An Analysis of Home-Team Bias in the Brazilian Soccer League Using Panel Data

Danilo Lopomo Beteto

April 25, 2010

Abstract

The literature on endogenous preference formation acknowledges that the social environment can affect individual behavior. In this spirit, sports, in particular soccer, constitute a natural setting where such a phenomenon can be studied. The present paper tests for the hypothesis of referees’ home team bias and the influence of the crowd on their behavior, using panel data collected from the brazilian soccer league of 2008. The variables focus of the analysis are amount of injury time, penalty kicks and red cards. Similarly to the results obtained in other leagues, the brazilian one also shows evidence of favoritism. In particular, the number of penalty kicks given to the home team is significantly larger than to the visitor, while the number of red cards is significantly smaller. Moreover, the referees tend to add more extra time when the home team is losing by one goal and less when it’s winning. The crowd doesn’t seem to exert any influence on the referees’ behavior.

Keywords: bias; favoritism; endogenous preferences; referees, soccer.

JEL Classification: D21; D73.
1 Introduction

The principal-agent paradigm is widely used in economics. Different types of situations can be thought of in terms of a principal and an agent. Given this flexibility, it can be applied to contexts where the outcomes are concern to fields other than economics. Sports is one such a field and a particular type of sport, soccer, is the main interest of the present work.

As with any other sport, there’s a well defined body of rules to which soccer is subject. The Fédération Internationale de Football Association, FIFA, is the entity responsible for establishing the rules of the game. Every country has its own Fédération which, among other things, is supposed to enforce the observance of those rules in the national competitions.

Again, as in any other sport, in soccer the referees play a major role in the process of enforcing the rules. However, the referees in soccer have more room to act at their own discretion and hence influence the final result of a match. For their decisions are once and for all, with no change in case they are shown to be wrong, which sometimes happens in other games (e.g. american football and basketball), with the help of the TV or any other electronic device. Among the variables under the “control” of the referees are penalty kicks, red and yellow cards, and the amount of injury time\(^1\).

In this context, it can be said that FIFA, or any other national body in charge of soccer, is a principal while the referee is an agent. The principal demands from the agent to be fair applying the rules of the game, in a way that a victory or a defeat is not to be blamed on the referee. The decisions of the referees should not depend on the teams playing, the place where the game is played or any other factor exterior to the match.

The very fact that referees are human beings, hence prone to errors, makes it hard to judge their decisions. A wrong decision of a referee acting upon bad faith and of a referee who happened to make a true mistake are observationally equivalent. This can be said not only of mistakes but also of any other act of the referees. However, if such mistakes or certain decisions are systematic, then it might be the case that the behavior of the referees are not due to bad faith or true mistakes but rather to other factors involved in the play

\(^1\)In soccer, the clock does not stop for stoppages during the game, e.g. fouls and substitutions. To compensate for the time lost the referees simply add more time at the end of the regular period in both halves of the match. This time added is usually referred to as injury time.
of the game, like the crowd or the place where the match takes place, for example.

The literature on endogenous preferences studies the formation of preferences in response to external factors, or how the social environment can affect individual behavior\(^2\). In soccer, it can be argued that the crowd can exert a lot of pressure on the referees, to the point of influencing their decisions, or simply that the referees might have a home team bias, for example. These two external factors would then impair the fairness of the observance of the rules, therefore influencing the final result of games.

The fact that preferences might be endogenous has testable implications and for the case of soccer, given the availability of data, those implications are subject to statistical scrutiny. For instance, if there’s any sort of home team bias, the referees should favor more frequently the home team than the visitor, which can be inferred, for example, by analyzing the distribution of red cards and penalty kicks, two of the most explicit ways the result of a game can be impacted. The amount of injury time also provides some evidence of home team bias: in close matches, if the home team is losing the referee should add excess injury time, giving more time to the home team to tie the game, with the very opposite happening in the case the home team is winning by a small margin.

Following the above discussion, the purpose of the present paper is to analyze a panel data from the brazilian soccer league in order to test for the presence of favoritism. In the panel data, the individual variables are the referees and the time variables are the rounds of the brazilian league, which consists of 20 teams playing home and away against each other during the season, with a total of 38 rounds. Each referee is assigned to at most one game in any round, but they are not necessarily present in all the rounds\(^3\). As will be detailed below, the data set contains variables such as the amount of injury time, number of yellow and red cards, number of substitutions and number of penalty kicks, among others.

As a preview of the results, there’s evidence of home team bias in the brazilian soccer league of 2008. In particular, using both fixed effects and random effects models, we find that:

- The amount of injury is significantly larger (smaller) when the home team is losing (winning) the game by one goal;

---

\(^2\)See Garicano, Palacios-Huerta, and Prendergast (2005) and the references therein.

