The Woman in the Arena: Evidence of Nonsexual Harassment on a Male-dominated Field *

Thiago de Lucena†

December 2, 2024

*I would like to thank Keisuke Teeple, Kangoh Lee, Ryan Abman, and Rafael Rosa for their valuable contributions to this study.

†San Diego State University

Abstract

This study investigates nonsexual harassment differentials in inter-gender interactions in the workplace. Exploring virtual quasi-random assignment of referees to matches and analyzing over 1,200 Brazilian top professional soccer divisions match logs, I find that the presence of a female assistant referee at a match is associated with a statistically significant 51% increase in disciplinary warnings awarded to players for harassment of match officials. I present a simple toy model that contemplates two potential mechanisms for this bias: perceived different expected returns of harassment based on the target gender and intrinsic motivations. Results suggest that players believe their chances of influencing decisions through harassment are higher when targeting female officials, particularly in high-stakes situations. The introduction of video assistant referees (VAR) eliminates differential harassment, supporting the notion that different perceived returns to harassment, rather than intrinsic motivations, drive these behaviors. My findings highlight the prevalence of nonsexual harassment in labor markets, suggesting that women may face greater challenges than men in high-stakes environments. This has implications for gender composition in decision-making roles, as disproportionate relative harassment could be a complementary cause for the lack female preference for high-risk and decision-making positions.

Keywords: Gender Gaps, Discrimination, Football

JEL Codes: P16; J15; J71; J78

1 Introduction

Over the last 60 years, the equality of legal rights obtained by women in Western countries has led to one the most drastic revolutions in labor markets: the entry of women into the workforce (Goldin (2023)). This entry gave rise to the interaction of men and women in the workplace, a now common phenomenon that was unknown to men before this time. The same time period also saw a considerable increase in interest on the analysis of the causes and consequences of bullying and harassment in the workplace (Einarsen et al. (2020)), with light being cast on the negative consequences of these interactions (Folke and Rickne (2022)). Hence, it is fair to ask, do genders differ in the probability of being a target of nonsexual harassment? If so, what are the potential implications? While a few studies have provided evidence of gender-based differential treatment on a professional environment (Card et al. (2020), Dupas et al. (2021), Eberhardt et al. (2023), Wu (2020), Folke and Rickne (2022), Chakraborty and Serra (2024)), real-world causal evidence of differential levels of harassment across genders is virtually nonexistent.

This study provides a causal answer to these question in the context of Brazilian male soccer's top divisions. Given the country's passion for the sport, Brazil's top leagues are highly competitive, high-stakes environments. Referees and assistant referees are frequently the targets of harassment and verbal abuse by players.[1] Furthermore, in contrast to soccer in other countries, soccer in Brazil has been a male-dominated sport until very recently, making it the perfect laboratory to study the interplay between female presence and male behavior.

Using data extracted from more than 1,200 match log files for the pre-video assistant referee (VAR) introduction period and the virtual quasi-random assignment of soccer referees to matches to estimate a fixed effects model, I show that having a female assistant referee is associated with a statistically significant

[1] In the context of this study, harassment is not a synonym of sexual harassment. While the latter implies the former, the converse is not true. In fact, in the context of this work, evidence points to harassment of nonsexual nature, commonly described as bullying. Einarsen et al. (2020) provides a great discussion about the differences in terminology: While bullying and harassment are frequently used interchangeably, Zapf and Einarsen (2005) argue that harassment is a broader concept and encompasses sexual harassment as well as single episodes of severe mistreatment. The frequency dimension of this definition seems particularly suitable to my context; hence, I refer to the mistreatment taking place in this study as harassment, but I caution the reader that this harassment is of particular nonsexual nature and more closely associated to the concept of bullying.

increase in the number of disciplinary warnings (yellow cards) of 0.413 per game awarded to players for harassing match officials. This increase is equivalent to a 70% increase relative to the mean and is robust to a series of fixed effects and controls (51% of the mean when controlled for rank of the assistant referee in the most conservative scenario). I interpret these results as causal evidence of player gender bias in the treatment of assistant referees. A major advantage of my study's context is that assistant referees are not responsible for handling warnings and perform an extremely standard and objective task; biases are unlikely to arise from differences in behavior between the target of the harassment——in this case, women.

Furthermore, I break down cards by subcategories. I show that both subcategories grouped under the definition of harassment-related warnings experience increases when a female is serving as an assistant referee. Warnings due to excessive protesting see a statistically significant increase of 0.364 standard deviations. Cards due to unauthorized field entrance, typically given to players who enter the field to protest a decision, also see an increase of 0.314 standard deviations. All other warning subcategories, such as violent fouls, delaying gameplay, and repeated fouling, among others, do not show any evidence of statistically significant increases under the presence of a female assistant referee. These findings reassure the causal interpretation of my estimates.

To better understand my findings, I develop a model that contemplates two reasons for harassment. First, workers may harass a coworker because they believe that harassment may increase the probability of a favorable outcome in which there is uncertainty in the workplace. For example, a worker may use foul language and offensive words in a call to an information technology (IT) worker in a company because they believe that this could lead to their computer being fixed sooner. Under this mechanism, differences in harassment levels based on the target's gender may arise because men believe that there will be higher marginal returns of harassment when the target is a female worker.

Another reason for harassment contemplated by my model is intrinsic motivation. Workers may harass others simply because they derive utility in doing so, regardless of any favorable outcome. Under this mechanism, differentials in harassment based on the gender of the target may arise because men may simply derive more utility from harassing women. This could happen, for example, due to men deriving utility in boosting their own identities relative to the underrepresented out-group, as suggested by Wu

(2020).

To detangle the two channels, I start by analyzing evidence of the first mechanism. Presumably, if differentials in expected marginal returns of harassment are the reason for different observed levels of harassment, a change in the utility of the outcome should exacerbate these differences. I utilize the second half of matches as a variation for the utility of this outcome. Since Brazilian top soccer divisions are highly competitive, a missed call or mistaken decision in the second half can be crucial for the result of a match. The presence of a female referee in the second half is associated with an additional 0.161 harassment warnings per match, suggesting that players do believe in different expected returns to harassment based on the target gender and implying that higher-stakes situations are associated with even more harassment asymmetries.

I also use the introduction of a VAR to empirically assess the contribution of each mechanism. This technology assisted referees in making key decisions, such as penalties, goals, and red cards, by reviewing video footage, virtually reducing the error rate to zero. The introduction of VAR eliminates any possibility of complaining, protesting, or harassment leading to a change in referee decision, since it means the decision is made after the review of video footage. Hence, any harassment observed in this case must be due to intrinsic motivations. My results show that differences in harassment warnings completely disappear after the introduction of this technology, suggesting little to no economic evidence of the intrinsic motivation channel and that different expected returns to harassment based on the gender of the target are the main driver of results.

Finally, given that I find evidence of the difference in expected returns mechanism, I investigate if these beliefs arise from differences in performance. In other words, if male assistant referees are more likely to make accurate calls, players may be less likely to harass them. A priori, this mechanism is plausible as female assistant referees could be less experienced. To test this hypothesis, I control for referees' quality, proxied by different ranks within the Brazilian Football Federation. I find little evidence that performance drives differences in expectations.

Put together, my main results present causal evidence of gender-based harassment in labor markets. Given the unique nature of my setup, I believe that this result presents to date one of the clearest pieces of

evidence of biases and differential treatment in labor markets. There are two reasons for this claim. First, in my context, harassment warnings are handled by a third party, which rules out differences in harassment driven by differential definitions of harassment across genders, which could be the case in self-reported data. Second, in my context, targets of harassment perform a very standard task, ruling out differences in target behavior as a mechanism.

My mechanism analysis deserves particularly close attention due to its implications. I show that the main mechanism behind these differential harassment levels is differences in expected outcome returns of harassment based on gender. In higher stakes or uncertain situations, men are disproportionately more likely to engage in harassment based on the gender of the target. This result implies that, ceteris paribus, in situations with higher outcome uncertainty or higher stakes, women are more likely to get harassed relative to men. This result suggests that a complementary explanation for gender composition gaps in top or decision-making positions is that women are disproportionately more subject to harassment in these cases. In a perfectly competitive market, this could lead women to require a larger premium for these positions relative to men. Evidence of female response to sexual harassment has already been documented in Folke and Rickne (2022); hence, this mechanism could also be present in a broader harassment context.

It is important to also mention the limitations of my results. First, in terms of setup, I cannot distinguish between direct and indirect harassment in this context. Second, my analysis is done in a highly competitive, highly male-dominated environment; if results in my mechanism section provide any insight, it is that these are the environments in which one expects to see the strongest effects.

This paper makes a few contributions to the literature. First, this paper contributes to the literature that analyzes differential gender treatment and bias in labor markets. Beyond the seminal study about bias in call-back rates by Bertrand and Mullainathan (2004), studies have documented asymmetries in several other dimensions. Particularly close to this study are studies that document discrimination in interactions, such as reference letters, audience behavior, and editorial treatment, among others (Card et al. (2020), Dupas et al. (2021), Eberhardt et al. (2023), Wu (2020)). An often-suggested possibility in this literature is that bias in male behavior arises from an initial deviation from the norm by the female agent. For example, if males interrupt female presenters more frequently, it is because female presenters have a different presentation

style that is more open to questions. I contribute to this literature by conducting my analysis in a situation in which the woman's task is almost fully devoid of subjectivity. Assistant referees perform a task that is almost fully objective and standardized, raising a flag when certain game conditions are met. Hence, in my context, this concern is most likely unfounded.

Next, this work contributes to the literature that studies gender asymmetries in decision-making positions. It is commonly argued that a driver of income and composition gaps between genders is the tendency of female workers to self-select into lower-risk positions. Higher risk aversion among female workers has been well documented (Ertac and Gurdal (2012), Alan et al. (2020), Eckel and Grossman (2008)) and is often cited as a mechanism for this self-selection.[2] My findings suggest a complementary explanation: In high-risk environments, female workers are more likely to experience differential treatment in the form of harassment. If women react to this differential by selecting out of these positions, as documented in Chakraborty and Serra (2024), differential harassment may be yet another reason for gender gaps in top positions.

This study is organized as follows: Section 2 discuss the context for my study. Section 2 describes the data. Section 5 presents a theoretical model and Section 4 discusses the methodology. Section 6 and 7 discuss main and auxiliary results. Finally, Section 8 discuss alternative mechanisms and Section 9 concludes.

2 Context

2.1 Gender Biases in Brazil

According to the World Bank's Women Business and the Law Index 2023, Brazil scores 85.0 on equal legal rights across genders. This places the country just slightly below the average for Europe and Central Asia (85.8), but well above the world average (77.9) and the Latin America average (81.2). Hence, Brazil is representative of a country that, while it has made significant strides toward gender equality, still needs more progress in some sectors to achieve full equality, much like many developed countries in the West.

[2] A non-exhaustive list of other factors linked to the gender gaps in top positions includes negotiation patterns (Bowles et al. (2007), Exley et al. (2020)), competitive behavior (Gneezy et al. (2003), Preece and Stoddard (2015)), and self-stereotyping Coffman (2014).

One of these sectors of low gender equality is soccer. Soccer refereeing has been an area of particular gender imbalances, with the first female refereeing appearance in a Série A division match only taking place in 2003.[3] Female presence in the assistant refereeing position has been historically low, although much more consistent, yet still subject to bias and discrimination from players, coaches, and media.

