may be more prominent. This is in line with the results of Garicano et al. (2005) who find referee bias in favor of home teams in close Spanish La Liga matches increases as the season progresses. Dawson and Dobson (2010) find that referee award more cards to both teams, but more so to away players in elimination than in group rounds in the UEFA Champions League and UEFA Cup. Once they control for competition reputation, they find no effect on home penalizations, but the effect on away cards persists. This is similar to the results of Dawson, Dobson, Goddard and Wilson (2007) for the English Premier League.

Referees may also implicitly adjust their call precision as the stakes change within a given match. Mario Cesar de Oliveira, Rogerio Orbetelli and Turibio de Barros Neto (2011) sampled 321 foul calls from the São Paulo State Football Federation in Brazil and concluded referees were more precise in their correct calls in the last third of a given match half than in the first two thirds. Similar results come from the body of work on biased stoppage time. In close games, referees award longer stoppage times on average than in games with a "clear winner" at its close. (Dohmen & Sauermann, 2016, pp. 681-83; Riedl, Strauss, Heuer, & Rubner, 2015; Lago-Peñas & Gómez-López, 2016)

Referee bias may be dependent on player or team reputation. High status players or teams may benefit from referee's decisions as has been documented in the National Basketball Association and Major League Baseball (Kim & King, 2014; Caudill, Mixon Jr, & Wallace, 2014) A referee may adjust their expectations about a player's or team's performance based on their status. This refers to either the expectation of high-quality play or mode of play, e.g., if a player or team is known to be aggressive or confrontational.

On the one hand, biased referees will award or allow for scoring opportunities for high status players or teams. Those are the findings of Kim and King (2014) and Caudill, Mixon Jr and Wallace (2014).[5] In the football context, this bias may appear in referee's

[4] The Libertadores Cup has a group stage prior to its elimination stages.
[5] Christian Deutscher (2015) found no evidence of referee bias in favor of high-status NBA players. Ryan M. Rodenberg and Choong Hoon Lim (2009) found that a single referee made biased calls against the Dallas Mavericks in the NBA playoffs, hindering their performance, but otherwise no evidence of referee bias affecting team performance in seven season of games of the NBA. Rodenberg (2011) found that

awarding more fouls in the attacking third or penalties if a team’s player is known to be a good dribbler or skilled striker. Attacking fouls and penalties contribute to more goal scoring. If Neymar attempts to advance past a defender and upon contact falls, the referee may be more likely to attribute his collapse to illegal fouling than simply tripping or being disarmed.

On the other hand, biased referees may change their foul calling behavior based on the player’s or team’s reputation for aggressiveness or unsportsmanlike conduct. If Neymar has a reputation for being not just a skilled footballer, but also a diver (faker), referees may refrain from calling fouls against him that would have been awarded to another player. The evidence for status bias in football is still scarce. In a controlled experiment, football referees issued a team more cards when they were told ahead of time that that team had a reputation for aggressive behavior than they were not told that information. (Jones, Paull, & Erskine, 2010)

The best evidence we have so far about the benefits of high status in referee calling comes from Constantinou et al (2014). The authors found that being title-contenders in the English Premier League awarded Manchester United and Manchester City a greater degree of favoritism than their model would predict. Constantinou et al (2014) conclude that “it is possible that the combination of home advantage and being a title-favourite team (which Manchester United have been since the Premier League inception) in a close title race is what is more predictive of positive referee bias for penalty kicks awarded.” Similar results come from Erikstad and Johansen (2020), who found that Norwegian high-performing teams were more likely to have penalties called for them than low-performing ones.

Referee biases are context-specific and differ in their sources, as mentioned above. Moreover, the extent of the biases changes from referee to referee. Not only do baselines vary between referees, i.e., referees are more or less strict, but also some referees may be more susceptible to crowd pressure, status effects, stakes, etc., as summarized in Dohmen and Sauermann (2016). Factors affecting referee-dependent match outcomes may be referee personality, age, experience, or nationality.[6] The latter reflects differences in professional preparation and training as well as institutional idiosyncrasies that develop and are reinforced by national or local football federations and other actors.

perceived referee bias against Miami Heat was absent despite team claims, suggesting confirmation bias might affect perceptions of referee influence on game outcomes.

[6] Boyko, Boyko and Boyko (2007) have found that home advantage decreases with more experienced refereeing.

Match outcomes differ based on referee nationalities, as we show in Section 3. Curiously, even referee height is associated with differences in decision-making. McCarrick et al. (2020) found that shorter referees award more yellow cards on average than taller ones, despite no significant differences in fouling. Shorter referees also award more red cards in the English lower leagues, but fewer in the higher ones relative to taller referees.

Just as referees differ, so do teams. Roster quality, playing style and defending stance are some of the team characteristics that may affect match outcomes, regardless of referee bias. Some teams may play a possession-style or a counter-attacking strategy. Some managers allow players behave more aggressively in the field, while others instruct their players to refrain from getting stuck in. For example, in our sample, average home and away team possession in a season varied from 32% to 73% and 24% to 65%, respectively. The number of fouls in a season also varied widely, with the least aggressive team fouling on average four times in a game and the most aggressive on fouling close to 25 times. These factors, as well as other characteristics of play, may lead to an advantage (or disadvantage) which is independent of referee bias.

Another source of bias is the sequential effects of referee calls. Theoretically, each referee decision should be independent. However, the probability of a caution often increases or decreases depending on the sequence of the events during a match. This means that each referee decision is not independent, and calls are biased during a given match. There is evidence to suggest referees are less likely to award a penalty or issue a card to a team that has had a penalty or card awarded previously in the same match. (Plessner & Betsch, 2001; Buraimo, Simmons, & Maciaszczyk, 2012)

While the literature on the effects of language differences is small (if non-existent), there are several studies dealing with difference refereeing based on nationality. In cricket, Sacheti, Gregory-Smith and Paton (2015) find that a significant home advantage bias in cricket officiating existed when umpires were of the same nationality as the home team, but that the bias disappeared after the introduction of neutral umpires. In football, Dawson and Dobson (2010)'s findings suggest the rate at which home and away teams receive cards in UEFA continental competitions vary depending on referee nationality. They suggest that in an international setting, team and referee identity can influence the referee's decision-making under uncertainty. Unfortunately, Dawson and Dobson (2010) fail to identify the exact mechanism through which differences in nationality creates bias. By using language, an explicitly observable characteristic of teams and referees, we can delineate a plausible cause-and-effect mechanism to explain biased officiating.