\(^3\)This way we have an unbalanced panel data set. However, the referees are assigned by lottery to the matches, hence randomly, and then this issue is not consequential for our analysis.
• In the games where a penalty kick is whistled to the visitor team, a penalty kick is consistently whistled to the home team as well. However, for each penalty kick whistled to the home team, less than one penalty kick is whistled to the visitor, again putting the home team in advantage;

• In the games where there’s a red card for the home team, there’s consistently a red card for the visitor team as well. However, for each red card given to the visitor team, less than one red card is given to the home team, which gives an advantage to the home team;

The paper is structured as follows: section 2 makes a literature review; section 3 presents the data and statistical evidence of favoritism; section 4 analyses which factors might influence referees’ decisions, and section 5 concludes.

2 Literature Review

The joint study of economics and sports is not a new theme. There’re several ways one can make use of economics to analyse outcomes in different games, which has already been done in other works. Just to cite a few, Wolfers (2006) provides an analysis of corruption in NCAA basketball, through the practice of point shaving, in the same spirit of Duggan and Levitt (2002); Walker and Wooders (2001) and Palacios-Huerta (2003) use data from a natural strategic play in professional sports, namely serve-and-return in tennis and penalty kicks in soccer, respectively, to provide an empirical test of the minimax theorem; Chiappori, Levitt, and Groseclose (2002), also analyzing penalty kicks in soccer, study the use of mixed strategies by the players; Brocas and Carrillo (2004) provide a game-theoretical analysis of the incentives for offensive playing coming from two specific changes in the rules of the game, namely the three-point victory and the golden goal; Kahn (2000) (and the references therein) provide an account of the use of professional sports as a natural setting for labor market research.

Regarding endogenous preference formation, Garicano, Palacios-Huerta, and Prendergast (2005) find evidence of favoritism in the spanish soccer league. By analyzing the amount of injury time in each match, they find that when the home team is losing by one goal the referees give on average two minutes more of injury time than when the home team is ahead by one goal, even after controlling for variables that represent true
stoppages during the match like player substitutions and yellow/red cards. They also find evidence that social pressure influences referees’ behavior: the larger (smaller) the crowd, the larger (smaller) is the amount of injury time added when the home team is behind (ahead) by one goal.

It turns out that adding excess injury time is not very consequential, as in very few occasions a goal is scored in the injury time, occasionally changing the final result of a match. More than that, what can be fundamental to the final score of a game is the award of penalty kicks and the distribution of red cards: the former offers a clear possibility of a goal while the later forces the team to play with one less player, which might severely impair the tactical scheme of the team. Sutter and Kocher (2004) take that into account and in the same spirit of Garicano et al. offer additional evidence of home team bias by the referees.

The approach to be followed in the present paper is the same of Garicano et al. and Sutter and Kocher. The data collected allows for the same type of analysis and is described in the next section, together with some summary statistics that provide the main evidence of favoritism in the brazilian soccer league of 2008.

3 Data and Evidence of Favoritism

The main evidence of favoritism found by Garicano et al. is the amount of injury time added at the end of the game. Basically what they show is that if the home team is behind by one goal the injury time is significantly larger than when the home team is ahead by one goal, even after controlling for events that represent true stoppages during the game, like player substitutions and/or yellow/red cards. Sutter and Kocher go further and find that the home team is more likely to be awarded with penalty kicks and less likely with red cards. We first describe the data and then try to see if in the brazilian soccer league there’s evidence of the same type of systematic behavior by the referees.

The data collected is from the brazilian soccer league played in the year of 2008. Twenty teams play a double round-robin tournament, where a victory gives three points and a draw one, no points being awarded in the case of a defeat. After 38 rounds the team with the largest number of points is declared the champion, qualifying for the Libertadores

\footnote{For a detailed description of the brazilian league see http://www.cbf.com.br (in portuguese).}
Cup\textsuperscript{5} together with the other 3 best ranked teams in the final standing; the teams from the 5th to the 12th position in the final standing qualify for the Copa Sudamericana\textsuperscript{6}; finally, the last four descend to the lower division in the following year.

The data collected comes from the official reports of the referees after the games, the so called Súmulas, document made available for every match in the website of the Confederação Brasileira de Futebol, CBF, the brazilian body in charge of soccer in Brazil\textsuperscript{7}. These documents provide all the information about the games played, like the squads, the final score and the time of the goals, the stadium where the match took place, number and time of yellow and red cards, minutes of injury time added at the end of both halves, the name of the referees and an evaluation by them of the behavior of the players, the coaching staff and the crowd, and the report of any other incident that might have occurred during the game, among other things.