For example, former Vasco da Gama manager Ramon Diaz attributed a poor decision in his squad's defeat to Bragantino in 2023 to the presence of a woman in the refereeing team by saying that it was "tough that a woman had to make a decision [on the topic]."[4] In 2015, a prominent blog dedicated to discussing referee performance in Brazil attributed the behavior of a female assistant referee to premenstrual syndrome.[5] Therefore, it is possible, if not likely, that this type of bias is reflected in on-field actions.

2.2 Refereeing in Soccer

The field referee trio in a soccer match is composed of a center referee and two assistant referees. The center referee serves as the primary authority on the soccer pitch, responsible for enforcing the laws of the game, making key and final decisions regarding fouls, misconduct, and penalties, and managing the overall flow of play. Their authority includes issuing disciplinary actions, such as yellow and red cards. A yellow card serves as a warning to a player for committing a minor infraction or displaying unsportsman-like actions. Examples of these infractions might include committing a tactical foul, delaying gameplay, committing a violent foul, or complaining about a referee's decision. A red card results in immediate expulsion from the match; hence, it indicates a serious infraction. A referee can award a red card to a player without awarding a yellow card first; however, a player who receives a second yellow card in a match is automatically shown a red.

Assistant referees, formerly referred to as linesmen, also have several key responsibilities in a soccer match. One of their most crucial duties is evaluating the application of the offside rule. Since offsides

[3]This only happened again in 2019, meaning that during the period of my main analysis, 2016-2017, there was no female presence in center refereeing.

[4]https://www.cnnbrasil.com.br/esportes/futebol/vasco-da-gama/apos-fala-machista-sobre-var-ramon-diaz-do-vasco-se-desculpa-veja/

[5]https://www.apitonacional.com.br/noticias/elas-estao-descontroladas.html. The target of these comments, Tatiane Sacilotti, is part of the study sample.

depend uniquely on the relative positioning of attacking and defending players at the moment the ball is played, enforcing the rule is a purely objective task. Furthermore, assistant referees are responsible for signaling when the ball has completely crossed the goal line or sideline, indicating whether a throw-in or a corner kick should be awarded. Finally, they are also responsible for assisting center referees in signaling fouls near their own sideline. While this latter responsibility is subject to some level of subjectivity, it only represents a minor duty for assistant referees, implying that the tasks performed by them are mainly of an objective nature.

It is noteworthy to mention that assistant referees have no authority to award players cards of any color but are subject to the same level of respect in treatment as any other authority in the game. This means that if a player vehemently or aggressively protests a decision made by an assistant referee, the central referee should show the player a yellow card. Unlike in other sports, this type of player behavior is very common in soccer, as players frequently protest both referees' and assistant referees' decisions, often using disrespectful and vulgar language, despite this behavior sometimes leading to warnings. This behavior is particularly common after an offside call, as these situations tend to be pivotal in a game, representing the difference between a goal being scored or not. In these cases, players usually rush to the sideline to confront the assistant referee, often surrounding him or her (see Figure A.1).

2.3 Soccer in Brazil

Soccer is the most popular sport and a national passion in Brazil. According to recent estimates, 68% of the Brazilian population with internet access reports being interested in the sport.[6] Although differences have narrowed over the past few years, this interest is highly heterogeneous between genders, with males reporting considerably more interest in the sport, while women actually faced legal barriers to playing it for nearly 40 years.[7] Given the high interest and participation among its population, it is unsurprising that Brazil has been very prolific at both the national and club levels in the male modality of the sport. At

[6]https://kantaribopemedia.com/conteudo/68-dos-brasileiros-com-acesso-a-internet-sao-fas-de-futebol/

[7]From 1941 to 1979, women in Brazil were prohibited from practicing sports that were "incompatible with their nature." https://www.terra.com.br/esportes/futebol/por-que-futebol-ainda-e-esporte-so-para-homem-no-brasil,854d92eb673c47c921ef5f7cdb8898ea6mbs7xzy.html

the national level, Brazil is the only country to have won five World Cups. At the club level, Brazil's top division, Série A, usually ranks among the top domestic leagues in the world, despite lacking the financial resources to compete with richer European leagues.[8]

The Campeonato Brasileiro Série A was established in 1971 and now features 20 teams that compete in a round-robin format, similar to other top leagues in the world. It is generally played from April to December, with the team with the most points at the end of the season crowned champion, while the bottom four teams are relegated to Série B. Série B, also established in 1971, serves as the second tier and includes 20 teams competing similarly, with the top four earning promotion to Série A and the bottom four facing relegation to Série C.

The soccer pyramid in Brazil, including Série A and Série B, is governed by the Confederação Brasileira de Futebol (CBF). Founded in 1914, it oversees all aspects of the sport in the country at the national level, including the national team, domestic leagues, and compliance with international standards set by the Fédération Internationale de Football Association (FIFA). As such, the CBF is responsible for assigning referees to matches, and it does so through its refereeing branch, the Refereeing Commission (*Comissão de Arbitragem*), which is responsible for the training, appointment, and oversight of referees and match officials. The Refereeing Commission also maintains the Refereeing Board, a ranking of top referees in the country who are certified to referee professional matches at the national level. To reach this board, referees must first register at their state-level federation, display outstanding performance for many years, and then have their associated federation appoint them to the Refereeing Board. Hence, any national-level referee in the Refereeing Board is generally considered part of the right tail of the quality distribution of referees in Brazil.

State federations are the entry point for a refereeing career in Brazil. They differ greatly in structure, governance, and norms but generally offer initial training and certification for match officials as well as act as a marketplace to link supply and demand for this type of service. For example, to be part of the local board of the Rio Grande do Sul State Refereeing Commission, applicants must first enroll in a six-month course with classes three times per week. They must then pass a theoretical exam and many physical tests,

[8] https://iffhs.com/posts/3336

which include running more than 2,700 m in 12 minutes (2,400 for women). Once admitted to the board, they are required to complete updating courses, certifications, and exams. The criteria for appointment to the national Refereeing Board is unclear; however, Rio Grande do Sul State can only appoint a handful of officials to the national ranks. Either way, referees at the national level are individuals who possess a considerable amount of training and skill.

Once at the national Refereeing Board, referees are awarded ranks that qualify them to officiate matches at different levels. These ranks have changed over time, but during the period of my study, they could be classified into four categories: *CD*, the lowest qualification, qualifies referees to work only in matches of the third and fourth national divisions. *AB* extends this qualification to the first and second divisions, while *Master* does the same with an additional "badge of quality."[9] Finally, *FIFA* is the most prestigious rank and allows referees to work in international matches in addition to any national-level competition.

These ranks are important for the assignment of referees to matches. This assignment is done largely, although not entirely, at random. First, the CBF excludes referees associated with a certain state federation from refereeing matches from that state. Then, the CBF claims to select referees to participate in a draw for each match, publicly releasing the name of the center referee drawn (and assistants). In practice, the CBF has faced a lot of backlash from the media and fans for the lack of transparency and randomness in this process. In Section 4, I provide evidence that this is the case; assistants are chosen mostly at random, in line with the selection process in other sports (Price and Wolfers (2010), Parsons et al. (2011)).

3 Data

This study uses data for two seasons of Brazilian soccer. I focus on the 2017 and 2018 seasons for a few reasons. First, female presence prior to 2017 was low, and this was the earliest year for which data was available at the time of collection. Second, the post-2018 season saw the implementation of VAR, a technology that dramatically altered the landscape of the interactions between players and referees.[10]. Match

[9]This rank is typically used for older referees in a later stage of their careers.

[10]This work eventually uses data for this period in the mechanisms section

reports for all matches were downloaded in pdf format from the official website of the *Confederação Brasileira de Futebol*. Finally, the period post 2018 also saw a change in the system that assigns referees to matches. The reports for 2017 and 2019 are fairly uniform and were converted into data using statistical software. They contain information about players, referees, goals, and other match measures (see the appendix for an example of one of these reports). Importantly, these reports contain detailed and standardized information about disciplinary measures; they indicate the player, color, minute, and reason that a card was awarded (see Figure A.2).

CBF allows referees to award a yellow card for seven different reasons. Figure 1 displays the histograms for each reason. By far, the most common reason for awarding a player a card is violent fouling, with an average of more than three cards per game awarded for that reason. The second most common reason for awarding a player a card is complaining, with more than 0.7 cards a game awarded for that reason. Unauthorized field entrance is another important reason in our context, as the name implies referees should award a player or a coach a yellow card if they enter the field without authorization. In practice, this is done almost exclusively when substitute players or managers want to protest or complain; that is, they enter the field to talk to the referee and end up receiving a yellow card for unauthorized entrance. Since this category is also related to complaining, my main results include it in the protesting category.

Some variables used in this study were not directly listed in these match reports. This is the case for the gender of referees and some match controls. For referee gender, I predict gender from the first name of officials, using the Brazilian census to make this prediction. As for match controls, I create categories based on the clubs involved in a given match. I define a club as being popular if it was part of an organization called *Clube dos Treze*, a now-defunct organization that was responsible for acting in the interest of the 13 largest clubs in Brazil. Furthermore, I define a derby as a match between any two clubs in any city in the first or second Brazilian division. Descriptive statistics for the sample can be seen in 1.

Finally, weather data was scraped from open-meteo.com. I extract game day weather rounded to the closest hour for all stadium coordinates in the Brazilian Série A and B. This dataset includes wind speed, cloud coverage, rain, and temperature.

4 Methodology

To better understand the impact of female presence in the refereeing trio on the behavior of athletes, I estimate the following model:

$$Y_m = \beta_0 + \delta_1 F_m + \beta X'_m + \theta_r + \kappa_h + \tau_y + \rho_d + u_m \tag{1}$$

where Y_m is outcome Y of match m, X_m is a vector of match specific characteristics, θ_r are referee fixed effects, θ_r are year fixed effects, κ_h are home team fixed effects, and ρ_d are division fixed effects. The referee fixed effects implies that this model allows for the possibility that female assistant referees associate themselves with harsher referees, despite there being little evidence of that. κ_h allows that CBF assigns female referees to stadiums that are more hostile. The identifying assumption is that, conditional on the controls, the female presence on a trio in a given match is random. While I cannot prove this assumption, I can provide evidence of its validity. Figure 2 and Table 2 display results of estimating my model with several different match characteristics as dependent variables (and excluding X'_m). Confidence intervals indicate that referee trios with female presence are not any more likely to be assigned to matches with of clubs, derbies, primetime matches, or others. These results increase our confidence in interpreting estimates as causal.