We wish to highlight two ways in which language differences may lead to bias. First, differences in languages reflect social distance, which in turn determines how much individuals care about others. (Buchan, Johnson, & Croson, 2006) Referees may implicitly bias their decision-making against players and clubs perceived to be further socially separated from them and favor those who are closest, which is in line with the literature on the economics of identity coming out of Akerlof (1997). Evidence of this type of bias comes from the NHL, in which Canadian Francophone referees call penalties on English Canadian players at faster rates than do English-speaking Canadian referees. (Mongeon & Longley, 2015)

Second, language itself fundamentally affects the ways in which individuals process information and judge circumstances. Returning to the evidence from the NHL, Mongeon and Longley (2015) suggest that, on top of their language difference, French Canadian players, "are distinctly different from the Anglophone majority … along social, political, and cultural lines." As Lera Boroditsky (2012) put it broadly "speakers of different languages may learn to attend to and encode different aspects of the world." If Brazilian players behave differently on the pitch than their Spanish-speaking counterparts, then Spanish-speaking referees may on the margin interpret Brazilian behavior negatively, as it deviates from their language- and culturally- shaped expectations. We would expect this bias to be reduced if the referee has international experience.

Ultimately, as with most biases, it is the referee's perceptions of the players and clubs that change the costs of calling a foul or disallowing a legal play or awarding a card. Home bias is undoubtedly due to this implicit but undue reponse. We would expect that same, should a language bias exist. With that, we do not forget to point out that Spanish and Portuguese are very similar languages in their origin, grammar and vocabulary. We caveat, therefore, that the absence of a statistically distinguishable language bias may be due to their linguistic closeness.

Given that biases in refereeing appear to exist, football-controlling entities, such as FIFA and CONMEBOL, have made changes to the rules of the game and referee selection as well as implemented technologies designed to mitigate the effects of biases. One such technology is the recently implemented VAR or Video Assistant Referee. Since VAR is a recent implementation of the main international soccer organizations, there are no systematic studies of how the technology has changed the way the game is played and refereed. VAR trials began in Europe in the mid-2010s and implementation across leagues and cups has been piece-meal. The 2018 FIFA World Cup was the first official competition featuring VAR technology for every match. In South American continental competitions, VAR was implemented slowly. In 2017, two Libertadores games were chosen as test runs of the new technology. In 2018 and 2019, VAR was used in playoff games (quarter finals and beyond) in both Libertadores and Sudamericana cups. (Cruz, 2019) In our sample, VAR was used in 2.2% of 2017 games, in 10% of 2018 and in 22% of 2019 games. There were no VAR games in 2016 in our sample.

Given VAR is a developing technology, since we are interested in how it may affect game outcomes, we must look at theory and at how other sports changed with the introduction of VAR-like technologies. The theory, and justification, for systems like VAR is that they raise the costs of rule violation and therefore deter "bad" behavior from players, while decreasing referee mistakes. They also increase monitoring ability. So VAR-like systems can increase detection rates due to more monitoring, and also rule violation. (Dawson, Dobson, Goddard, & Wilson, 2007, p. 233) In other words, VAR can both increase the costs of deliberate deception and correct unintentional mistakes by referees.

The evidence regarding monitoring and deterrence suggests players adjust their behavior marginally in the face of incentive changes. In his study of the National Hockey League, Allen (2002) found that after the introduction of a second penalty-calling referee in some games in 1999, violent offenses, such as high-sticking, were more likely to occur or be detected in two-referee games than in single referee games, while the incidence of non-violent penalties was unchanged, which suggests two referees increased monitoring more than they deterred illegal violent behavior, but that monitoring and deterrence may have been equally relevant in determining non-violent behavior.

Robert Witt (2005) found that English Premier League players adjusted their fouling behavior between the 1997–8 and 1998–9 seasons in response to a rule change that sanctioned red card punishment for tackles from behind, considered to be violent fouling. There was no significant increase in red cards after the introduction of the new rule, but the number of fouls that resulted in no card (least violent) and yellow cards (less violent) rose as a result. In this case, the rule change deterred players from committing the violent foul but caused them to change behavior and switch to lesser-violent tackling strategies at the margin.

Furthermore, VAR-like systems are expected to mitigate the effects of idiosyncratic (or biased) referee behavior. While the referee is the final decision-maker, VAR assistants can alert the referee to a variety of unobserved (or mistakenly so) events on the pitch. Those include violent fouling or aggressive behavior, missed fouling in defensive box, which would lead to a penalty-kick violation, or illegally scored goals, mostly due to missed offside calling on the part of an assistant referee, to fouling or hand balling by the attacking team. The ability (or willingness) of the referee to influence the result of the match should therefore be reduced with the advent of VAR. (Dohmen & Sauermann, Referee Bias, 2016, p. 692)

Parsons et al. (2011) analyzed MLB games with and without video technology and found racial discrimination in strike-calling decreased in games with video technology and high attendance. Their results indicate that referee bias is contingent on external monitoring. Moreover, baseball players adjusted their pitching to account for known biased umpire calling. According to Parsons et al. (2011)," [pitchers] who match the umpire's race [or] ethnicity attempt to 'paint the corners,' throwing pitches allowing umpires the most discretion. This tendency is much stronger in low-scrutiny situations, when umpires face a lower cost of indulging their preferences." (Parsons, Sulaeman, Yates, & Hamermesh, 2011, p. 1411) Not only do umpires allow themselves to bias their strike-calling, but knowledgeable pitchers adjust their playing style accordingly.

In cricket, the introduction of the Decision Review System (DRS) in 2009 has altered umpire decision-making incentives. With DRS, cricket teams can challenge an on-field umpire's decision, which will be reassessed by a third umpire off-field with the assistance of video technology. Ram Shivakumar's (2018) analysis of DRS implementation concludes that "with the advent of the DRS, on-field umpires appear to be less willing to give the benefit of the doubt to the batsman, a tradition, though not a rule, in cricket for more than a century" (p. 317) Moreover, Shivakumar finds that the third umpire's decisions appear to be unbiased against away teams, suggesting DRS works well to monitor on-field and off-field umpire behavior.