Given the double round-robin structure of the tournament and the number of teams, the initial data set consisted of 380 observations. Some of these observations had to be discarded: Botafogo vs Sport in the first round, where 22 minutes of injury time were added by reason of a electricity shortage in the stadium during the second half of the game; Flamengo vs Santos, also in the first round, as Flamengo was being punished with one closed-gates game\textsuperscript{8}; Náutico vs Coritiba (15th round), Atlético Mineiro vs Goiás (21st round), Vasco da Gama vs Cruzeiro (24th round), Figueirense vs Palmeiras (29th round), Figueirense vs Fluminense (32nd round), Internacional vs Fluminense (36th round) and Internacional vs Cruzeiro (37th round), all these either because a page containing an important piece of data included in the analysis was missing in the referee report or the data itself wasn’t readable. After discarding these matches the data set reduced to 371 observations.

Another document made available by CBF is the financial report provided by the teams hosting the matches, the so called Borderôs, whereby they inform the total revenue and expenses incurred in the organization of the match and also the total number of supporters. These numbers will also be considered in the analysis to be made in the

\begin{footnotesize}
\footnotesize\textsuperscript{5}The Libertadores Cup is the analogous of the Champions League played in Europe.
\footnotespace
\footnotesuperscript{6}Copa Sudamericana is the analogous of the UEFA League played in Europe.
\footnotespace
\footnotesuperscript{7}For the games played in the brazilian soccer league of 2008, these documents can be found at \url{http://www.cbf.com.br/php/tabela.php?ct=1&cc=38&aa=2008}.
\footnotespace
\footnotesuperscript{8}The reason is that a Flamengo supporter threw a plastic cup in the pitch during a Flamengo’s match at Maracanã in the previous year.
\end{footnotesize}
Table 1 provides the summary statistics of the variables included in the analysis. Noticeably, on average the home team scores almost one goal more than the visitor and is awarded almost twice the number of penalty kicks of its opponent; conversely, visitors receive almost twice the number of red cards of home teams; 3 minutes of injury time is added, on average, at the end of a match. These numbers, specially the ones for penalty kicks and red cards, suggest some sort of favoritism or home team bias and are further explored in the next subsections.

[Table 1 here]

3.1 Injury Time

According to the Law 7, regarding allowance for time lost:

“All allowance is made in either period for all time lost through:

• substitutions
• assessment of injury to players
• removal of injured players from the field of play for treatment
• wasting time
• any other cause

The allowance for time lost is at the discretion of the referee.”

Home team bias would imply the referee adding excess injury time in the case the home team is losing and doing the opposite in the case it’s winning, as in this way there would be more time for tying the game and less for the visitor to recoup, respectively. If the score margin is too large, any reasonable amount of injury time would not be enough

---

9In Stata 9.2, the command used is Statistics / Summaries, tables, & tests / Summary statistics / Summary statistics, selecting then the appropriate variables in the data set.


11Bold emphasis added.
for a significative change in the final result of a match, and presumably in this case the time added indeed corresponds to the time lost due to stoppages during the game\textsuperscript{12}.

Following the above discussion, the hypothesis to be tested in this subsection is:

**Hypothesis 1**: Injury time in the second half depends on the score margin.

In particular, favoritism would imply the extra time to be systematically larger when the home team is behind by one goal and systematically shorter when the home team is ahead by one goal.

The aforementioned papers by Garicano et al. and Sutter and Kocher find evidence of favoritism, as expressed in Hypothesis 1, in both the spanish and german soccer leagues, respectively. In the spanish league the results indicate that, on average, the referees act lengthening close games by almost two minutes, for the case where the home team is behind; in the german league the result is not so pronounced, but it’s still confirmed: referees add around half a minute more injury time when the home team is losing by one goal.

Figure 1 plots for the brazilian league the average amount of injury time for each possible score margin. As we can see, it’s still valid that on average the referees add excess injury time in the case the home team is losing by one goal. The difference in the amount of injury time when the home team is behind and when it’s ahead by one goal doesn’t seem to be visually significant, though.

![Figure 1 here](image_url)

Tables 2 and 3 present the estimation results for the panel data model with fixed effects and random effects, respectively, where the dependent variable is the amount of injury time in the 2nd half of the match\textsuperscript{13}. The models are estimated controlling for (1) the score difference, (2) number of substitutions, yellow and red cards and penalties,

\textsuperscript{12}The case of a match being tied at the moment the referee decides the amount of injury time is ambiguous: excess injury time can be interpreted as the referee giving more chances to the home team to score and hence win the game, while a shorter injury time could be interpreted as the referee decreasing the chances of the visitor to score, which would make the home team to be defeated. Hence, both cases could be interpreted as favoritism and therefore such situations won’t be included in the discussion of injury time.

\textsuperscript{13}In Stata 9.2, the command used is Statistics / Panel Data / Count Outcomes / Poisson regression, selecting then the appropriate variables according to the model specification being used.
(3) attendance and (4) rank difference. Regarding the controls, (1) corresponds to the hypothesis that we’re testing, (2) represents all the events that are true stoppages during the 2nd half of the game, (3) controls for a potential effect of the crowd in the referee behavior and (4) controls for the identities of the teams involved in the match.