5 Harassment Toy Model

Consider a worker facing the choice of optimal amount of harassment to exert, $t \in [0, 1]$, about a decision. One can think about a worker calling an IT department in order to get his personal computer fixed sooner rather than later. In this case, the difference is whether the computer will be fixed sooner or later and the worker's goal is to have it fixed sooner. The worker believes that the IT attendant has some discretion over the decision. The worker can also receive an warning for harassment. Hence, there are two events in this context, which for simplicity, I model independently: $\gamma = \{Goal, \overline{Goal}\}$ and $\upsilon = \{Warning, \overline{Warning}\}$. Let P_γ the set of probabilities associated with γ outcomes, such that $P_\gamma = \{p_g(t), (1 - p_g(t))\}$ and P_υ the set

of probabilities associated with v outcomes, such that $P_v = \{p_w(t), (1 - p_w(t))\}$.[11]

Now assume that the player has preferences over the outcome such that $Goal \succcurlyeq \overline{Goal}$ and $\overline{Warning} \succcurlyeq Warning$. Since preferences are continuous, transitive and complete, it is possible to show that utility functions U_g: $\gamma \rightarrow \mathbb{R}$ such that $U_g(Goal) \geq U_g(\overline{Goal})$ and U_w: $v \rightarrow \mathbb{R}$ such that $U_w(\overline{warning}) \geq U_w(warning)$ exist. Assume that the worker's utility function is:

$$V(G, W, t) = U_g(G) + U_w(W) + \theta * t \tag{2}$$

And its expected value:

$$E(V(G, W, t)) = [p_g(t)U(g) + (1 - p_g(t))U(\overline{g})] + [p_w(t)U(w) + (1 - p_w(t))U(\overline{w})] + \theta * t \tag{3}$$

Assume, for simplicity that $U(\overline{G}) = U(\overline{W})$= 0. Notice that his implies that $U_g(g) \geq 0$ and $U_w(w) \leq 0$. A necessary condition for the optimal, if it exists, is that[12]

$$\frac{\delta E(V(G, W, t))}{\delta t} = p'_g(t^*)U(g) + p'_w(t^*)U(w) + \theta = 0 \tag{4}$$

Consider now the case in which a worker faces this choice under the presence of a female or male worker, s = {m,f}. Assume that the utility of reaching a goal or receiving a warning is the same in both cases, but players may have different intrinsic utilities of harassment θ^s. This difference may arise for several reasons; for example, men may draw more utility from harassing women relative to other men because they like to boost their own identities relative to the underrepresented out-group in a profession, much as in the effect documented in Wu (2020) or because they have desire to protect their occupational status as documented in Goldin (2014). A player can also have different probability functions over having a goal validated p_g^s based on the gender of the assistant referee. The rationale for this is that it is possible that the worker believes that a woman is more likely to change her decision for a given level of harassment.

[11] In terms of notation $p_g(t) = p(G = Goal|t)$ and $U(g) = U_g(g) = U_g(Goal)$.

[12] Assume that $U(G)$ is large enough such that $\frac{\delta^2 E(V(G,W,t))}{\delta t^2}|_{t=t^*} = p''_g(t^*)U(g) + p''_{\overline{g},w}(t^*)U(w) < 0$ and that V(t^*) > V(0) and V(t^*) > V(1) so that we have an maximum a t^*

$$\frac{\delta E(V^m(G,W,t))}{\delta t} = p_g^{m\prime}(t^*)U(g) + p_w'(t^*)U(w) + \theta^m = 0 \tag{5}$$

$$\frac{\delta E(V^f(G,W,t))}{\delta t} = p_g^{f\prime}(t^{**})U(g) + p_w'(t^{**})U(w) + \theta^f = 0 \tag{6}$$

Subtracting equation 5 from 6:

$$(p_g^{f\prime}(t^*) - p_g^{m\prime}(t^{**}))U(g) + (p_w'(t^*) - p_w'(t^{**}))U(w) + (\theta^m - \theta^f) = 0 \tag{7}$$

This equation leads us to two very important testable implications:

1. If $\frac{\delta t^* - t^{**}}{\delta U} \neq 0$ and $(\theta^m = \theta^f)$ then it must be that $p_g^{f\prime}(t) \neq p_g^{m\prime}(t)$

   To see this assume by way of contradiction that $\frac{\delta t^* - t^*}{\delta U} \neq 0$ and $(\theta^m = \theta^f)$ but $p_{g,\overline{w}}^{f\prime}(t) = p_{g,\overline{w}}^{m\prime}(t), \forall \in \mathbb{R}^+$. Since both $(\theta^m = \theta^f)$ and $p_{g,\overline{w}}^{f\prime}(t) = p_{g,\overline{w}}^{m\prime}(t), \forall \in \mathbb{R}^+$, it must be that $t^* = t^{**}$ which imply $(p_{g,\overline{w}}^{f\prime}(t^*) - p_{g,\overline{w}}^{m\prime}(t^{**})$ which in turn leads to the contradiction with $\frac{\delta t^* - t^*}{\delta U} \neq 0$. In other words, if $(\theta^m = \theta^f)$ and $\frac{\delta t^* - t^*}{\delta U} \neq 0$ then $p_{g,\overline{w}}^{f\prime}(t) \neq p_{g,\overline{w}}^{m\prime}(t)$. I will evaluate $\frac{\delta t^* - t^*}{\delta U} \neq 0$ in Section 7.1 assuming $(\theta^m = \theta^f)$ and then evaluate this assumption in Section 7.2.

2. If $p_g^{s\prime}(t) = 0 \ \forall t \in [0,1], s \in \{f,m\}$ then any difference between t^{**} and t^* is a consequence of differences in $(\theta^m - \theta^f)$

   To see this assume without loss of generality $(\theta^f > \theta^m)$. If $p_g^{s\prime}(t) = 0 \ \forall t \in [0,1], s \in \{f,m\}$ then $(p_g^{f\prime}(t^*) - p_g^{m\prime}(t^{**})) = 0$ for any given t^* and t^{**} . Hence, it must be that $(p_w'(t^*) - p_w'(t^{**}))U(w) > 0$. Since $U(w) < 0$ it must be that $(p_w'(t^*) - p_w'(t^{**})) < 0$. Since $\frac{d^2p}{dt^2} < 0$, it must be that $t^{**} < t^*$

6 Main Results

I start my analysis by estimating the model mentioned in Section 4. These estimates can be seen in Figure 3 and Table 3 for several different specifications.[13] My preferred specification controls for referees,

[13]These results do not include other controls, although they are robust to it and can be seen in Figure A.3 in the appendix. These controls are mostly uncorrelated with the outcome variable.

division and year fixed effects, and cluster standard errors at home team level.[14] The results imply that having a female as an assistant referee is associated with a significant increase in violations for harassment of 0.413 per game; this is equivalent to an 70% increase relative to the mean. Furthermore, this result is robust to a series of different fixed effects specifications, including the exclusion of all fixed effects. Interestingly, adding referee fixed effects actually slightly increases point estimates, suggesting that, if anything, female assistant referees tend to associate themselves with more lenient referees.

These results present initial evidence of bias in the treatment of female assistant referees. Point estimates are not only statistically significant but also economically significant. Given the extremely large magnitude of these estimates, especially relative to the mean, it is important to discuss the external validity of these findings. Two factors can be crucial in contributing to the large magnitude of these estimates. First, the context of these findings is one of large gender inequality; that is, only 6% of assistant referees are female, meaning that men may feel even more comfortable engaging in harassment based on gender. Second, this is a context in which decisions are very high stakes, which, if the results of my next section can be trusted, also suggests that this may increase harassment. Hence, my results cannot be extrapolated to all sectors of the economy. Nevertheless, while not all sectors of modern labor markets have these characteristics, it is important to note that many still do, such as top positions and science, technology, engineering, and math fields.

Next, I analyze whether the female presence affects cards awarded for other reasons by presenting a breakdown of cards by category. This breakdown casts light on the behavior of center referees in the presence of female assistants. If referees are becoming more rigorous in the presence of female assistants, one would expect to see impacts in other categories of cards. It is noteworthy to point out that there is no theoretical backing for this argument; that is, there is no formal guideline for center referees suggesting an increase in rigor in the presence of a female assistant referee. Nevertheless, this analysis can help confirm that no such increase occurs. This analysis also allows us to break down the cards for protesting into two sub-categories——excessive complaining and unauthorized field entry. While both of these categories are grouped under warnings for complaining, they have a somewhat different nature. While excessive

[14]The inclusion of home team fixed effects increases standard errors while leading to small changes in coefficients.

complaining has a somewhat subjective definition, leaving room for interpretation from the center referee, unauthorized field entry is a very objective category: if a player or a coach enters the field of play to complain, they should be awarded a warning. Hence, this category can provide evidence about players' behavior while keeping fixed referees' actions.

Results for this analysis are displayed in Figure 4 and Table 4. First, we notice that for all non-complaining related categories, having a woman as an assistant referee is not associated with a significant increase in repeated fouling, removing a jersey, delaying gameplay, not respecting wall distance, or committing a violent foul. It is noteworthy to mention that this last category is the most common warning category, providing a sizable amount of variation and power to test my hypothesis. I interpret this as evidence that referees are not changing their behavior in response to having a female assistant. Moreover, when we analyze protesting categories, we observe that both sub-categories present significant results; a female presence increases cards for complaining by 0.364 standard deviations and cards for unauthorized field entrance by 0.314 standard deviations. With the latter being a category with very little subjectivity and discretion for referee decision-making, I find no evidence that my main results are driven by changes in center referees' behavior.

It is important to take the time to ensure the correct interpretation of this result. Two stories are possible in my context: First, it is possible that after a controversial decision by a female referee, players or coaches harass the assistant referee directly with offensive language and tone, prompting the center referee to issue a warning in the form of a yellow card. Alternatively, it is possible that players and coaches are more aggressive directly towards the referee in response to the decision of a female assistant referee. In my context, while anecdotal evidence suggests that in the case of off-sides the former is more common, it is impossible to distinguish between direct and indirect harassment of this type. Nevertheless, while it may appear that indirect harassment is non-consequential to women in this or broader contexts, it is easy to imagine a world in which, at equilibrium, workers refrain from working with other women because they get harassed on behalf of female workers.

7 Mechanisms

Given that differences in referee behavior were already ruled out, differences in amount of harassment can arise for two reasons: Differences on expected marginal returns of harassment ($p_g^{f\prime}(t^*) - p_g^{m\prime}(t^{**})$) and differences on intrinsic preferences ($(\theta^m - \theta^f)$). To detangle these two reasons, I follow the implications of the model presented in Section 5. To summarize the discussion, the latter can be thought of as the difference in probability that an assistant referee changes their decision based on their gender. For example, players may believe that females are more likely to change their decision if they get harassed. The former simply reflects a difference in taste for harassment in the presence of each gender: Players may chose higher harassment in the case of a female referee because they may want to boost their own identities relative to the underrepresented out-group (Wu (2020)). If $(p_g^{f\prime}(t^*) - p_g^{m\prime}(t^{**})) \neq 0$, then changes in the utility of scoring a goal should lead to relative changes in the amount of harassment. If $(p_g^{f\prime}(t^*) - p_g^{m\prime}(t^{**})) = 0$ then any observed difference in harassment must be a consequence of differences intrinsic preference for harassment.

7.1 Expected marginal returns of harassment

My first mechanism test is to analyze the heterogeneity regarding half of play.[15] The assumption is that given the lack of time left, having a favorable decision awarded to a squad in the second half is far more likely to be pivotal on the determination of the final result, leading to considerable more utility compared to the first half.[16] This provides a great source of variation in utility to detangle the mechanisms. If part of overall differences in harassment by players arise from differences on expected marginal returns of harassment, a change in utility of the outcome should lead to relative changes in harassment. Hence, if we observe any difference in harassment across halves, we have evidence that $(p_g^{f\prime}(t^*) - p_g^{m\prime}(t^{**})) \neq 0$ and that there are differences in expected marginal returns of harassment across genders. To do so I first modify my outcome variable to now reflect cards by halves rather than cards by game. Table 5 shows the results for this analysis for my preferred specification.

[15] Soccer is played in two halves of 45 minutes plus stoppage time.

[16] An assumption here is that matches are competitive in this context. This assumption is backed by empirical evidence—70% of matches on the Brazilian Serie A in 2017 and 2018 ended with a score difference equal to or lower than one goal. For a histogram of the goal difference distribution, see Figure A.4 in the appendix.

First, with a point estimate of 0.260 it is important to notice that second halves are associated with more yellow cards for harassment in general. Next, we see that female presence is still associated with more harassment cards on the first half—female presence is associated with an increase in harassment warnings of 0.124 . Finally, female presence in the second half is associated with a increase in cards of of 0.161 significant at 10%. This point estimate suggests that female presence on a referee trio is associated with more than twice the additional number of cards as a female presence on the first half.

