Universität St.Gallen

Rot-Jaune-Verde. Language and Favoritism: Evidence from Swiss Soccer

Richard Faltings, Alex Krumer, Michael Lechner

October 2019 Discussion Paper no. 2019-15

School of Economics and Political Science, Department of Economics

University of St.Gallen

Editor:	Vanessa Pischulti University of St.Gallen School of Economics and Political Science Department of Economics Müller-Friedberg-Strasse 6/8 CH-9000 St.Gallen Phone +41 71 224 23 07 Email seps@unisg.ch
Publisher:	School of Economics and Political Science Department of Economics University of St.Gallen Müller-Friedberg-Strasse 6/8 CH-9000 St.Gallen Phone +41 71 224 23 07
Electronic Publication:	http://www.seps.unisg.ch

Rot-Jaune-Verde. Language and Favoritism: Evidence from Swiss Soccer [1]

Richard Faltings, Alex Krumer, Michael Lechner[2]

Author's address:

Richard Faltings
Department of Economics
The University of Texas at Austin
2225 Speedway
US-Austin, Texas 78712
Email richard.faltings@utexas.edu

Author's address:

Alex Krumer
Faculty of Business Administration and Social Sciences
Molde University College
Britvegen 2
NO-Molde, 6402
Email alex.krumer@himolde.no

Author's address:

Michael Lechner
Swiss Institute for Empirical Economic Research (SEW)
University of St.Gallen
Varnbüelstrasse 14
CH-9000 St.Gallen
Email michael.lechner@unisg.ch
Website michael-lechner.eu

[1] We would like to thank Daniel Goller, Guy Elaad, Kjetil Haugen, Lars Hvattum, Michael Knauss, and Ignacio Palacios Huerta for helpful comments. The usual disclaimer applies.

[2] Michael Lechner is also affiliated with CEPR, London, CESIfo, Munich, IAB, Nuremberg, IZA, Bonn, and RWI, Essen.

Abstract

Switzerland is a multi-lingual developed country that provides an attractive stage to test in-group favoritism that is driven by linguistic differences. To that end, we utilize data from soccer games in the top two Swiss divisions between the seasons 2005/06 and 2017/18. In these games, the referee was from the same linguistic area with one team, whereas the other team was from a different linguistic area. Using very rich data on teams' and games' characteristics, our causal forest-based estimator reveals that referees assign significantly more penalties in the form of yellow and red cards to teams from a different linguistic area. This form of in-group favoritism is large enough so that it is likely to affect the outcome of the game. As evidence, we find that the difference in points in favor of the home team increases significantly when a referee is from the same linguistic area.

Keywords

Favoritism, discrimination, soccer, language

JEL Classification

D00, J71, L00, Z13, Z20

1 Introduction

"The relation of comradeship and peace in the we-group and that of hostility and war towards others-groups are correlative to each other ... The closer the neighbors, and the stronger they are, the intenser is the warfare ... Loyalty to the group, sacrifice for it, hatred and contempt for outsiders, brotherhood within, warlikeness without – all grow together, common products of the same situation. These relations and sentiments constitute a social philosophy."[1]

-William Graham Sumner, American social scientist

By our nature, humans are species who join together in groups (Sumner, 1904; Yuki, 2003). Therefore, behavior that can serve the interest of one's group is likely to be an inherent feature of humans. For example, Efferson, Lalive and Fehr (2008) showed that even different signs on shirts were enough to divide people into groups and create in-group favoritism, according to which members of one's group favor their in-group members over out-group members. This evolutionary pattern violates the liberal idea of equality, where people of equal ability or merit should not be treated differently. The principle of equality is made explicit in the laws that govern most modern democracies.

In reality, many different features can create group identities, such as religion, nationality, skin color, etc. In this paper, we study another basic feature that divides people into groups; namely, language that may serve as a marker of social group identity, in the same way as ethnicity or nationality, all of which may provoke in-group favoritism and tension between different groups. As evidence, throughout history, linguistic differences have created many conflicts all around the world.

1 From Sumner (1906, pp. 12–13).

This paper investigates possible linguistic bias in Switzerland, a country that has four official languages: German, French, Italian and Romansh. The federal government's working languages comprise the first three (Swiss Federal Government (1999), hereinafter (SR 101), Art. 70), all of which are included in the title of this paper. The Swiss Constitution explicitly prohibits any linguistic discrimination (SR 101, Art. 8). Even in the absence of further legislation, this binds the country's judicial system as well as the state in its role as an employer against practicing linguistic discrimination. Nevertheless, over the years, Switzerland has experienced many tensions over language issues. A recent example concerns the usefulness of the English language versus French or German in schools. For instance, French cantons objected to English becoming the first non-native language taught in German-speaking cantons.[2] Because it can be difficult to systematically identify instances of discrimination in legal and employment proceedings, the existence and extent of linguistic discrimination in Switzerland is an open question.