[Table 2 here]

[Table 3 here]

As we can see, the score difference seems to indeed matter for the amount of injury time in the second half of the game. In both the FE and the RE models, the coefficient of the variable score difference is negative and significant at the 5% level, indicating that the higher (lower) the margin by which the home team is winning the lower (higher) is the amount of extra-time added in the end of the 2nd half. This reinforces the results obtained in Garicano et al. and Sutter and Kocher.

A word of qualification is necessary: (i) given that the referee is supposed to point out the amount of injury time before the 45th minute of the game, the score margin we consider for the analysis is the one prevailing at the 44th minute, which seems not to be the case in the other papers\(^{14}\) (apparently not consequential as in very few matches a goal was scored in the last minute AND the score margin turned to +1 or −1); (ii) in the so called derby matches, it’s difficult to point a home team, and this effect is not being controlled for in either the present analysis or in the other papers.

Even though injury time is a variable that can provide some evidence of home team bias, the very fact that only in few games a goal is scored after the regular time makes it not a good choice in the case a referee is indeed willing to favor the home team. For instance, in 55 out of 380 games in the data set a goal was scored after the referee indicated the amount of injury time and, among these, in only three the game came to a draw after the home team being behind by one goal. In the next subsections we turn to other variables that might be more fundamental in determining the final score of soccer matches and hence picked by referees to favor home teams.

\(^{14}\)For example, 4 minutes of injury time when the home team is ahead by one goal at the 45th minute of the second half doesn’t seem to qualify for home team bias. However, it might be the case that the goal that put ahead the home team was scored during the 44th minute, at a time when the referee had already indicated the amount of injury time: 4 minutes of injury time was may be in an attempt to give the home team more time to win the game, constituting home team bias.
3.2 Penalty Kicks

A penalty kick is awarded whenever a foul is committed by a player inside his own penalty area. In this case the team awarded the penalty kick is granted a direct free kick from the penalty mark, which is made 12 yds from the midpoint between the goalposts and equidistant to them. The distance between the posts is 8 yds and the distance from the lower edge of the crossbar to the ground is 8 ft. In other words, once a team is awarded a penalty kick it’s not difficult to score a goal\(^\text{15}\).

In the brazilian league played in 2008, in 71 out of 380 matches (18.7\%) the home team was awarded with one or more penalty kicks. Among those, in 52 games (73.2\%) the home team won the game and it was defeated in only 9 (12.7\%). These initial numbers suggest that a possible way to favor home teams would be by awarding them a penalty kick. Following this discussion, the hypothesis to be tested is:

**Hypothesis 2**: Visitor teams are awarded significantly less penalty kicks than home teams.

In Table 4, a paired t-test comparing the average number of penalty kicks awarded to home teams and visitors is performed\(^\text{16}\). As we see, on average home teams are awarded almost twice the number of penalty kicks of visitors. The test indicates that this difference is significant at the 1\% level.

Tables 5 and 6 present the estimation results for the panel data model with fixed effects and random effects, respectively, where the dependent variable is penalty kicks to the visitor team. The models are estimated controlling for (1) red cards to the visitor team, (2) penalties given to the home team, (3) attendance and (4) rank difference. Regarding the controls, (1) captures the idea that a referee might whistle a penalty kick to the visitor to compensate for a red card, (2) represents the hypothesis we’re trying to test (3) controls for a potential effect of the crowd in the referee behavior and (4) controls for the identities of the teams involved in the match.

[Table 4 here]

\(^{15}\)For instance, in the 1417 penalty kicks analysed by Palacios-Huerta (2003) from 1995 to 2000, mainly from the Italian, Spanish and English leagues, the scoring rate is 80.1\%.

\(^{16}\)In Stata 9.2, the command used is Statistics / Summaries, tables, & tests / Classical tests of hypotheses / Mean comparison test, paired data, selecting then the appropriate variables in the data set.
The results suggest that Hypothesis 2 is valid for the Brazilian soccer league in 2008. In the FE model, the coefficient of the variable penalties home is positive and significant at the 5% level, whereas for the RE model it’s significant at the 10% level. This indicates that a referee tends to give a penalty kick to the visitor team whenever a penalty kick is given to the home team, but for each penalty kick given to the home team the referee gives less than one penalty kick to the visitor, which gives an advantage to the home team.

In the next subsection we turn attention to yet another variable at the discretion of the referee that might be influential in the final result of a match: red cards.