We cannot, however, interpret the borderline significant coefficient for female presence as evidence of intrinsic motivation. Even in the first half, there are still reasons for players to attempt to alter an assistant referee decision by harassing them. Hence, all that we can say is that I find evidence of differences on expected marginal returns of harassment across genders. Despite serving as a test of only one of the mechanisms, this analysis, it is important to point out, has the advantage that it holds fixed several match characteristics—since we are are comparing halves within games, characteristics such as weather and game type are kept fixed—decreasing potential for omitted variables in this analysis.

7.2 Intrinsic Utility

Given that mechanisms are not mutually exclusive I now evaluate if my main results can also be explained by changes in differences in intrinsic utilities of harassment. The ideal setup to conduct this analysis would be a situation in which changing harassment levels would not lead to any changes in the probability of a favorable decision ($p_g^{f\prime}(t^*) = p_g^{m\prime}(t^{**}) = 0$) also implying that $p_g^{f\prime}(t^*) - p_g^{m\prime}(t^{**}) = 0$. In this case, any difference in warnings observed is due to differences in intrinsic utility of harassment. Luckily, there is a natural experiment that allows us to empirically conduct this analysis.

Introduced in Serie A in Brazil in 2019, the video assistant referee system represents a significant advancement in officiating technology within professional soccer, aimed at enhancing the accuracy and fairness of match decisions. Introduced to mitigate the impact of human error, VAR operates under a framework that allows for video review of specific match incidents. The primary areas of intervention include the validation of goals, penalty decisions, direct red card incidents, and the rectification of mistaken identity among players. By establishing a formal protocol for these situations, VAR seeks to reduce controversy and

ensure that critical decisions reflect the true nature of the game.

The VAR process involves a team of officials situated in a centralized video operations room, equipped with access to multiple camera feeds and advanced replay technology. When a potential reviewable incident occurs, the on-field referee is called the VAR team, who analyze the footage and communicate their findings. The on-field referee retains ultimate authority over the decision, but can choose to review the footage on a pitch-side monitor if necessary. This collaborative approach aims to uphold the integrity of the game while providing referees with additional resources to make informed decisions. In practice, this technology virtually reduced the marginal increase in a favorable decision probability of harassment to zero, given that now VAR would ultimately have inaccurate calls reversed. In particular, for objective decisions such as offsides, players were now assured that the right decision would be made. Hence, after VAR introduction any difference in player behavior must be due to intrinsic motivation. To analyze harassment due to intrinsic motivation I estimate the following canonical difference-in-difference model:

$$Y_m = \beta_0 + \delta_1 F_m + \delta_2 Post_VAR_m + \beta X'_m + \theta_r + \kappa_h + \eta y + u_m \tag{8}$$

The main limitation of this approach is that, different from the test described in Subsection 7.1 this analysis does not does not keep some match conditions fixed. Rather, it relies on the timing of an event for identifying the effect of a certain mechanism. Any other shock that happens in the same time period and has differential impacts across genders is likely to be confounded to the effects of VAR. [17] Furthermore, another limitation of this analysis is that it focus only on the first division, as it had a clear introduction date.

Figure 6 presents this analysis. We observe that in the years post-VAR introduction the point estimate for female presence on a referee trio decreases by nearly 0.5 warnings per game. This is almost exactly the same coefficient for female presence, meaning that, for the years post-VAR introduction the total effect of a female presence is null. This result suggests that there is no evidence of differences in intrinsic motivation for the harassment channel. Practically, this result is suggestive that once VAR was introduced, all uncertainty regarding a final referee decision was removed, as now referees almost always make the right call

[17] One of these shocks, the COVID-19 pandemic, has its impacts evaluated in Figure A.5 in the appendix.

using video. This reduced the incentives that players may have had to harass assistant and center referees in order to try to make them change their call. Once these incentives were gone we do not observe any difference in harassment based on the gender of the assistant referee, suggesting that players have little to no intrinsic motivation for harassment in this context.

8 Alternative Mechanisms

8.1 Performance Asymmetries

A complementary mechanism for the primary and secondary findings of this study is the presence of performance asymmetries among assistant referees based on gender; that is, male assistant referees may simply be more efficient in signaling off-sides or throw-ins. This explanation does not imply that female referees are intrinsically less efficient; rather, it suggests that they may have less experience, which could contribute to their reduced performance in certain tasks. Therefore, a priori, this explanation is plausible and deserves attention.

To evaluate this mechanism, I control for the rank of assistant referees. If female assistant referees are, on average, less efficient than males, this should be reflected in their rank, and controlling for rank would help address this difference. Finding similar results while controlling for rank suggests that there are no gender-based performance asymmetries; however, finding considerable differences across models implies that gender-based performance asymmetries or other mechanisms may be at play.[18]

Results for this analysis follow the format of the main table of this study and can be seen in Figure 7. Overall, we observe a small decrease in the magnitude of point estimates compared to Figure 3 (0.347 compared to 0.413 of the main results table), suggesting that some small performance asymmetry may be taking place. Nonetheless, results are still significant at the 5% significance level and with a magnitude that still represents almost 58% of the mean. Even in the most conservative scenario (column 6) the magnitude of the effects is 51% of the mean. Hence, I conclude that performance asymmetry is not a main driver of my

[18] Another theory consistent with this result would be that players display less respect towards lower-ranked assistant referees, regardless of whether their performance is equal to that of higher-ranked referees.

findings.

9 Conclusion

In conclusion, this study provides compelling evidence of gender-based harassment in labor markets, specifically within the context of Brazilian male soccer. By using a virtual quasi-random assignment of assistant referees to matches, I analyze the behavior of players towards female assistant referees. I show that a female presence in the assistant referee position leads to a statistically significant 70% (51% of the mean in the most conservative scenario) increase in warnings for harassment. My findings shed light on how biases manifest in workplace interactions. Unlike other contexts, my setup allows me to rule out that bias arises due to differences in target behavior or definitions of harassment. This suggests that men are more likely to engage in harassment when the target is female, further reinforcing the notion that gender dynamics significantly influence workplace behavior.

I present a model that contemplates two explanations for differential levels of harassment based on target gender: different perceived returns of harassment based on target gender and different intrinsic motivations. I test these two mechanisms using a situation in which stakes are higher and the introduction of VAR. In both cases, results indicate that differential perceived returns of harassment based on target gender are the drivers of the results. Finally, I analyze whether this perception can be driven by differences in the performance of the target, finding no evidence to support this theory.

Hence, this study indicates that men are more likely to resort to intimidation tactics against women in high-stakes environments where such behavior may influence outcomes. Extensive research has shown that women often self-select into lower-risk jobs, a trend frequently cited as a contributing factor to gender gaps in labor markets. If the findings of this study can be generalized to other sectors, they suggest that one underlying reason for this risk aversion may be that, in competitive settings, men perceive stronger incentives to employ intimidation strategies to achieve their objectives when facing female colleagues. This dynamic could help explain the persistent disparities in workplace representation and advancement between genders.

References

Alan, S., Ertac, S., Kubilay, E., and Loranth, G. (2020). Understanding gender differences in leadership. *The Economic Journal*, 130(626):263–289. 7

Bertrand, M. and Mullainathan, S. (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American economic review*, 94(4):991–1013. 6

Bowles, H. R., Babcock, L., and Lai, L. (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and human decision Processes*, 103(1):84–103. 7

Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2020). Are referees and editors in economics gender neutral? *The Quarterly Journal of Economics*, 135(1):269–327. 3, 6

Chakraborty, P. and Serra, D. (2024). Gender and leadership in organisations: The threat of backlash. *The Economic Journal*, 134(660):1401–1430. 3, 7

Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, 129(4):1625–1660. 7

Dupas, P., Modestino, A. S., Niederle, M., Wolfers, J., et al. (2021). Gender and the dynamics of economics seminars. Technical report, National Bureau of Economic Research. 3, 6

Eberhardt, M., Facchini, G., and Rueda, V. (2023). Gender differences in reference letters: Evidence from the economics job market. *The Economic Journal*, 133(655):2676–2708. 3, 6

Eckel, C. C. and Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of experimental economics results*, 1:1061–1073. 7

Einarsen, S. V., Hoel, H., Zapf, D., and Cooper, C. L. (2020). *Bullying and harassment in the workplace: Theory, research and practice*. CRC press. 3

Ertac, S. and Gurdal, M. Y. (2012). Deciding to decide: Gender, leadership and risk-taking in groups. *Journal of Economic Behavior & Organization*, 83(1):24–30. 7

Exley, C. L., Niederle, M., and Vesterlund, L. (2020). Knowing when to ask: The cost of leaning in. *Journal of Political Economy*, 128(3):816–854. 7

Folke, O. and Rickne, J. (2022). Sexual harassment and gender inequality in the labor market. *The Quarterly*

Journal of Economics, 137(4):2163–2212. 3, 6

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The quarterly journal of economics*, 118(3):1049–1074. 7

Goldin, C. (2014). A pollution theory of discrimination: male and female differences in occupations and earnings. In *Human capital in history: The American record*, pages 313–348. University of Chicago Press. 14

Goldin, C. (2023). Why women won. Technical report, National Bureau of Economic Research. 3

Parsons, C. A., Sulaeman, J., Yates, M. C., and Hamermesh, D. S. (2011). Strike three: Discrimination, incentives, and evaluation. *American economic review*, 101(4):1410–1435. 11

Preece, J. and Stoddard, O. (2015). Why women don't run: Experimental evidence on gender differences in political competition aversion. *Journal of Economic Behavior & Organization*, 117:296–308. 7

Price, J. and Wolfers, J. (2010). Racial discrimination among nba referees. *The Quarterly journal of economics*, 125(4):1859–1887. 11

Wu, A. H. (2020). Gender bias among professionals: an identity-based interpretation. *Review of Economics and Statistics*, 102(5):867–880. 3, 4, 6, 14, 18

Zapf, D. and Einarsen, S. (2005). Mobbing at work: Escalated conflicts in organizations. 3

Tables

TABLE 1: DESCRIPTIVE STATISTICS OF MATCHES

	2017		2018	
	Mean	S.D.	Mean	S.D.
Type:				
Large Club Derby	0.174	(0.379)	0.205	(0.404)
Large Club	0.455	(0.498)	0.426	(0.495)
Derby	0.042	(0.201)	0.053	(0.223)
Primetime	0.142	(0.349)	0.143	(0.351)
Weekend	0.537	(0.499)	0.543	(0.498)
Weather:				
Rain	0.129	(0.466)	0.151	(0.571)
Wind Speed	12.919	(6.462)	12.803	(6.152)
Cloud Coverage	39.984	(33.672)	42.750	(32.912)
Temperature	23.574	(4.508)	23.332	(4.534)
Center Referee Status:				
A/B	0.741	(0.438)	0.614	(0.487)
FIFA	0.229	(0.420)	0.228	(0.420)
Other	0.030	(0.171)	0.158	(0.365)
Assistant Referee Status:				
Female Assis. Ref.	0.071	(0.257)	0.054	(0.226)
A/B	0.680	(0.360)	0.705	(0.348)
FIFA	0.152	(0.273)	0.167	(0.292)
Other	0.168	(0.320)	0.128	(0.249)
Warnings:				
Violent or Tactical Foul	3.647	(1.950)	3.963	(1.835)
Harassment	0.593	(0.827)	0.584	(0.804)
Other	0.487	(0.778)	0.533	(0.766)
N	760		760	

Note: Simple difference in means; no fixed effects are included. Standard error in parenthesis. Harassment warnings are defined as yellow cards for complaining or unauthorized field entry. Large clubs are defined as those members of *Clube dos Treze*. Source: CBF match logs.