Since there are no VAR studies available yet, we must look at changes in other aspects of professional football to assess the extent to which monitoring and deterring occurs. One of those changes is the live broadcasting of matches, which increases scrutiny over referee decision-making and player behavior. If referees and players care about reputation, live match broadcasting should lead to changes in behavior on the pitch. Oddly, Dawson et al (2007) found that live broadcasting of English Premier League matches did not alter players' or referees' behavior on the pitch. In basketball, however, the evidence suggests that national live broadcasting changes game outcomes by lowering score differentials in the NBA and WNBA compared to non-televised games. (Wang, Hilsman, & Caudill, 2014)

The extent to which and the channels through which referees bias their decisions on the pitch matter for the evolution of the sport of football. Each league's reputation matters for viewership, as seen in the MLB strike of 1994, which caused viewership to drop and not recover to pre-strike levels into well into the mid-2000s. (ESPN, 2004) Similarly, NHL Stanley Cup TV ratings were much lower following the 2004/05 lockout than in preceding years. (Bleacher Report, 2013) Finally, in Italian football, the *Calciopoli* cheating scandal resulted in revenue losses and attendance decline. (Buraimo, Migali, & Simmons, 2016) The perception of fairness influences a league's reputation and its viability. In the next section, we develop our econometric model.

3. Econometric model

Our econometric goal is to assess if there are any differences in referee-contingent outcomes between Brazilian and non-Brazilian clubs playing in elimination-style, high stakes South American competitions. Our main data come from the website Footystats.org (2020), which has compiled match-level data from 2016 through 2019. We transformed each original match observation into two separate observations, one for each club involved. This allows us to test for home advantage effects. Additional control variables come from sources outlined below.

We use a fixed effects model to estimate if referees bias their decisions against (or for) non-Spanish speaking clubs. We want to find out if referees treat a Brazilian team playing non-Brazilian team (B vs. N) differently than they would a non-Brazilian team playing another non-Brazilian team (N vs. N) and if referees treat a non-Brazilian team playing a Brazilian team (N vs. B) differently than they would a Brazilian team playing another Brazilian team (B vs. B). We split our main sample into two subsamples to conduct these tests. In the first specification, the Brazilian team is the treatment, while in the second, the non-Brazilian team is the treatment. These estimations allow us to separate possible bias effects.

We therefore estimate two separate equations, for two separate samples, one containing observations where a Brazilian team is playing a non-Brazilian team (B vs. N) and a non-Brazilian team is playing another non-Brazilian (N vs. N), and the other containing all observations where a non-Brazilian team is playing a Brazilian team (N vs. B) and Brazilian teams are facing each other (B vs. B). The first sample contains 1,221 observations and the second 399, since there are fewer B vs. B encounters in the data.

We recognize our results may be driven by unobserved referee- or team-specific characteristics that correlate independently with our dependent variables. For example, Paraguayan and Peruvian referees award more fouls and cards than other nationalities while Chilean and Brazilian less so, as shown in Table 1. We control for referee-specific characteristics and a standardized skill differences measure for each match to lessen team-specific differentials. We find our measure is more precise in isolating skill differentials to traditional betting odds variables, because it is uncorrelated with home playing (r=-0.00223), which is taken into account in betting odds (r=0.601). We further control for team violence in our CardRate estimations and for playing style in the goal differential (NetGoals) regressions.

TABLE 1 HERE

The estimated equations are as follows:

$$Outcome_i = \beta_0 + \beta_1 * Brazil_i + \beta_2 * Away_i + \gamma * \boldsymbol{X}_i + \alpha_r + \varepsilon_i$$

and

$$Outcome_i = \beta_0 + \beta_1 * Foreign_i + \beta_2 * Away_i + \gamma * \boldsymbol{X}_i + \alpha_r + \varepsilon_i,$$

where our dependent variables (*Outcome*) are *CardRate* and *NetGoals*, in which *CardRate* is total number of cards divided by the total number of fouls a team *i* commits in each match, and *NetGoals* is the goal differential between the team *i* and its opponent. The independent variables of interest are *Brazil*, *Foreign* and *Away*. *Brazil* takes a value of one if the team is Brazilian, zero otherwise. *Foreign* takes the opposite values. Similarly, Away takes a value of one if the team is playing away, zero otherwise. $\mathbf{X}_i$ is a vector of other covariates including measures of team quality, playing style, aggressiveness and stadium condition (attendance and presence of track), all of which we describe below. We include referee fixed effects to control for referee-specific characteristics (α_r). We provide a table of variable names, descriptions and sources in Appendix A.

We have chosen CardRate as our main measure of punishment instead of the traditional point system because the CardRate variable already accounts for a team's relative propensity for fouling. The awarding of cards during a match is at the referee's discretion. The reader may wonder if the decision to call a foul is also potentially biased against a particular team. A simple t-test suggests away teams are no more likely to foul than home teams, but they are more likely to receive cards as a result (see Table 2).There appears to be a difference in fouls and yellow cards received between Brazilian and non-Brazilian teams, with the latter being more likely to foul and be punished than the former (see t-test results in Table 3).

TABLE 2 HERE

TABLE 3 HERE

Given non-Brazilian teams (Ns) both foul more and are punished more, there is no clear indication of bias either in calling fouls or awarding cards. The same cannot be said of home bias in carding since home teams appear to receive fewer cards per foul than away teams while fouling at similar rates. In our CardRate model, we include average team possession in match and season (Possession and Mean Possession) and following season team controls: fouls (Mean Fouls), cards (Mean Cards), CardRate (Mean CardRate), and red cards (Mean Reds).

The differences between Brazilian and non-Brazilian teams in scoring and defending are significant, as shown in Table 3 above. Brazilians score more and get scored on less than non-Brazilian teams. These statistics suggest that without controlling for differences in team quality, our main results will capture these skills differentials as well as any potential referee bias. We therefore control for playing style and skill differential using team possession rate in each match (Possession), average team possession in season (Mean Possession), average goal differential in season (Mean NetGoals) and average goals scored in season (Mean Goals) in our NetGoals regressions.

To further control for differences in team quality, we use the Ranking CONMEBOL Libertadores. Each season participating teams are given a score based on their (discounted) performance in the competition in the last ten years, their winning record in the competition between 1960 and 2006 and the number of local or national championships won in the prior year. The detailed methodology of the Ranking is available online. (CONMEBOL, 2017) Since no such ranking is available for the CONMEBOL Sudamericana, we have taken the Ranking CONMEBOL Libertadores methodology and modified it to generate a scoring system for teams participating in the Sudamericana Cup each year.

The ranking data differ between the two competitions and over time. In 2016, the average number of points in the Ranking CONMEBOL Libertadores was 1,740 points, compared to 178 points in the Sudamericana ranking we created. In 2019, the average had increased to 1,992 points team in the Libertadores ranking compared to 317 points in the Sudamericana ranking. These differences complicate our analysis since we use data from both competitions. Therefore, we standardize the Ranking variables to make them comparable between competitions and seasons. We take the difference between the teams' standardized ranking for each match to create our variable of interest NetPoints, which assesses skills differentials for each match as distance from the standardized mean.