### 3.3 Red Cards

Another factor that might influence the final result of a match is the distribution of red cards during the game. Upon receiving a red card a player is sent off and hence cannot participate in the game anymore. According to the Law 12:

“A player, substitute or substituted player is sent off if he commits any of the following seven offences:

- serious foul play
- violent conduct
- spitting at an opponent or any other person
- denying the opposing team a goal or an obvious goal-scoring opportunity by deliberately handling the ball (this does not apply to a goalkeeper within his own penalty area)
- denying an obvious goal-scoring opportunity to an opponent moving towards the player’s goal by an offence punishable by a free kick or a penalty kick
- using offensive, insulting or abusive language and/or gestures
- receiving a second caution in the same match.”
Even following the above guidelines, the referees still have discretion to act upon his own interpretation of situations like a serious foul play or a violent conduct. For instance, they cannot count on any sort of electronic device to better assess how serious a foul play was, and they might argue that from their position in the pitch the foul looked like a serious one even though it was not, providing an excuse for an unwarranted red card.

A red card might severely impair the tactical scheme of a team. In general, all the eleven players of each squad have a well defined position in the pitch and those are defined in accordance with players’ skills. In this sense, a red card implies not only one less player to face the opponent but also the need to put players in positions that they might not be well suited for.

Ridder, Cramer, and Hopstaken (1994), analyzing data on 140 red card games in the dutch professional soccer league from 1989 to 1992, show that a red card increases substantially the probability of victory of the team who was benefited, specially if the red card is given in the beginning of the match.

In the brazilian league played in 2008, in 86 out of 380 matches (22.6%) the visitor received one or more red cards. Among those, in 56 games (65.1%) the home team won the game and it was defeated in only 12 (14.0%). These initial numbers suggest that a possible way to favor the home team would be by giving red cards to the visitor.

The above discussion suggests the following hypothesis to be tested:

**Hypothesis 3**: The number of red cards to home teams is systematically smaller than the number of red cards to visitors.

In Table 7, a paired t-test comparing the average number of red cards given to home teams and visitors is performed. As we see, on average home teams receive almost half the number of red cards of visitors. The test indicates that this difference is significant at the 1%.

Tables 5 and 6 present the estimation results for the panel data model with fixed effects and random effects, respectively, where the dependent variable is red cards given to the home team. The models are estimated controlling for (1) red cards to the visitor team, (2) penalties given to the home team, (3) attendance and (4) rank difference. Regarding the controls, (1) represents the hypothesis we’re trying to test, (2) captures the idea that a referee might give a red card to the home team to compensate for a potential mistake in a
penalty kick given, (3) controls for a potential effect of the crowd in the referee behavior and (4) controls for the identities of the teams involved in the match.

[Table 7 here]

[Table 8 here]

[Table 9 here]

The results suggest that Hypothesis 3 is valid for the brazilian soccer league in 2008. In the FE model, the coefficient of the variable red cards visitor is less than one and significant at the 1% level, whereas for the RE model it’s also significant at the 1% level, but close to one. At least in the FE model, this indicates that a referee tends to give a red card to the home team whenever a red card is given to the visitor team, but for each red card given to the visitor team the referee gives less than one red card to the home team, which gives an advantage to the home team, again.

In the next section we analyze a potential channel through which home team bias might take place: social pressure.

4 Social Pressure: the Influence of the Crowd

Anyone who ever attended a sport’s match with a stadium full of supporters knows the pressure the crowd can exert on the home team, the visitors or the referee. Usually in soccer the supporters are very passionate about the game and the pressure can be even higher than what one would expect it to be\textsuperscript{17}.

\textsuperscript{17}In an interview to the soccer website Lancenet.com.br, the former Palmeiras and brazilian midfielder Zinho remembers going to take his kick during the decision by penalty kicks of the quarter-finals of Libertadores Cup 2002, in a match against Corinthians, one of the most important derbys in Brazil: “I still don’t know how I managed to walk from the halfway line to the penalty mark to kick that ball (...). It was the greatest pressure ever in my life.” (in the decision by penalty kicks, the players from both teams stand in the halfway line inside the centre circle, and from there they go alternately to the penalty mark to take their respective kicks). As a remark, Zinho was one of the kickers in the decision by penalty kicks in the World Cup of 1994, Brazil vs Italy, the most important tournament in the career of a professional soccer player but where presumably the pressure exerted by the crowd is much smaller.
The fact that the social environment can influence individual behavior is acknowledged in economics by the literature on endogenous preference formation. In sports, the literature shows evidence of this phenomenon happening in different ways, like referees trying to please the supporters and the crowd noise serving as a (biased) cue for the referee’s decision, impairing the use of heuristics in their decision making.

The hypothesis to be tested is:

**Hypothesis 4**: The number of spectators in the game has a significant effect on the behavior of the referee.

Some comments are in order about Hypothesis 4. First, it’s being taken for granted that the size of the crowd is the factor capturing social pressure or, equivalently, the number of spectators has a one-to-one relationship with the factor of the social environment that in fact influences the referee’s behavior. Second, by referee’s behavior it’s understood the distribution of penalty kicks and red cards between home teams and visitor.