TABLE 2: FALSIFICATION TEST

	Match Type					Weather			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	CB 13 Derby	CB 13	Derby	Primetime	Weekend	Rain	Temp.	Cloud	Wind S.
Female Assistant Ref.	0.096	0.055	0.180	-0.046	0.137	0.057	-0.118	-0.046	0.020
	(0.113)	(0.085)	(0.128)	(0.141)	(0.129)	(0.129)	(0.128)	(0.137)	(0.143)
Mean	0.177	0.450	0.042	0.145	0.518	0.140	23.453	41.367	12.861
N	1514	1514	1514	1514	1514	1514	1514	1514	1514

Note: Includes referee, division, and year fixed effects. The sample includes all matches in Brazilian first and second divisions for the 2017 and 2018 seasons. Standard error in parenthesis. Derbies are defined as matches played by teams within the same city. Primetime games are defined as matches played either Wednesday at 9:30pm or Sunday at 4:00pm as these are typically the matches broadcast on open TV networks. Large clubs are defined as those members of *Clube dos Treze*; see the appendix for a list of both *Clube dos Treze* members and derbies list. Source: CBF match logs.

TABLE 3: IMPACT OF FEMALE ASSISTANT PRESENCE ON HARASSMENT WARNINGS

	(1)	(2)	(3)	(4)	(5)	(6)
Female Assistant Ref.	0.248***	0.418***	0.417***	0.413***	0.381***	0.342***
	(0.071)	(0.083)	(0.083)	(0.083)	(0.081)	(0.128)
Ref. FE	No	Yes	Yes	Yes	Yes	No
Division FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	No
Ref p/ Year FE	No	No	No	No	No	Yes
Home Team FE	No	No	No	No	Yes	Yes
Mean	0.59	0.59	0.59	0.59	0.59	0.59
N	1520	1514	1514	1514	1514	1509

Note: Standard error in parentheses. Fixed effects used are described on the table. The sample includes all matches in Brazilian first and second divisions for the 2017 and 2018 seasons. Standard error clustered at the home team level when home team fixed effects were not included. Robust standard errors were used when home team fixed effects were included. Harassment warnings are defined as yellow cards for complaining or unauthorized field entry. For a figure including controls see the appendix. Source: CBF match logs.

TABLE 4: IMPACT OF FEMALE ASSISTANT PRESENCE ON PROTESTING WARNINGS (CARDS) BY CATEGORIES

	Harassment Related		Non-Harassment Related				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Complaining	Unaut. Entr.	Foul	Delay	Repeated	Wall	Jersey
Female Assistant Ref.	0.364***	0.314	-0.031	0.143	-0.074	-0.009	0.037
	(0.075)	(0.199)	(0.132)	(0.114)	(0.062)	(0.108)	(0.156)
N	1514	1514	1514	1514	1514	1514	1514

Note: The sample includes all matches in Brazilian first and second divisions for the 2017 and 2018 seasons. Includes referee, division, and year fixed effects. Standard error in parentheses. Harassment warnings are defined as yellow cards for complaining or unauthorized field entry. All subcategories are defined by CBF. Source: CBF match logs.

TABLE 5: IMPACT OF FEMALE ASSISTANT PRESENCE ON PROTESTING WARNINGS BY HALF

	(1)	(2)
Female A. Ref.	0.124*	0.108
	(0.069)	(0.069)
2nd Half.	0.260***	0.260***
	(0.021)	(0.021)
Female A. Ref. x Second Half	0.161*	0.161*
	(0.083)	(0.083)
N	3040	3040

Note: The sample includes all matches in the Brazilian first and second divisions for the 2017 and 2018 seasons. Includes referee, division, and year fixed effects. Standard error in parentheses. Harassment warnings are defined as yellow cards for complaining or unauthorized field entry. All subcategories are defined by CBF. Source: CBF match logs.

TABLE 6: IMPACT OF VAR AND FEMALE ASSISTANT PRESENCE ON HARASSMENT WARNINGS

	(1)	(2)
Female A. Ref.	0.310**	0.317**
	(0.138)	(0.161)
Post-VAR	0.000	0.000
	(.)	(.)
Fem. A. Ref.xPost-VAR	-0.334*	-0.328*
	(0.176)	(0.187)
Ref. FE	Yes	Yes
Division FE	Yes	Yes
Year FE	No	Yes
Ref p/ Year FE	No	No
Home Team FE	Yes	Yes
N	1892	1892

Note: The sample includes all matches in the Brazilian first and second divisions for the 2017 to 2021 seasons. Different fixed effects. Standard error in parentheses. Harassment warnings are defined as yellow cards for complaining or unauthorized field entry. Source: CBF match logs.

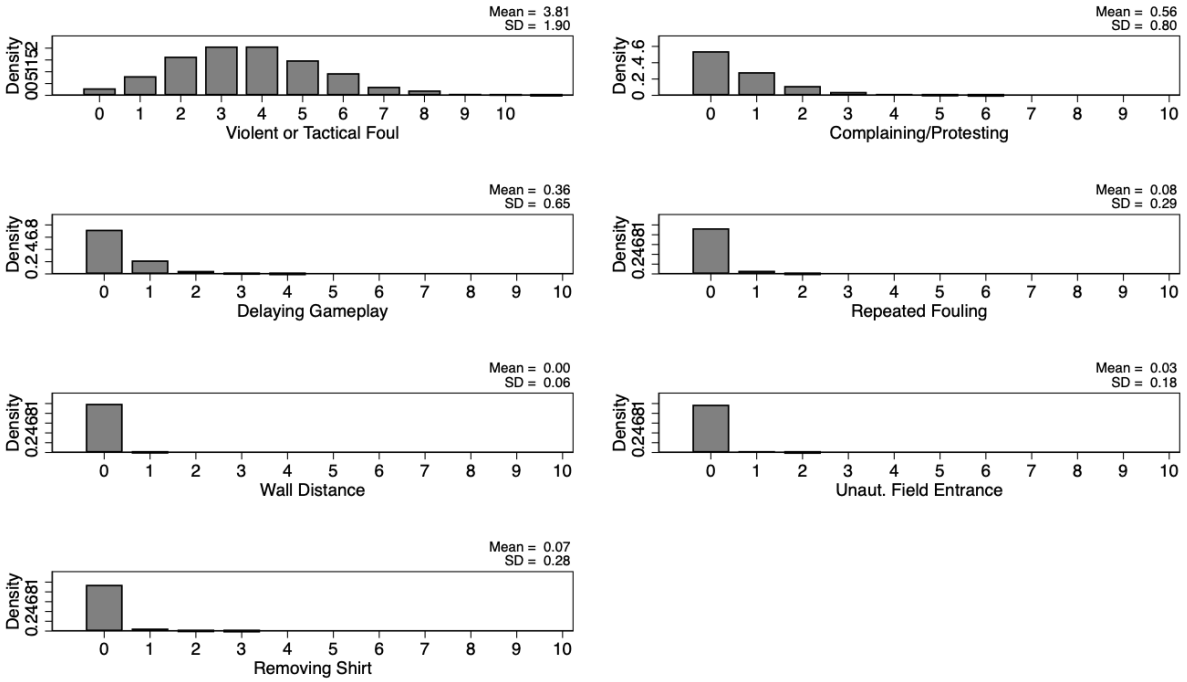
TABLE 7: IMPACT OF FEMALE ASSISTANT PRESENCE ON HARASSMENT WARNINGS

	(1)	(2)	(3)	(4)	(5)	(6)
Female Assistant Ref.	0.215**	0.414***	0.411***	0.392***	0.347***	0.301**
	(0.102)	(0.110)	(0.108)	(0.111)	(0.130)	(0.128)
Ref. FE	No	Yes	Yes	Yes	Yes	No
Division FE	No	No	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	No
Ref p/ Year FE	No	No	No	No	No	Yes
Home Team FE	No	No	No	No	Yes	Yes
Mean	0.59	0.59	0.59	0.59	0.59	0.59
N	1520	1514	1514	1514	1514	1509

Note: Standard error in parentheses. Fixed effects used are described on the table. The sample includes all matches in the Brazilian first and second divisions for the 2017 and 2018 seasons. Standard error is clustered at the home team level when home team fixed effects are not included. Robust standard errors are used when home team fixed effects are included. Harassment warnings are defined as yellow cards for complaining or unauthorized field entry. For a figure including controls see the appendix. Source: CBF match logs.