The advantage of using these ranking variables is that they are not confounded with home and away effects, since the score for each team does not vary between matches, but NetPoints does, since each match-up is unique. The ranking variables do not account for form during the season, which would be ideal. However, if we were to control for form before each match, then we would have had to assess each team's performance in their local or national competitions. This would require compiling information on the quality of the opponents they faced in those other events, which is not possible with the current quality of data available, especially for lower reputation local and national leagues. We find that even by forgoing an explicit measure of form, we can capture the relevant quality differences by using the ranking variables as above.

Our attendance data, Crowd, come from WorldFootball.net (2020). The CONMEBOL does not require local match officials to report attendance data for either competition, so the staff at WorldFootball.net has compiled these data from a sports news websites and local newspapers. We are grateful for the work that they have done and shared with us for this project but use the information with caution as their data are approximations. The data on matches played in Brazil are far more precise than from all other countries in our sample. We have inspected the data closely and for all matches that appeared to have far too-low attendance, we conducted a video check of match highlights and a news search. The visual and reporting evidence was in line with the low attendance numbers.

Several studies estimate that distance from pitch influences referees due to crowd pressure. (Dawson & Dobson, 2010; Buraimo, Forrest, & Simmons, 2010; Scoppa, 2008; Dohmen, 2008) We therefore include a dummy variable, Track, to assess if the supporters' distance from the field affects referee decision-making. Track equals one if the stadium has a track, zero otherwise, with data compiled from Worldstadiums.com (2020).

Tables 4 and 5 below present the descriptive statistics for our main variables by club nationality (Brazilian vs. non-Brazilian) and by home status (home vs. away). Brazilian clubs, compared to non-Brazilian ones, tend to have greater goal differentials in their favor (0.55 vs. -0.15), score more goals (1.44 vs. 1.12) and have higher standardized ranking points (0.21 vs. -0.06). Brazilian clubs tend to foul slightly less (14.22 vs. 14.99) and receive fewer cards (2.48 vs. 2.70).

TABLE 4 HERE

TABLE 5 HERE

In Table 5, we see that home teams receive fewer cards per foul and tend to possess the ball more often and win more. Taken together, these statistics begin to paint a picture of our econometric exercise below. There exist differences between our groups (Brazilian and non-Brazilian and home and away). As we show below, we cannot rule out referee home biases.

We drop Brazilian referees from our main regressions, as Brazilian referees only work in non-Brazilian (N vs. N) matches and we wish to estimate the behavior differences of non-Brazilian referees only. We recognize that future research into the differential behavior of Brazilian referees in local (B vs. B) and continental (N vs. N) competitions will be valuable as well but is beyond the scope of this work.

Our sample includes Mexican teams and referees, who participated in the 2016 Libertadores. No Mexican teams or referees participated in any CONMEBOL competition thereafter. Since the language barrier between Brazilian teams and Mexican teams and referees is identical to that of the rest of our sample, we do not drop Mexican team and referee observations from our main estimations. We find no significant differences in coefficients and statistical significance (not reported) when we exclude Mexican referees from the regressions.

4. Results

Our main results suggest there is no bias against Brazilian teams in B vs. N matchups relative to N vs. N games. Nor is there evidence of bias in favor of non-Brazilians in N vs. B matchups relative to B vs. B games. Our estimations of differences in card issuance suggest there is no difference in the rates at which Brazilian and non-Brazilian teams are punished, as we show in Tables 6 and 7.

TABLE 6 HERE

TABLE 7 HERE

Referees appear to be fair regardless of language barrier, though they punish away teams at a greater rate than home teams. If we interpret CardRate to be the probability of being awarded a card for any given foul committed, then away teams are on average 20% more likely to receive a card than a comparable team playing at home. Playing away is associated with between 3.1% and 3.6% greater carding probability, or one third of a standard deviation. A club's season CardRate, which captures aggressiveness in fouling independent of referee selection or home status, is associated with similar increases in CardRate. Our measures of crowd pressure have no statistical relationship with CardRate.

With regards to match outcomes, i.e. NetGoals, there again appear to be no differences in match outcomes between Brazilian and non-Brazilian clubs once we control for skills differentials. Brazilian teams appear to perform better than non-Brazilian teams when facing a non-Brazilian adversary after controlling for playing style by including Possession and Mean Possession (columns 1 and 2 of Table 8). This performance difference disappears once we control for skill differentials (columns 3 and 4 of Table 8). The addition of Mean NetGoals erases the statistical significance of the Brazil coefficient. This result suggests that, once quality differences are taken into account, there remains no variation in match outcome that one might attribute to referee bias.

TABLE 8 HERE

We derive similar conclusions from Table 9, in which we report our results from comparing the performance of non-Brazilian teams and Brazilian teams facing Brazilian adversaries. While non-Brazilians facing Brazilians appear to underperform, once skill differentials are controlled for, this performance difference disappears. Crowd density does not appear to matter when facing Brazilian adversaries (Table 9) but, curiously, has a negative relationship with NetGoals when teams face non-Brazilian adversaries, as shown in Table 8. This may be due to the lack of precise attendance data that we discussed in the previous section, though we cannot be sure. The presence of a track separating supporters from the pitch has no statistically significant relationship to NetGoals.

A key finding is that there exists a strong home advantage effect that does not appear to be related to differences in skill differentials or playing style. In all NetGoals specifications, we find a statistically significant negative relationship between playing away and winning (NetGoals). Playing away lowers NetGoals by between three-fourths and six-sevenths of a standard deviation, an economically significant relationship. The average winning team has a 1.78 goal differential. The Away coefficient suggests a reduction of 1.44 goals, reducing the average goal differential to 0.34. All else equal, away teams are more likely to tie than they would have been playing a comparable team at home. While referees have far less discretion in affecting match results compared to awarding cards, the possibility of a pro-home team bias cannot be ruled out.