As we can see in the tables from the previous sections, in none of the models specified the variable attendance turns out to be significant. This might be a result from an specification error or a true lack of any influence whatsoever coming from the supporters and impacting the likelihood of the referee in awarding more extra-time, penalty kicks or red cards. Apparently, if there’s any social pressure on the referee that might potentially be influential to his behavior, this is not being captured by the number of supporters in the stadium.

## 5 Concluding Remarks

Economic theory offers many tools to analyze outcomes that are of interest to fields other than economics itself. Sports is one such a field and the goal of the present paper was to analyze the results observed in a particular sport, soccer, through the lenses of economics.

Soccer is a suggestive case where one potentially observes preferences subject to endogenous formation or, in simple terms, behavior being affected by the social environ-

---

18See Akerlof (1980); Becker and Murphy (2005); Bernheim (1994).
19See Sutter and Kocher (2004) and the references therein.
20For instance, if noise is what influences referee’s behavior and noise is driven by passion, in the case the largest crowd is not the most passionate then Hypothesys 4 would not necessarily be valid.
The literature offers strong evidence of referees in soccer favoring home teams by all sorts of means, in particular using the injury time for shortening (lengthening) matches where the home team is ahead (behind) the visitor by a close margin and awarding more penalty kicks to home teams than to visitors. Moreover, the crowd seems to be the channel through which favoritism takes place: the larger the crowd the larger the bias observed in the referee’s behavior.

The present paper tested the hypothesis of referees’ home team bias in the Brazilian soccer league of 2008. The variables focus of the analysis were amount of injury time, penalty kicks and red cards. Similarly to the results obtained in other leagues, the Brazilian one also shows evidence of favoritism. In particular, the amount of extra-time given in the end of the 2nd half of the game is negatively related to the score margin of the home team, being significant at the 5% level. Moreover, the number of penalty kicks awarded to home teams is significantly larger than that to visitors, the converse happening with red cards.

As a remark, more than trying to estimate the size of the home team bias or the influence of the crowd on the behavior of referees, the purpose of the paper was to verify the mere presence of those effects in the data collected. The specifications chosen need to be more carefully elaborated in order to assent the results in more solid grounds.

Also, a more detailed data on the Brazilian soccer league of 2008 is necessary before more incisive statements can be made. For instance, in the analysis of penalty kicks and red cards, one of the arguments is that, being disfavoured by a red card or a penalty kick to the visitor, the supporters of the home team start to claim unwarranted penalty kicks or red cards to the opponent, putting more pressure on the referee and potentially influencing his decisions. However, the data set consists of aggregate variables and does not indicate the timing of the events, which would be more in accordance with the type of argument being used.

References


Table 1: Summary statistics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score Difference</td>
<td>371</td>
<td>.7412399</td>
<td>1.654771</td>
<td>-6</td>
<td>5</td>
</tr>
<tr>
<td>Score Home</td>
<td>371</td>
<td>1.749326</td>
<td>1.343339</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Score Visitor</td>
<td>371</td>
<td>.9892183</td>
<td>1.047464</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Penalties Home</td>
<td>371</td>
<td>.2156334</td>
<td>.478594</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Penalties Visitor</td>
<td>371</td>
<td>.1239892</td>
<td>.3686956</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Extra Time 1st half</td>
<td>371</td>
<td>1.469003</td>
<td>.8485178</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Extra Time 2nd half</td>
<td>371</td>
<td>3.051213</td>
<td>1.031958</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Red Cards Home</td>
<td>371</td>
<td>.1509434</td>
<td>.4335497</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Red Cards Visitor</td>
<td>371</td>
<td>.2587601</td>
<td>.4963625</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Attendance</td>
<td>371</td>
<td>17107.2</td>
<td>12471.02</td>
<td>500</td>
<td>77387</td>
</tr>
<tr>
<td>Att./Stadium Capac.</td>
<td>371</td>
<td>.3913348</td>
<td>.2504902</td>
<td>.0166667</td>
<td>.9792929</td>
</tr>
</tbody>
</table>