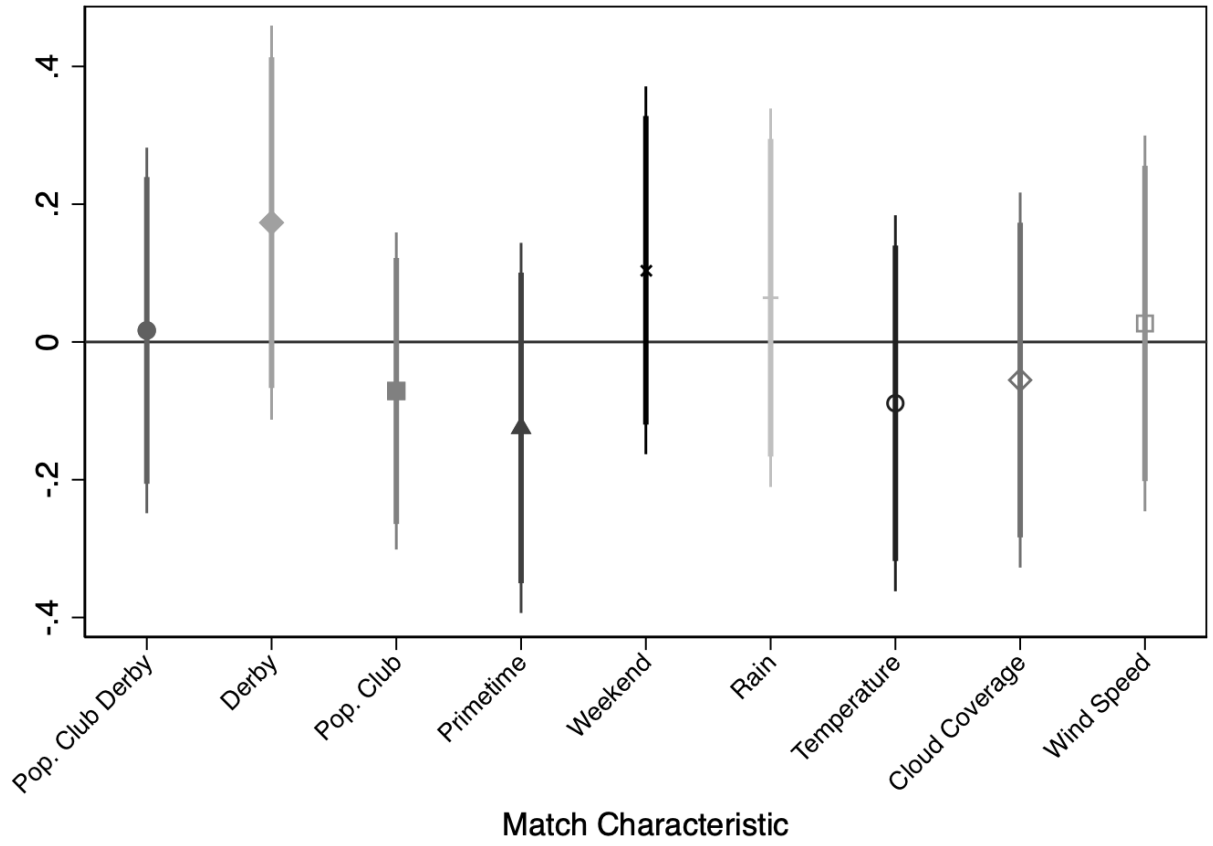
Figures

FIGURE 1: HISTOGRAM OF YELLOW CARDS BY CATEGORY



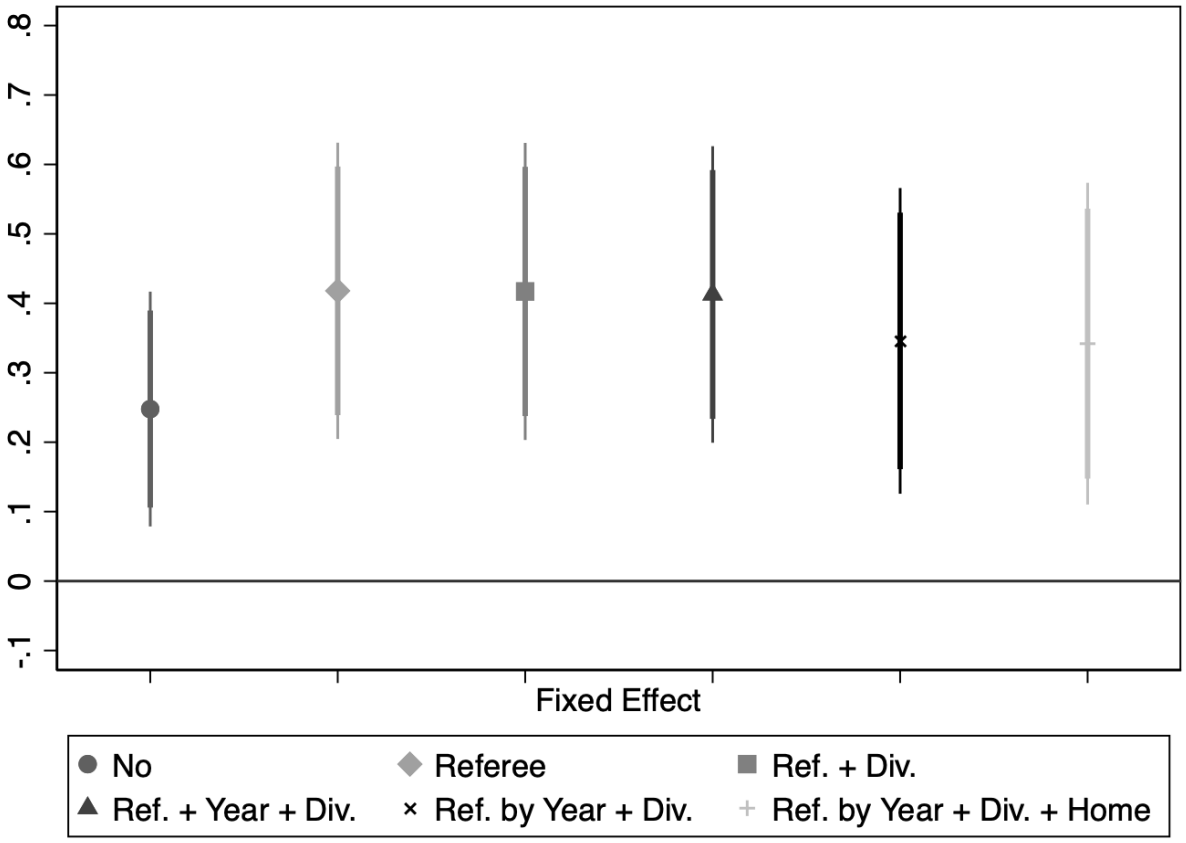
Note: The figure displays histograms of all cards awarded to players according to CBF's card categories in the 2017 and 2018 season of the Brazilian first and second divisions. Mean and standard deviation of each variable are reported on the upper right corner. Source: CBF Match Logs

FIGURE 2: CORRELATION BETWEEN MATCH CHARACTERISTICS AND FEMALE PRESENCE



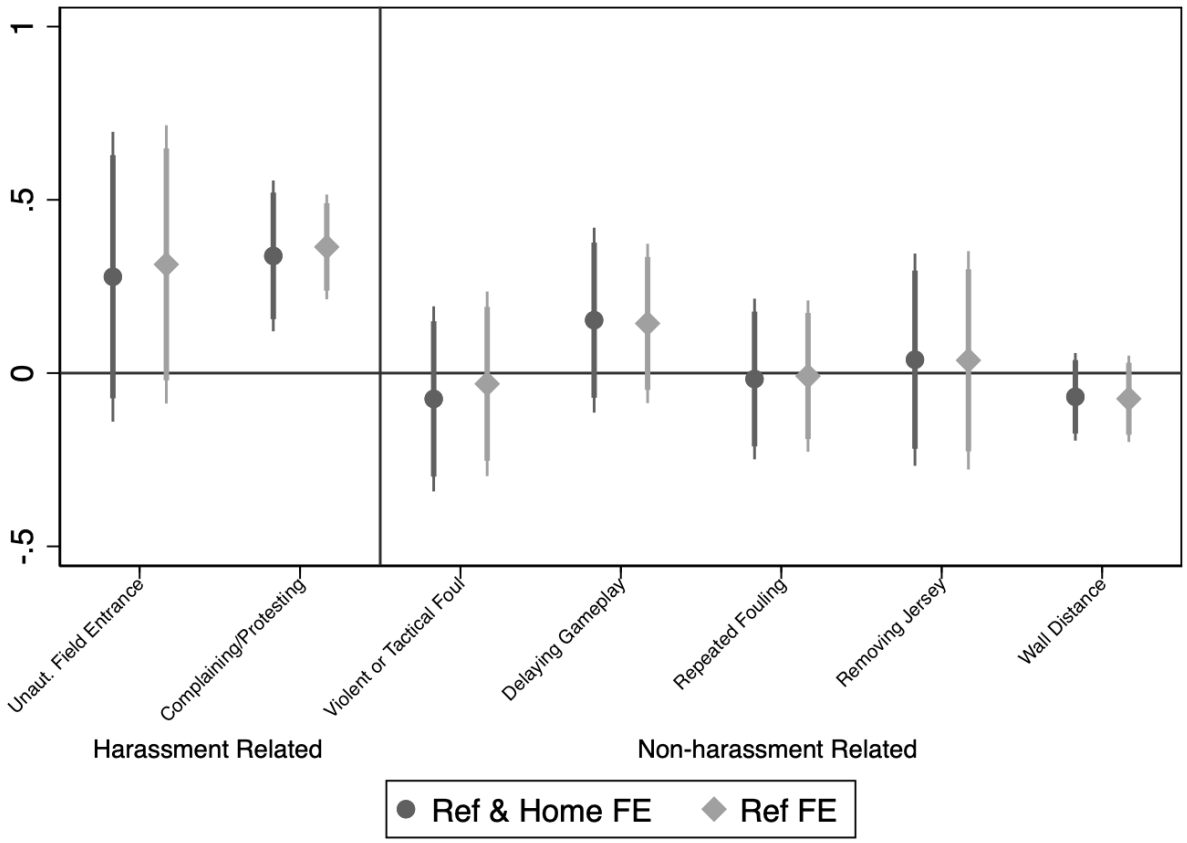
Note: Includes referee, division, and year fixed effects. The sample includes all matches in Brazilian first and second divisions for the 2017 and 2018 seasons. Harassment warnings are defined as yellow cards for complaining or unauthorized field entry. Large clubs are defined as those members of *Clube dos Treze*. Derbies are defined as a match between two clubs from the same city and are listed in the appendix. Source: CBF Match Logs & open-meteo.com

FIGURE 3: EFFECT OF FEMALE PRESENCE ON WARNINGS FOR HARASSMENT



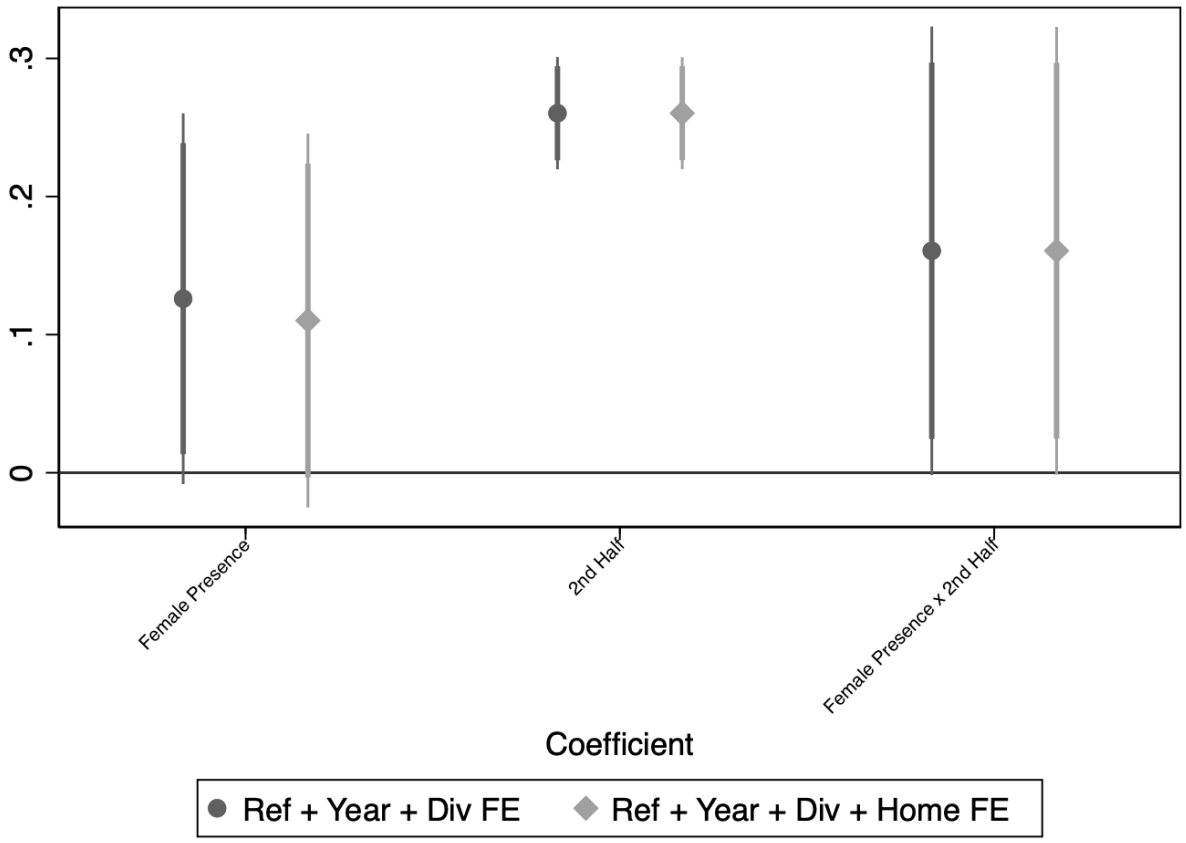
Note: Figure displays point estimates for the impact of having a female assistant referee on the total number of harassment yellow cards per game using different fixed effect specifications. Harassment cards are defined as cards due to excessive complaining and unauthorized field entry. Standard errors are clustered at the home team level unless home team fixed effects are included, in which case robust standard errors are used. Source: CBF Match Logs