TABLE 9 HERE

In unreported estimations, we assess the role of competition stakes, team status and betting odds and find no significant changes (economically or statistically) in our coefficients when adding status and odds controls. We find that when stakes are lower (group phase games of Libertadores), the absolute size of Away coefficients increases between 13.9% and 25% for CardRate regressions and 8.5% and 12.6% for NetGoals estimations. This may be due to greater club skill differentials or lower quality refereeing. The larger the stakes, the smaller the home bias. In elimination phase matches only, the Away coefficient is between 11.1% and 12.9% in CardRate estimations and between 8% and 9.8% smaller in NetGoals regressions. In sum, there are no differences in outcomes that would suggest a disadvantage against Brazilian teams, only against away teams. We find therefore that complaints of referee bias based on language barriers are unfounded.[7]

Given that home bias is prevalent, one hopes for increased fairness from VAR technologies. We have conducted some early tests on its effectiveness. Biased referees might award a red card to an away team player more often than to home team player. If

[7] See, for example, ESPN Brasil (2020), Yahoo! Sports (2018) and GloboEsporte (2014).

referee bias in NetGoals occurs due to the referee's decision to allow an illegal goal to stand, VAR should eventually help correct this bias. Our results (not reported) indicate VAR does not have a statistically significant effect on CardRate, but VAR is associated with lower NetGoals by between 0.33 (0.19 of a standard deviation, p=0.06) and 0.61 goals (0.34 of a standard deviation, p=0.01). More importantly, controlling for VAR does not modify the Brazil, Foreign or Away coefficients in any of the specifications. These early estimations indicate that, in our limited sample, VAR did not mitigate home bias.

We suspect that upon further investigation greater differences may emerge due to VAR's potential in disallowing goals scored by offside players, a violation often outside of referee control, or by illegal fouling and handballing. Perhaps a further innovation is needed in the use of VAR: transparency. Currently, VAR audio recordings between the VAR booth and the on-pitch referee are not always available to the clubs or the public. If VAR officials and referees know each other, which they often do, it is possible that there is implicit collusion in decision-making between the VAR booth and the pitch. Releasing these recordings after each match may not only shed some light on the VAR reviewing process, but also increase monitoring of the match officials' behavior on and off the field.

5. Conclusion

Football fans will always, especially in South America, look for explanations for each week's losses on the pitch. Referees are easy pickings. It appears no one likes them. Good refereeing goes unnoticed ("they did nothing more than their job"), while bad, but potentially honest mistakes (or no mistakes at all) are scrutinized and vilified by fans and football commentators. Still, as our results indicate, referees tend to be fair in their decision-making, with the well-reported exception of home bias. So far, we see no effects of VAR technology on mitigating this bias, but more research and data are needed. Ultimately, we do not expect to change the minds of fans, club managers and partisans, but we bring to light evidence that if referees make mistakes, they appear to do so randomly, at least in high stakes, high quality international competitions.

The question of language bias is of course not fully answered. Are there other means to assess language differences? Does the presence of a Spanish-speaking player in a Brazilian team's roster matter? Does the experience and certification of referees explain their fairness in professional continental competitions? Football is a global sport and as players move around the world to play the beautiful game, lovers of the sport expect fairness and equal treatment on the pitch. Continuing to answer this question is fundamental to understanding the success of this global enterprise from the perspective not only of clubs and their supporters, but also from the players', who are moving across the world more than ever.

6. References

Akerlof, G. (1997). Social Distance and Social Decisions. Econometrica.

Allen, W. D. (2002). Crime, Punishment, and Recidivism Lessons from the National Hockey League. Journal of Sports Economics.

Balch, M. J., & Scott, D. (2007). Contrary to Popular Belief, Refs are People Too! Personality and Perceptions of Officials. Journal of Sport Behavior.

Billaut, F., & Aughey, R. (2013). Update in the understanding of altitude-induced limitations to performance in team-sport athletes. British Journal of Sports Medicine, i22-i25.

Bleacher Report. (2013). Has the NHL Officially Won Back Its Fans After the Lockout? Retrieved September 15, 2020, from https://bleacherreport.com/articles/1624740-has-the-nhl-officially-won-back-its-fans-after-the-lockout

Boroditsky, L. (2012). How the Languages We Speak Shape the Way We Think. In M. Spivey, K. McRae, & M. Joanisse (Eds.), The Cambridge Handbook of Psycholinguistics. Cambridge University Press.

Boyko, R., Boyko, A., & Boyko, M. (2007). Referee bias contributes to home advantage in English Premiership football. Journal of Sports Sciences, 1185-1194.

Brutsaert, T., Spielvogel, H., Soria, R., Araoz, M., Caceres, E., Buzenet, G., . . . Vargas, E. (2000). Performance of Altitude Acclimatized and Non-Acclimatized Professional Football (Soccer) Players at 3,600 M. Journal of Exercise Physiology.

Buchan, N. R., Johnson, E. J., & Croson, R. T. (2006). Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. Journal of Economic Behavior & Organization.

Buraimo, B., Forrest, D., & Simmons, R. (2010). The 12th man?: refereeing bias in English and German soccer. Journal of the Royal Statistical Society A.

Buraimo, B., Migali, G., & Simmons, R. (2016). AnAnalysis of Consumer Response to Corruption: Italy's Calciopoli Scandal. Oxford Bulletin of Economics and Statistics.

Buraimo, B., Simmons, R., & Maciaszczyk, M. (2012). Favoritism and Referee Bias in European Soccer: Evidence from the Spanish League and the UEFA Champions League. Contemporary Economic Policy.

Burtscher, M. (2005). The athlete at high altitude: Performance diminution and high altitude illnesses. International SportMed Journal, 215-223.

Carmichael, F., & Thomas, D. (2005). Home-Field Effect and Team Performance. Journal of Sports Economics.

Caudill, S. B., Mixon Jr, F. G., & Wallace, S. (2014). Life on the Red Carpet: Star Players and Referee Bias in the National Basketball Association. International Journal of the Economics of Business.

CBF. (2021). RGC - Regulamento Geral das Competições.

Chapman, R. F. (2013). The individual response to training and competition at altitude. British Journal of Sports Medicine, i40-i44.

CONMEBOL. (2017). Ranking Conmebol de Copa Libertadores. Retrieved June 6, 2020, from http://www.conmebol.com/es/ranking-conmebol-libertadores2017

CONMEBOL. (2020). Libertadores Ranking. Retrieved from http://www.conmebol.com/es/ranking-conmebol-libertadores2020

Constantinou, A. C., Fenton, N. E., & Pollock, L. J. (2014). Bayesian networks for unbiased assessment of referee bias in Association Football. Psychology of Sport and Exercise.

Cruz, I. (2019, April 1). VAR en la Copa CONMEBOL Libertadores 2019. Retrieved May 11, 2020, from https://www.copalibertadores.com/es/noticias/var-en-la-copa-conmebol-libertadores-2019-cuando-se-usa-como/hn4x1eo93k6k1790e7qsghst0

Dawson, P., & Dobson, S. (2010). The influence of social pressure and nationality on individual decisions: Evidence from the behaviour of referees. Journal of Economic Psychology, 181-191.