Figure 1: Injury Time Per Score Margin
Table 2: Extra-time 2nd half (FE regression).

```
.xtpoisson extra_time_2ndhalf score_diff_befor subst_2ndhalf yellowcards_2ndhalf 
   red_cards_2ndhalf tot_pen att_ratio rank_diff, fe 

note: 7 groups (7 obs) dropped because of only one obs per group

Conditional fixed-effects Poisson regression  Number of obs  =  364
Group variable (i): code_ref  Number of groups  =  38

Obs per group: min =  2
avg =  9.6
max =  22

Wald chi2(7)  =  11.57
Log likelihood  = -494.98454  Prob > chi2  =  0.1155

| Extra-time 2nd Half | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|---------------------|--------|-----------|-------|------|----------------------|
| Score Difference    | -.0459587 | .0195099  | -2.36 | 0.018 | -.0841974 - .00772 |
| Substitution 2nd Half | .0027844 | .0333033  | 0.08  | 0.933 | -.0624888 .0680576 |
| Yellow cards 2nd Half | .0118671 | .0201341  | 0.59  | 0.556 | -.0275959 .0513293 |
| Red Cards 2nd Half  | .0892777 | .0507096  | 1.76  | 0.078 | -.0101113 .1886667 |
| Penalties           | -.0569116 | .0535372  | -1.06 | 0.288 | -.1618427 .0480194 |
| Attendance          | .0626815 | .1281805  | 0.49  | 0.625 | -.1885477 .3139108 |
| Rank Difference     | -.0007836 | .0039329  | -0.20 | 0.842 | -.008492 .0069247 |
```
Table 3: Extra-time 2nd half (RE regression).

```
.xtpoisson extra_time_2ndhalf score_diff_befor subst_2ndhalf yellowcards_2ndhalf red_cards_2ndhalf tot_pen att_ratio rank_diff, re
```

Random-effects Poisson regression

- Number of obs = 371
- Number of groups = 45

Random effects u_i ~ Gamma
- Obs per group: min = 1
- avg = 8.2
- max = 22

Wald chi2(7) = 12.43
Log likelihood = -613.99875
Prob > chi2 = 0.0872

| Extra-time 2nd Half | Coef. | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|---------------------|-------|-----------|------|-----|----------------------|
| Score Difference    | -.0378563 | .0182393 | -2.08 | 0.038 | -.073604 7 - .0021079 |
| Substitution 2nd Half | .0040972 | .0310041 | 0.13  | 0.895 | -.0566697 .697 .0648641 |
| Yellow cards 2nd Half | .0254026 | .0173158 | 1.47  | 0.142 | -.00853 58 .059341 |
| Red Cards 2nd Half  | .0622032 | .0468739 | 1.33  | 0.184 | -.029668 4 .1540744 |
| Penalties           | -.0727215 | .0500455 | -1.45 | 0.146 | -.1708089 .0253658 |
| Attendance          | .0855881 | .1205823 | 0.71  | 0.478 | -.1507488 .3219251 |
| Rank Difference     | -.0009876 | .0036546 | -0.27 | 0.787 | -.0081505 .0061754 |
| Constant            | 1.01281 | .1644451 | 6.16  | 0.000 | .6905037 1.335116 |

/Inalpha  | -16.25043 | 456.6767 | -911.3203 | 878.8194 |

alpha  | 8.76e-08 | .00004 | 0 | . |

Likelihood-ratio test of alpha=0: chibar2(01) = 0.00 Prob>=chibar2 = 1.000
Table 4: Paired t-test: Penalty Kicks.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penalties Home</td>
<td>371</td>
<td>.2156334</td>
<td>.0248474</td>
<td>.478594</td>
<td>.1667737  .2644932</td>
</tr>
<tr>
<td>Penalties Visitor</td>
<td>371</td>
<td>.1239892</td>
<td>.0191417</td>
<td>.3686956</td>
<td>.086349  .1616294</td>
</tr>
<tr>
<td>Difference</td>
<td>371</td>
<td>.0916442</td>
<td>.0299199</td>
<td>.5762978</td>
<td>.0328098 .1504786</td>
</tr>
</tbody>
</table>

Ho: mean(Difference)=0 Ha: mean(Difference)>0 t=3.0630 df=370 Pr(T>t)=0.0012

Table 5: Penalty kicks (FE regression).

```
.xtpoisson penaltis_visitor red_cards_visitor penalties_home attendance rank_diff, fe
test: 7 groups (7 obs) dropped because of only one obs per group
note: 15 groups (95 obs) dropped due to all zero outcomes

Conditional fixed-effects Poisson regression
Number of obs = 269
Group variable (i): code_ref
Number of groups = 23
Obs per group: min = 3
             avg = 11.7
             max = 22
Wald chi2(4) = 9.16
Log likelihood = -92.751224
Prob > chi2 = 0.0571

Penalties visitor | Coef.  Std. Err.  z  P>|z|  [95% Conf. Interval]
------------------|--------|-------------|--------|--------|---------------------|
Red cards visitor | .2267725 | .2800493  0.81  0.418  -.3221141  .775659 |
Penalties Home    | .7011555 | .2837714  2.47  0.013   .1449738  1.257337 |
Attendance        | 7.58e-06 | .0000123  0.62  0.536  -.0000164  .0000316 |
Rank Difference   | .0288212 | .0192387  1.50  0.134  -.0088859  .0665284 |
```