FIGURE 4: EFFECT OF FEMALE PRESENCE ON WARNINGS FOR HARASSMENT—BY SUBCATEGORY



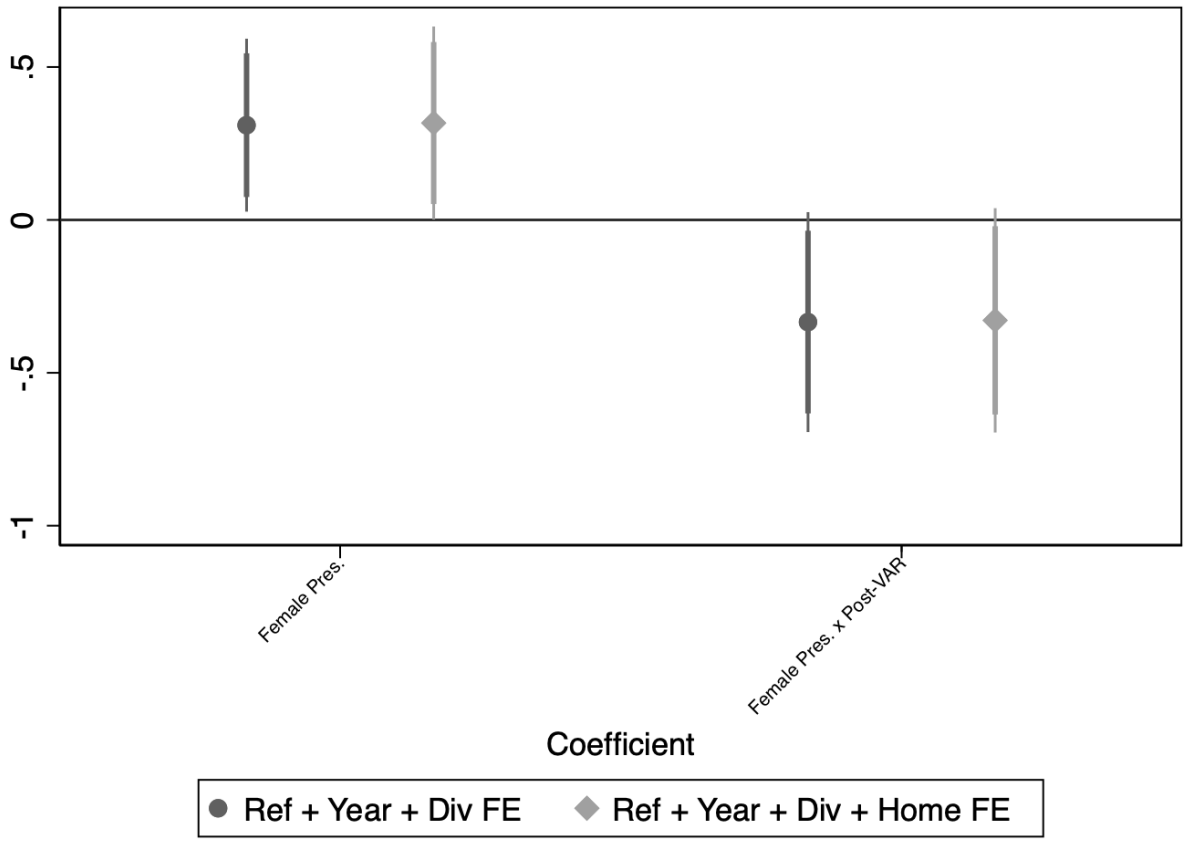
Note: Figure displays point estimates for the impact of having a female assistant referee on the total number of harassment yellow cards per game using preferred fixed effect specifications. Categories defined by CBF. Standard errors clustered at the home team level unless home team fixed effects are included, in which case robust standard errors are used. Source: CBF Match Logs

FIGURE 5: EFFECT OF FEMALE PRESENCE ON WARNINGS FOR HARASSMENT—BY HALF



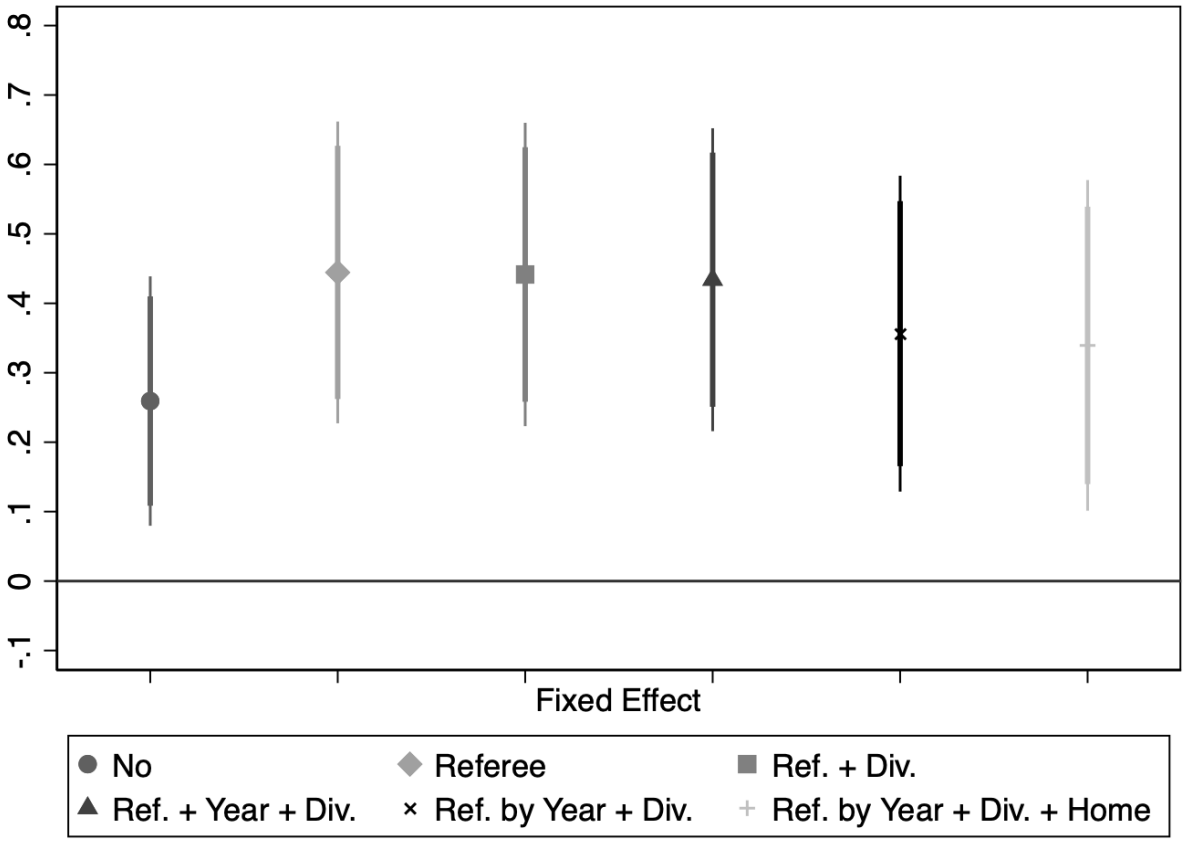
Note: Figure displays point estimates for the impact of having a female assistant referee on the total number of harassment yellow cards by half using preferred fixed effect specifications. Harassment cards defined as cards due to excessive complaining and unauthorized field entry. Standard errors clustered at the home team level unless home team fixed effects are included, in which case robust standard errors are used. Source: CBF Match Logs

FIGURE 6: EFFECT OF FEMALE PRESENCE ON POST-VAR PERIOD



Note: Figure displays point estimates for the impact of having a female assistant referee on the total number of harassment yellow cards post VAR period using preferred fixed effect specifications. Harassment cards defined as cards due to excessive complaining and unauthorized field entry. Standard errors clustered at the home team level unless home team fixed effects are included, in which case robust standard errors are used. Source: CBF Match Logs

FIGURE 7: EFFECT OF FEMALE PRESENCE CONTROLLING FOR REFEREE QUALITY



Note: Figure displays point estimates for the impact of having a female assistant referee on the total number of harassment yellow cards controlling for referee status using preferred fixed effect specifications. Harassment cards defined as cards due to excessive complaining and unauthorized field entry. Standard errors clustered at the home team level unless home team fixed effects are included, in which case robust standard errors are used. Source: CBF Match Logs

A Appendix



FIGURE A.1: EXAMPLE OF ASSISTANT REFEREE HARASSMENT I GOALKEEPER (LÉO MURICI ENDED UP BEING AWARDED A YELLOW CARD)



CBF - CONFEDERAÇÃO BRASILEIRA DE FUTEBOL
SÚMULA ON-LINE

Jogo: 38

Campeonato:	Campeonato Brasileiro - Série A / 2017			**Rodada:**	4
Jogo:	Bahia / BA X Atlético / GO				
Data:	05/06/2017	**Horário:**	20:00	**Estádio:**	Octávio Mangabeira / Salvador

Arbitragem		
Arbitro:	Rodrigo Batista Raposo (AB / DF)	ASSINATURA DIGITAL VÁLIDA
Arbitro Assistente 1:	Jose Reinaldo Nascimento Junior (CBF / DF)	ASSINATURA DIGITAL VÁLIDA
Arbitro Assistente 2:	Daniel Henrique da Silva Andrade (AB / DF)	ASSINATURA DIGITAL VÁLIDA
Quarto Arbitro:	Ciro Chaban Junqueira (AB / DF)	ASSINATURA DIGITAL VÁLIDA
Arbitro Assist Adic 1:	Savio Pereira Sampaio (AB / DF)	ASSINATURA DIGITAL VÁLIDA
Arbitro Assist Adic 2:	Christiano Gayo Nascimento (CD / DF)	ASSINATURA DIGITAL VÁLIDA
Inspetor:	Sergio Cristiano Nascimento (CBF / BRA)	ASSINATURA DIGITAL VÁLIDA

Cronologia

1º Tempo				2º Tempo			
Entrada do mandante:	19:49	**Atraso:**	Não Houve	**Entrada do mandante:**	21:00	**Atraso:**	Não Houve
Entrada do visitante:	19:49	**Atraso:**	Não Houve	**Entrada do visitante:**	20:59	**Atraso:**	Não Houve
Início 1º Tempo:	20:00	**Atraso:**	Não Houve	**Início do 2º Tempo:**	21:02	**Atraso:**	Não Houve
Término do 1º Tempo:	20:47	**Acréscimo:**	2 min	**Término do 2º Tempo:**	21:49	**Acréscimo:**	2 min
Resultado do 1º Tempo: 2 X 0				**Resultado Final: 3 X 0**			

Relação de Jogadores

Bahia / BA

Nº	Apelido	Nome Completo	T/R	P/A	CBF
1	Jean	Jean Paulo Fernandes ...	T(g)	P	420409
3	Tiago	Tiago Pagnussat	T	P	188260
6	Matheus Reis	Matheus Reis de Lima	T	P	337413
8	Allione	Agustin Lionel Allione	T	P	509926

Atlético / GO

Nº	Apelido	Nome Completo	T/R	P/A	CBF
1	Felipe	Felipe Garcia dos Pr...	T(g)	P	172028
3	Ricardo Si ...	Ricardo Cesar Dantas...	T	P	327667
4	Roger	Roger de Carvalho	T	P	184174
5	Marcos	Marcos Antonio Almei...	T	P	345210