Dawson, P., Dobson, S., Goddard, J., & Wilson, J. (2007). Are football referees really biased and inconsistent?: evidence on the incidence of disciplinary sanction in the English Premier League. Journal of the Royal Statistical Society, 231-250.

de Oliveira, M. C., Orbetelli, R., & de Barros Neto, T. (2011). Call Accuracy and Distance from the Play: A Study with Brazilian Soccer Referees. International Journal of Exercise Science.

Deutscher, C. (2015). No referee bias in the NBA: New evidence with leagues' assessment data. Journal of Sports Analytics.

Di Corrado, D., Pellarin, E., & Agostini, T. A. (2011). The phenomenon of social influence on the football pitch: Social pressure from the crowd on referees' decisions. Review of Psychology.

Dohmen, T. (2008). The Influence of Social Forces: Evidence from the Behavior of Football Referees. Economic Inquiry.

Dohmen, T., & Sauermann, J. (2016). Referee Bias. Journal of Economic Surveys, 679-695.

Downward, P., & Jones, M. (2007). Effects of crowd size on referee decisions: Analysis of the FA Cup. Journal of Sports Sciences.

Erikstad, M. K., & Johansen, B. T. (2020). Referee Bias in Professional Football: Favoritism Toward Successful Teams in Potential Penalty Situations. Frontiers in Sports and Active Living.

ESPN. (2004). 1994 strike was a low point for baseball. Retrieved from https://www.espn.com/espn/print?id=1856626&type=story

ESPN Brasil. (2020). Fonte: https://www.espn.com.br/video/clipe/_/id/6723221

Faltings, R., Krumer, A., & Lechner, M. (2019). Rot-Jaune-Verde. Language and Favoritism: Evidence from Swiss Soccer.

FiveThirtyEight. (2020). Global Club Soccer Rankings. Retrieved September 15, 2020, from https://projects.fivethirtyeight.com/global-club-soccer-rankings/

Football World Rankings. (2020). Football World Rankings. Retrieved from https://www.clubworldranking.com/

FootyStats. (2020). FootyStats. Retrieved from https://footystats.org/

Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2005). Favoritism Under Social Pressure. The Review of Economics and Statistics.

Garvican, L. A., Hammond, K., Varley, M. C., Gore, C. J., Billaut, F., & Aughey, R. J. (2014). Lower Running Performance and Exacerbated Fatigue in Soccer Played at 1600 m. International Journal of Sports Physiology and Performance.

GloboEsporte. (18 de January de 2014). 'Poliglotas' do Cruzeiro podem ajudar na conquista da Libertadores. Acesso em 1 de April de 2021, disponível em http://globoesporte.globo.com/futebol/times/cruzeiro/noticia/2014/01/poliglotas-do-cruzeiro-podem-ajudar-na-conquista-da-libertadores.html

Gore, C. J., Aughey, R. J., Bourdon, P. C., Garvican-Lewis, L. A., Soria, R., Claros, J. C., . . . Schmidt, W. F. (2013). Methods of the international study on soccer at altitude 3600 m (ISA3600). British Journal of Sports Medicine, i80-i85.

Goumas, C. (2014). Home advantage and referee bias in European football. European Journal of Sport Science.

Hlasny, V., & Sascha, K. (2015). Catch Me If You Can: Referee–Team Relationships and Disciplinary Cautions in Football. Journal of Sports Economics.

Hörzer, S., Fuchs, C., Gastinger, R., Sabo, A., Mehnen, L., Martinek, J., & Reichel, M. (2010). Simulation of spinning soccer ball trajectories influenced by altitude. Procedia Engineering.

Johansen, B. T., & Haugen, T. (2013). Anxiety level and decision-making among Norwegian top-class soccer referees. International Journal of Sport and Exercise Psychology.

Johnston, R. (2008). On referee bias crowd size and home advantage in the English soccer Premiership. Journal of Sport Science.

Jones, M. V., Paull, G. C., & Erskine, J. (2010). The impact of a team's aggressive reputation on the decisions of association football referees. Journal of Sports Sciences.

Kim, J. W., & King, B. G. (2014). Seeing Stars: Matthew Effects and Status Bias in Major League Baseball Umpiring. Management Science.

Lago-Peñas, C., & Gómez-López, M. (2016). The Influence of Referee Bias on Extra Time in Elite Soccer Matches. Perceptual and Motor Skills.

Levine, B. D., Stray-Gundersen, J., & Mehta, R. D. (2008). Effect of altitude on football performance. Scandinavian Journal of Medicine and Science in Sports, 76-84.

McCarrick, D., Brewer, G., Lyons, M., Pollet, T. V., & Neave, N. (2020). Referee Height Influences Decision Making in British Football Leagues. BMC Psychology.

McNamee, M. (2013). Sports officiating, linguistic bias and fair play. Sports, Ethics and Philosophy.

Mongeon, K., & Longley, N. (2015). Testing for Ethnicity Discrimination among NHL Referees: A Duration Model Approach. Eastern Economic Journal.

Nassis, G. (2013). Effect of Altitude on Football Performance: Analysis of the 2010 FIFA World Cup Data. Journal of Strength and Conditioning Research.

Parsons, C. A., Sulaeman, J., Yates, M. C., & Hamermesh, D. S. (2011). Strike Three: Discrimination, Incentives, and Evaluation. American Economic Review.

Plessner, H., & Betsch, T. (2001). Sequential Effects in Important Referee Decisions: The Case of Penalties in Soccer. Journal of Sport and Exercise Psychology.

Ponzo, M., & Scoppa, V. (2018). Does the Home Advantage Depend on Crowd Support? Evidence From Same-Stadium Derbies. Journal of Sports Economics.

Poulter, D. (2009). Home advantage and player nationality in international club football. Journal of Sports Sciences.

Rankings, F. W. (2020). All Time Top 250 Clubs. Retrieved September 2020, from https://www.clubworldranking.com

Reilly, B., & Witt, R. (2012). Red cards, referee home bias and social pressure: evidence from English Premiership Soccer. Applied Economics Letters.

Rickman, N., & Witt, R. (2008). Favouritism and Financial Incentives: A Natural Experiment. Economica.

Riedl, D., Strauss, B., Heuer, A., & Rubner, O. (2015). Finale furioso: referee-biased injury times and their effects on home advantage in football. Journal of Sports Sciences.

Rodenberg, R. (2011). Perception ≠ Reality: Analyzing Specific Allegations of NBA Referee Bias. Journal of Quantitative Analysis in Sports.