19
Table 6: Penalty kicks (RE regression).

```
.xtpoisson penaltis_visitor red_cards_visitor penalties_home attendance rank_diff, re

Fitting Poisson model:

Random-effects Poisson regression Number of obs = 371
Group variable (i): code_ref Number of groups = 45

Random effects u_i ~ Gamma Obs per group: min = 1
                 avg = 8.2
                 max = 22

Wald chi2(4) = 6.09
Log likelihood = -142.72788 Prob > chi2 = 0.1927

-----------------------------------------------------------------------------------
Penalties visitor | Coef. Std. Err. z P>|z|   [95% Conf. Interval]
-----------------+----------------------------------------------------------
Red cards visitor |  .2204547  .2715397  0.81   0.417   -.3117533   .7526627
Penalties home    |  .4193317  .2366923  1.77   0.076   -.0445762   .8832395
Attendance        |  5.78e-06  .0000118  0.49   0.623  -.0000173   .0000237
Rank difference   |  .0253357  .0179774  1.41   0.159  -.0098993   .0605708
Constant          |  -2.387994 .2865267 -8.33  0.000  -2.949576  -1.826412
-----------------------------------------------------------------------------------
/lnalpha | -15.30859  899.5507  -8.33  0.000  -2.949576  -1.826412
-----------------------------------------------------------------------------------
alpha |  2.25e-07  .0000202  0  .
-----------------------------------------------------------------------------------
Likelihood-ratio test of alpha=0: chibar2(01) = 0.00 Prob>=chibar2 = 1.000
```

20
Table 7: Paired t-test: Red Cards.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. Err.</th>
<th>Std. Dev.</th>
<th>95% Conf. Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Cards Home</td>
<td>371</td>
<td>.15094</td>
<td>.0225088</td>
<td>.4335497</td>
<td>.1066822 - .1952046</td>
</tr>
<tr>
<td>Red Cards Visitor</td>
<td>371</td>
<td>.25876</td>
<td>.0257699</td>
<td>.4963625</td>
<td>.2080864 - .3094339</td>
</tr>
<tr>
<td>Difference</td>
<td>371</td>
<td>-.10782</td>
<td>.0292804</td>
<td>.5639804</td>
<td>-.1653936 - .0502398</td>
</tr>
</tbody>
</table>

Ho: mean(Difference)=0 Ha: mean(Difference)<0 t=-3.6822 df=370 Pr(T<t)=0.0001

Table 8: Red cards (FE regression).

```
xtpoisson red_cards_home red_cards_visitor penalties_home attendance rank_diff, fe
```

Note: 7 groups (7 obs) dropped because of only one obs per group

Note: 17 groups (93 obs) dropped due to all zero outcomes

Conditional fixed-effects Poisson regression

|                   | Coef. | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|-------------------|-------|-----------|------|------|----------------------|
| Red cards home    | .88662| .223748   | 3.96 | 0.000| .4480795 1.325156   |
| Red cards visitor | -.23474| .3449828| -0.68| 0.496| -.910891 .4414169  |
| Penalties home    | -2.20e-06| .000013| -0.17| 0.866| -.0000277 .0000233 |
| Attendance        | -2.05 | 0.041    | -2.05| 0.041| -.0702759 -.0015527 |
```

21
Table 9: Red cards (RE regression).

```
.xtpoisson red_cards_home red_cards_visitor penalties_home attendance rank_diff, re

Random-effects Poisson regression
Group variable (i): code_ref

Number of obs = 371
Number of groups = 45

Random effects u_i ~ Gamma
Obs per group: min = 1
avg = 8.2
max = 22

Wald chi2(4) = 27.99
Prob > chi2 = 0.0000

Log likelihood = -152.1991

------------------------------------------------------------------
Red cards home | Coef. Std. Err. z P>|z| [95% Conf. Interval]
----------------+--------------------------------------------------
Red cards visitor | 1.020338 .2031851 5.02 0.000 .6221025 1.418573
Penalties home | -.3750043 .3305603 -1.13 0.257 -1.022891 .2728821
Attendance | -3.31e-06 .0000119 -0.28 0.781 -.0000266 .00002
Rank difference | -.0294093 .0171412 -1.72 0.086 -.0630054 .004190
Constant | -2.242156 .2918148 -7.68 0.000 -2.814102 -1.670209
------------------------------------------------------------------
/lnalpha | -1.210806 .7868631 -2.753029 .331474
------------------------------------------------------------------
alpha | .2979571 .2344514 .0637345 1.392941
------------------------------------------------------------------
Likelihood-ratio test of alpha=0: chibar2(01) = 3.55 Prob>chibar2 = 0.030
```

Red cards home | Coef. Std. Err. z P>|z| [95% Conf. Interval]
----------------+--------------------------------------------------