FIGURE A.2: CAPTION

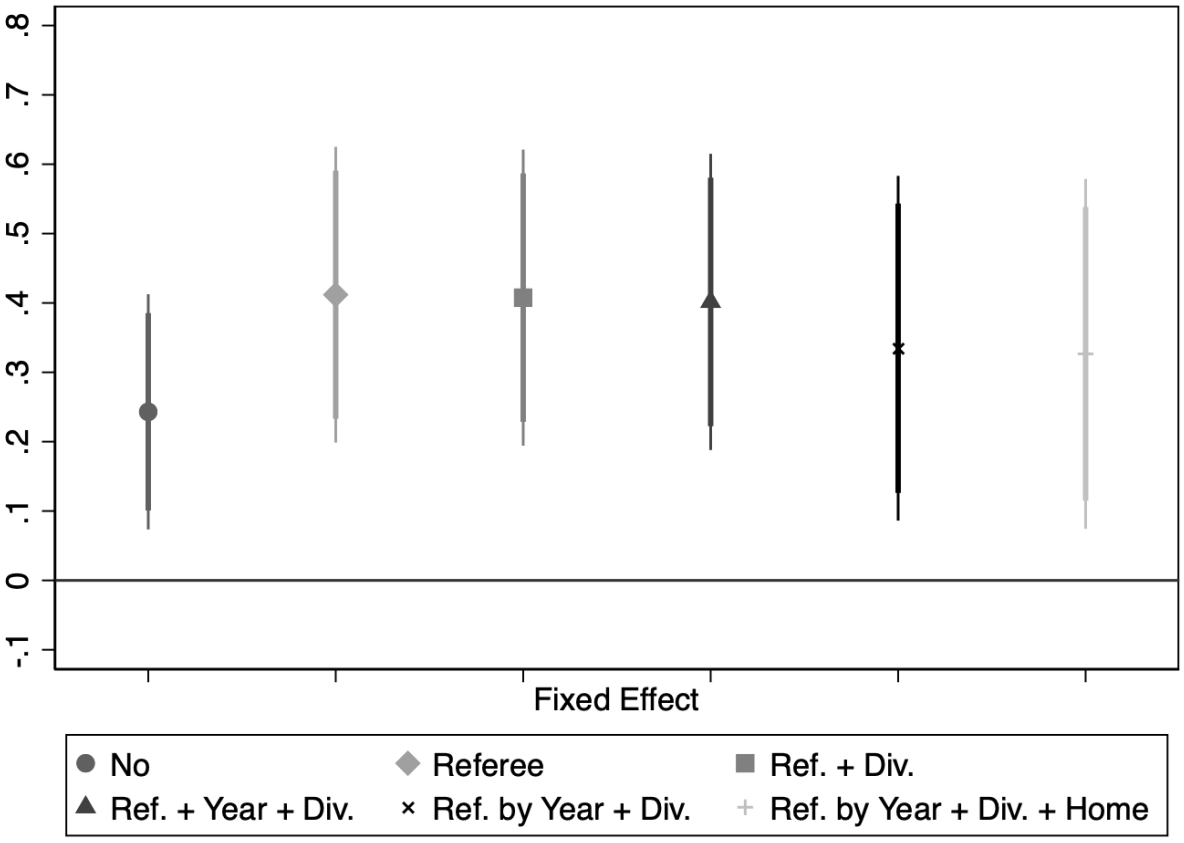


FIGURE A.3: IMPACT OF FEMALE PRESENCE ON THE HARASSMENT WARNINGS INCLUDING CONTROLS

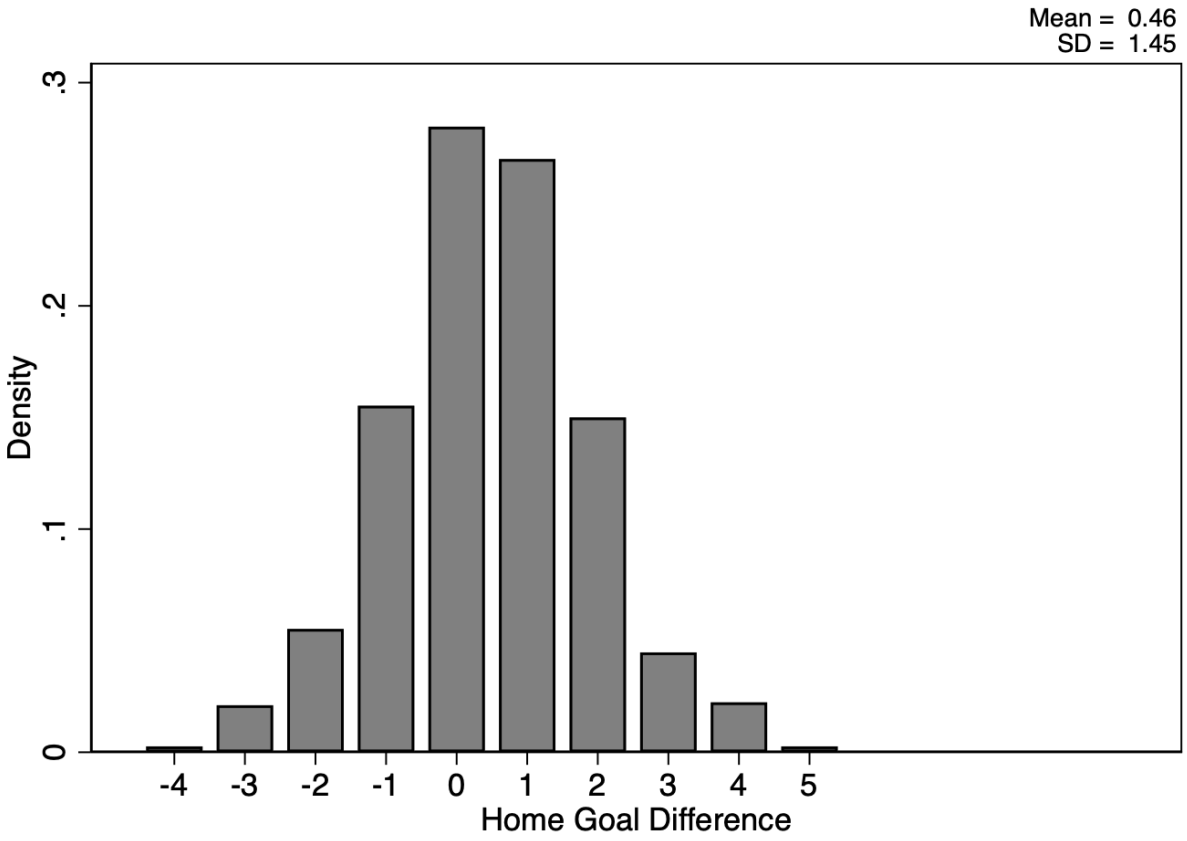


FIGURE A.4: CAPTION

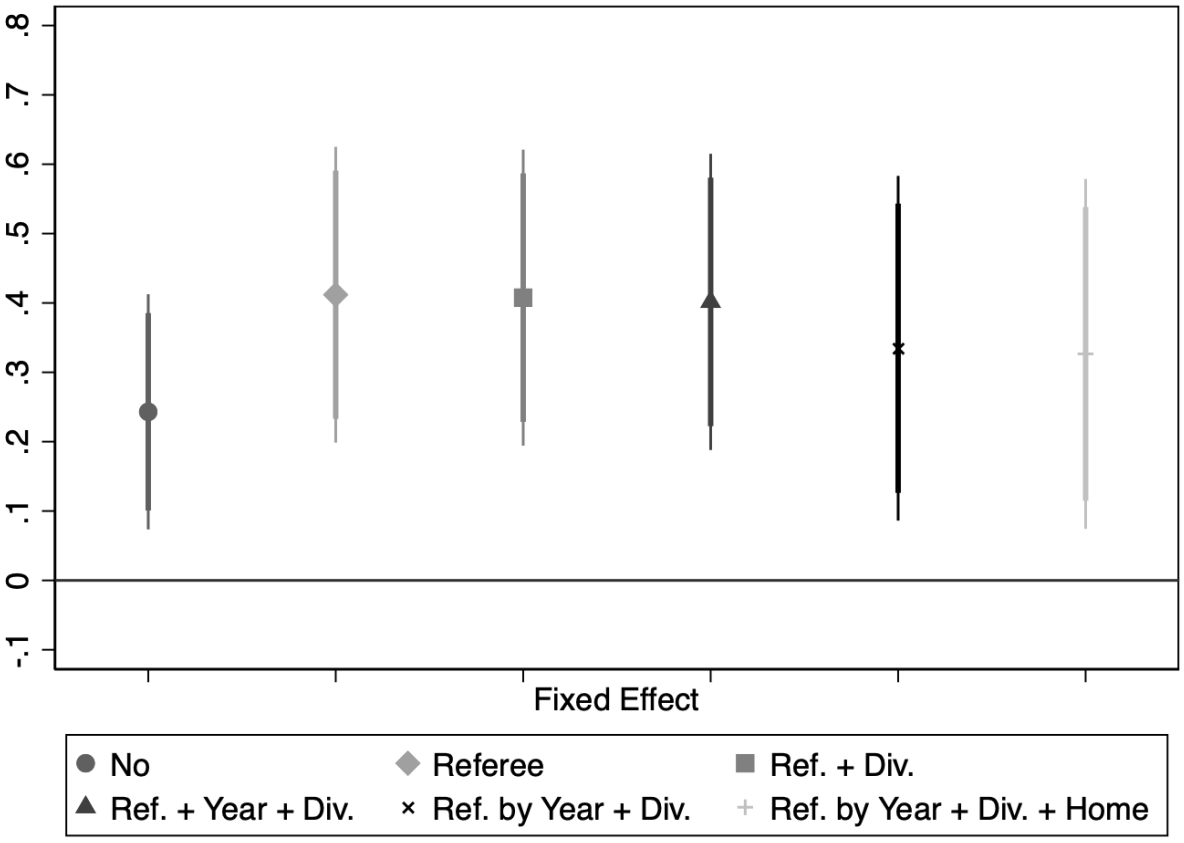


FIGURE A.5: DIFFERENCE-IN-DIFFERENCE—RESULTS EXCLUDING 2020

TABLE A.1: CLUBE DOS TREZE MEMBERS

Club	State
Atletico	MG
Bahia	BA
Botafogo	RJ
Corinthians	SP
Cruzeiro	MG
Flamengo	RJ
Fluminense	RJ
Gremio	RS
Internacional	RS
Palmeiras	SP
Sao Paulo	SP
Santos	SP
Vasco da Gama	RJ

TABLE A.2: DERBY LIST

Derby	City
Atletico vs. Cruzeiro	Belo Horizonte
Atletico-GO vs. Goias	Goiania
Atletico-GO vs. Vila Nova	Goiania
Atlhetico vs. Coritiba	Coritiba
Avai vs. Figueirense	Florianopolis
Bahia vs. Vitoria	Salvador
Botafogo vs. Flamengo	Rio de Janeiro
Botafogo vs. Fluminense	Rio de Janeiro
Botafogo vs. Vasco da Gama	Rio de Janeiro
Corithians vs. Palmeiras	Sao Paulo
Corithians vs. Sao Paulo	Sao Paulo
Corithians vs. Santos	Sao Paulo
Flamengo vs. Fluminense	Rio de Janeiro
Flamengo vs. Vasco da Gama	Rio de Janeiro
Fluminense vs. Vasco da Gama	Rio de Janeiro
Gremio vs. Internacional	Porto Alegre
Goias vs. Vila Nova	Goiania
Guarani vs. Ponte Preta	Campinas
Nautico vs. Sport	Recife
Palmeiras vs. Sao Paulo	Sao Paulo
Palmeiras vs. Santos	Sao Paulo
Sao Paulo vs. Santos	Sao Paulo