Rodenberg, R. M., & Lim, C. H. (2009). Payback Calls: A Starting Point for Measuring Basketball Referee Bias and Impact on Team Performance. European Sport Management Quarterly, 375-387.

Sacheti, A., Gregory-Smith, I., & Paton, D. (2015). Home bias in officiating: evidence from international cricket. Journal of the Royal Statistics Society Series A.

Scoppa, V. (2008). Are subjective evaluations biased by social factors or connections? An econometric analysis of soccer referee decisions. Empirical Economics.

Shivakumar, R. (2017). What Technology Says About Decision-Making: Evidence From Cricket's Decision Review System (DRS). Journal of Sports Economics.

Shivakumar, R. (2018). What Technology Says About Decision Making Evidence from Cricket's Decision Review System. Journal of Sports Economics.

Tovar, J. (2014). Gasping for air: soccer players' passing behavior at high altitude. Journal of Quantitative Analysis in Sports .

Wang, Y., Hilsman, W., & Caudill, S. B. (2014). Television Coverage and Outcome Uncertainty in Sports: Empirical Evidence from the NBA and WNBA. International Journal of Sport Science.

Witt, R. (2005). Do players react to sanction changes? Evidence from the English Premier League. Scottish Journal of Political Economy.

WorldFootball.Net. (2020). WorldFootball.Net. Retrieved from www.worldfootball.net

WorldStadiums.com. (2020). WorldStadiums.com. Retrieved from http://www.worldstadiums.com/

Yahoo! Sports. (19 de November de 2018). A arbitragem da Copa Libertadores prejudica os clubes brasileiros? Fonte: https://esportes.yahoo.com/noticias/arbitragem-da-copa-libertadores-prejudica-os-clubes-brasileiros-080129690.html

Tables

Table 1 – Match Outcome Differences by Referee Nationality

Nationality	Card Rate	Fouls	Yellow Cards	Red Cards	Match Goals
Argentina	0.19	13.53	2.41	0.16	2.16
Bolivia	0.20	14.64	2.60	0.13	2.47
Brazil	0.18	14.78	2.51	0.13	2.11
Chile	0.18	14.61	2.33	0.17	2.17
Colombia	0.22	14.47	2.85	0.20	2.61
Ecuador	0.16	14.97	2.14	0.15	2.46
Mexico	0.13	14.50	1.90	0.00	1.60
Paraguay	0.20	15.25	2.64	0.20	2.48
Peru	0.20	16.14	2.74	0.28	2.56
Uruguay	0.17	15.97	2.43	0.21	2.38
Venezuela	0.16	14.34	2.11	0.13	2.65
Average Total	0.18	14.84	2.42	0.16	2.33

Table 2 - Differences in Means (Home – Away)

Fouls	-0.112
	(-0.573)
Cards	0.635***
	(0.000)
Yellow Cards	-0.561***
	(0.000)
Red Cards	-0.074***
	(0.000)

p-values in parentheses: * p < 0.05, ** p < 0.01, *** p < 0.001

Table 3 - Differences in Means (Brazilian – Non-Brazilian)

Fouls	0.786**
	(-0.001)
Cards	0.228**
	(-0.006)
Yellow Cards	0.216**
	(-0.004)
Red Cards	0.012
	-0.622
NetGoals	-0.700***
	(0.000)
Goals For	-0.319***
	(0.000)
Goals Against	0.381***
	(0.000)

p-values in parentheses: * p < 0.05, ** p < 0.01, *** p < 0.001

Table 4 - Descriptive Statistics by Club Nationality

	Brazilian Clubs				Non-Brazilian Clubs			
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
CardRate	0.18	0.11	0.00	0.78	0.19	0.11	0.00	0.88
NetGoals	0.55	1.74	-5.00	7.00	-0.15	1.79	-8.00	8.00
Crowd	0.59	0.26	0.01	1.03	0.48	0.28	0.01	1.03
Track	0.23	0.42	0.00	1.00	0.32	0.46	0.00	1.00
Mean Fouls	14.22	2.40	9.50	21.00	14.99	2.46	6.00	24.50
Mean Cards	2.48	0.49	0.50	4.50	2.70	0.65	0.75	6.00
Mean CardRate	0.18	0.04	0.04	0.31	0.19	0.05	0.06	0.52
Mean NetGoals	0.55	0.55	-1.00	1.60	-0.15	0.78	-3.50	1.50
Mean Goals	1.44	0.45	0.00	2.50	1.12	0.48	0.00	2.75
Possession	0.50	0.09	0.22	0.77	0.50	0.09	0.20	0.80
Mean Possession	0.50	0.03	0.41	0.645	0.50	0.05	0.31	0.66
NetPoints	0.21	1.35	-3.40	4.02	-0.06	1.45	-5.13	5.13
Mean Reds	0.17	0.19	0	1	0.18	0.19	0	2
Observations	425				1599			

Table 5 - Descriptive Statistics by Home Status

	Home				Away			
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
CardRate	0.16	0.10	0.00	0.88	0.21	0.12	0.00	0.80
NetGoals	0.62	1.69	-5.00	8.00	-0.62	1.69	-8.00	5.00
Possession	0.55	0.08	0.31	0.80	0.45	0.08	0.20	0.69
Observations	1012				1012			

Table 6

Dependent Variable: CardRate

Comparison Groups: Brazilian (playing non-Brazilian) vs. Non-Brazilian (playing non-Brazilian)

	(1)	(2)	(3)	(4)	(5)
Brazil	0.000	-0.004	0.003	0.004	0.003
	(0.999)	(0.581)	(0.711)	(0.591)	(0.626)
Away	0.036***	0.036***	0.037***	0.036***	0.036***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
NetPoints	-0.006**	-0.006*	-0.005*	-0.005*	-0.005*
	(0.007)	(0.014)	(0.024)	(0.019)	(0.020)
Crowd	-0.004	-0.005	0.000	0.000	0.001
	(0.768)	(0.699)	(0.985)	(0.970)	(0.958)
Track	0.002	0.001	-0.002	-0.005	-0.005
	(0.724)	(0.871)	(0.794)	(0.489)	(0.455)
Possession	-0.160***	-0.164***	-0.147***	-0.148***	-0.148***
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)
Mean Possession	0.148	0.132	0.200*	0.186*	0.185*
	(0.085)	(0.122)	(0.013)	(0.019)	(0.020)
Mean Fouls		-0.005***	-0.012***	-0.001	-0.001
		(0.000)	(0.000)	(0.832)	(0.821)
Mean Cards			0.066***	0.001	0.000
			(0.000)	(0.923)	(0.981)
Mean CardRate				0.916***	0.897***
				(0.000)	(0.000)
Mean Reds					0.015
					(0.352)
Constant	0.175***	0.263***	0.146***	-0.016	-0.011
	(0.000)	(0.000)	(0.000)	(0.758)	(0.829)
Observations	1221	1221	1221	1221	1221
R^2	0.08	0.09	0.20	0.22	0.22

p-values in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table 7

Dependent Variable: CardRate

Comparison Groups: non-Brazilian (playing Brazilian) vs. Brazilian (playing Brazilian)

	(1)	(2)	(3)	(4)	(5)
Foreign	0.021	0.022	0.011	0.012	0.012
	(0.138)	(0.120)	(0.398)	(0.381)	(0.367)
Away	0.033*	0.033**	0.031**	0.031**	0.031*
	(0.011)	(0.009)	(0.010)	(0.010)	(0.010)
NetPoints	0.003	0.003	-0.000	0.000	0.000
	(0.414)	(0.474)	(0.971)	(0.905)	(0.961)
Crowd	0.006	-0.002	-0.006	-0.007	-0.005
	(0.800)	(0.922)	(0.771)	(0.748)	(0.811)
Track	0.007	0.007	0.005	0.006	0.006
	(0.614)	(0.563)	(0.658)	(0.626)	(0.635)
Possession	-0.080	-0.077	-0.105	-0.106	-0.106
	(0.325)	(0.336)	(0.171)	(0.163)	(0.162)
Mean Possession	-0.087	-0.127	-0.011	0.017	0.050
	(0.560)	(0.391)	(0.940)	(0.906)	(0.732)
Mean Fouls		-0.007**	-0.013***	-0.003	-0.003
		(0.001)	(0.000)	(0.546)	(0.500)
Mean Cards			0.058***	0.006	0.015
			(0.000)	(0.807)	(0.587)
Mean CardRate				0.785*	0.763*
				(0.033)	(0.039)
Mean Reds					-0.039
					(0.268)
Constant	0.233***	0.369***	0.260***	0.093	0.069
	(0.000)	(0.000)	(0.001)	(0.388)	(0.530)
Observations	397	397	397	397	397
R^2	0.06	0.08	0.17	0.18	0.18

p-values in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table 8

Dependent Variable: NetGoals

Comparison Groups: Brazilian (playing non-Brazilian) vs. Non-Brazilian (playing non-Brazilian)

	(1)	(2)	(3)	(4)
Brazil	0.613***	0.611***	0.013	0.013
	(0.000)	(0.000)	(0.905)	(0.903)
Away	-1.379***	-1.467***	-1.444***	-1.444***
	(0.000)	(0.000)	(0.000)	(0.000)
NetPoints	0.303***	0.292***	0.157***	0.156***
	(0.000)	(0.000)	(0.000)	(0.000)
Crowd	-0.157	-0.184	-0.559**	-0.559**
	(0.427)	(0.351)	(0.002)	(0.002)
Track	-0.021	-0.011	0.086	0.087
	(0.848)	(0.923)	(0.382)	(0.378)
Possession	-0.942	-1.786*	-1.551*	-1.551*
	(0.116)	(0.012)	(0.016)	(0.016)
Mean Possession		2.917*	0.786	0.800
		(0.029)	(0.515)	(0.509)
Mean NetGoals			1.034***	1.041***
			(0.000)	(0.000)
Mean Goals				-0.016
				(0.896)
Constant	1.237***	0.255	1.461**	1.472**
	(0.000)	(0.655)	(0.005)	(0.005)
Observations	1221	1221	1221	1221
R^2	0.21	0.21	0.36	0.36

p-values in parentheses, * p<0.05, ** p<0.01, *** p<0.001

Table 9

Dependent Variable: NetGoals Comparison Groups: non-Brazilian (playing Brazilian) vs. Brazilian (playing Brazilian)				
	(1)	(2)	(3)	(4)
Foreign	-0.719**	-0.716**	-0.051	-0.082
	(0.002)	(0.002)	(0.815)	(0.708)
Away	-1.347***	-1.523***	-1.542***	-1.544***
	(0.000)	(0.000)	(0.000)	(0.000)
NetPoints	0.193**	0.172**	0.079	0.076
	(0.003)	(0.009)	(0.182)	(0.198)
Crowd	0.470	0.377	-0.122	-0.108
	(0.198)	(0.300)	(0.712)	(0.744)
Track	-0.091	-0.070	0.158	0.154
	(0.673)	(0.741)	(0.414)	(0.425)
Possession	-2.493*	-4.302**	-4.858***	-4.883***
	(0.029)	(0.001)	(0.000)	(0.000)
Mean Possession		6.426**	5.238*	5.053*
		(0.009)	(0.018)	(0.022)
Mean NetGoals			1.030***	1.162***
			(0.000)	(0.000)
Mean Goals				-0.321
				(0.136)
Constant	1.730*	-0.445	0.170	0.667
	(0.011)	(0.676)	(0.859)	(0.510)
Observations	399	399	399	399
R^2	0.16	0.18	0.34	0.35

p-values in parentheses, * p<0.05) ** p<0.01, *** p<0.001

Appendix A: Variable Names and Definitions

Variable Name	Description	Original Data Sources
CardRate	Total number of team cards divided by the total number of team fouls	Footystats.org (2020)
NetGoals	goal differential between the team *i* and its opponent.	Footystats.org (2020)
Brazil	Dummy variable equal to one if team is Brazilian, zero otherwise	Created by the authors
Foreign	Dummy variable equal to one if team is not Brazilian, zero otherwise	Created by the authors
Away	Dummy variable equal to one if team is playing away, zero otherwise	Footystats.org (2020)
Crowd	Match attendance as a share of stadium capacity	WorldFootball.net (2020)
Track	Dummy variable equal to one if match stadium has track, zero otherwise	Worldstadiums.com (2020)
Mean Fouls	Average team fouls in season	Footystats.org (2020)
Mean Cards	Average team cards in season	Footystats.org (2020)
Mean CardRate	Average CardRate in season	Footystats.org (2020)
Mean NetGoals	Average goal differential in season	Footystats.org (2020)
Mean Goals	Average goals scored in season	Footystats.org (2020)
Possession	Team possession rate in each match	Footystats.org (2020)
Mean Possession	Average team possession in season	Footystats.org (2020)
NetPoints	Difference between the teams' standardized ranking for each match	CONMEBOL (2017) and authors
Mean Reds	Average red cards in season	Footystats.org (2020)