To answer this question, we use a real competition between professionals to investigate possible in-group favoritism that may be driven by linguistic differences between groups. For that purpose, we use data from Swiss soccer. The yellow and red card penalty decisions of referees are evaluated to determine the existence of systematic biases that favor teams belonging to the same linguistic area (region) as the referee or penalize teams that do not belong to the same area. For example, in the 2018 FIFA World Cup, the Serbian team complained against a German referee who, according to the Serbians, was biased in favor of the Swiss team. Moreover, Football Association of Serbia claimed that there was a linguistic bias by issuing the following statement: "We are not clear how the German referee could

[2] See https://www.swissinfo.ch/eng/multilingualism_language-vote-would-be--dangerous--for-switzerland/41082384 Last accessed on 16/07/2019.

have been appointed for the match between Switzerland and Serbia, when it is well known that one of Swiss confederation cantons is a German canton."[3]

In this study, we used the Modified Causal Forest (MCF) estimator recently proposed by Lechner (2018). This estimator exploits recent advances in causal machine learning literature and allows nonparametric estimation of causal effects. We used very rich data on the games from the top two Swiss soccer divisions from the 2005/06 season until the first half of the 2017/18 season, in which a referee shared the same linguistic area with one and only one team. Our MCF estimator reveals that having a referee from the same linguistic area as the home team increases the gap between the sides, in terms of the number of red and yellow cards in favor of a home team, by about 0.27. Moreover, having a referee from the same linguistic region as a home team also has a significant effect on the outcome of the game, because such teams gain about 0.23 points more than home teams that do not share the same linguistic area as the referee. Such a difference may have serious financial consequences for teams in a tight league (such as relegation to a lower division or non-participation in the European cups). Overall, the results provide empirical evidence that linguistic discrimination is a concern, with immediate implications for the design of Swiss soccer tournaments as well as for general anti-discriminatory policies.

The remainder of the paper is organized as follows. Section 2 briefly reviews the relevant literature before Section 3 describes the institutional settings. The data and some descriptive results are presented in Section 4. Section 5 presents the empirical strategy, the results are contained in Section 6, and Section 7 offers concluding remarks. Appendix A contains further descriptive statistics. Finally, Appendices B and C present some effect heterogeneity tests.

[3] See https://eu.usatoday.com/story/sports/soccer/2018/06/24/serbia-fa-complains-of-biased-referee-in-switzerland-match/36335241/. Last accessed on 16/07/2019.

2 Literature review

This article adds to the long literature on discriminatory biases in economics, dating back to Becker (1957), who introduced the notion of taste-based discrimination to explain differences in employment patterns across blacks and whites in the United States.[4] In the sports related context, Szymanski (2000) showed strong evidence of discrimination against black players in English soccer.

The other main class of economic models in this area involves so-called statistical discrimination, as introduced by Phelps (1972) and Arrow (1972). In these models, race (or any other discriminatory factor) is taken as a signal of a job market candidate's performance by an employer confronted with information asymmetries. In the context of sports, we are unlikely to find any kind of statistical discrimination in the sense presented here, as referees do not face information asymmetries when making decisions.

Another type of discrimination that was described in detail by Bertrand, Chugh, and Mullainathan (2005) is the so-called implicit discrimination that may arise due to the unconscious association of a social group with one or more negative attributes.[5] According to Price and Wolfers (2010), such an unconscious association may play a role in the type of split-second high-pressure situations that appear in NBA basketball games. The authors found that players have up to 4 percent fewer fouls called against them and score up to 2.5 percent more points when their race matches that of the refereeing crew. Similarly, Gallo, Grund and Reade (2013) found that non-white players from countries with a GDP per capita below 10,000 USD are more likely to be awarded a yellow card in English soccer. However,

[4] The name derives from the fact that the model simply assumes that agents have certain preferences (that is, tastes) for or against certain groups of individuals, based on race, for example.

[5] For additional details, see an excellent review of Greenwald and Banaji (1995).

a definition that relates a soccer player's skin color to his country's GDP per capita in a context of referee's decision making is not entirely intuitive.

Additionally, previous research has shown nationalistic in-group favoritism in various settings. For example, Zitzewitz (2006) found that judges assign significantly higher grades to athletes from their nationality in international ski jumping and figure skating competitions. Sandberg (2017) found similar result in dressage competitions. Finally, Pope and Pope (2015) showed that soccer referees favor their compatriot players by assigning them more beneficial foul calls in the UEFA Champions League games.

Also, several studies have related to linguistic differences in economic environments. For example, Eugster, Lalive, Steinhauer, and Zweimüller (2017) showed how cultural attitudes towards work can cause consistent differences in unemployment duration between different linguistic regions of Switzerland. However, that paper does not claim any in-group favoritism. On the other hand, Angerer, Glätzle-Rützler, Lergetporer, and Sutter (2016), who conducted an experiment in a Northern Italian bilingual city, did show in-group favoritism among children based on linguistic differences.[6] Two studies have sought to link in-group favoritism on the linguistic/ethnic basis in the NHL ice hockey league. In the first, Mongeon and Longley (2015) showed that two referees from French-speaking parts of Canada call penalties significantly faster against players from English-speaking parts of Canada. However, their findings have three possible caveats. Although the authors did not present the results of the appropriate tests, the results do not seem to differ from the case when there was one Canadian English and one American referee, suggesting that language differences do not drive the effect. In addition, only 3.5 percent of the games were refereed by two French referees. Finally, given the same dataset, in their follow-up study (Mongeon and Longley, 2017), the authors showed that two referees from the English part of Canada assign

---

[6] See also Fidrmuc, Ginsburgh and Weber (2009), who suggested reducing the number of core languages in the EU.

significantly more penalties to a team with more English Canadians on the ice, contradicting the possibility of in-group favoritism. Therefore, the external validity of the previous studies on in-group favoritism that is driven by linguistic differences remains an open question and we have sought to fill that gap by using a different environment and methodology.

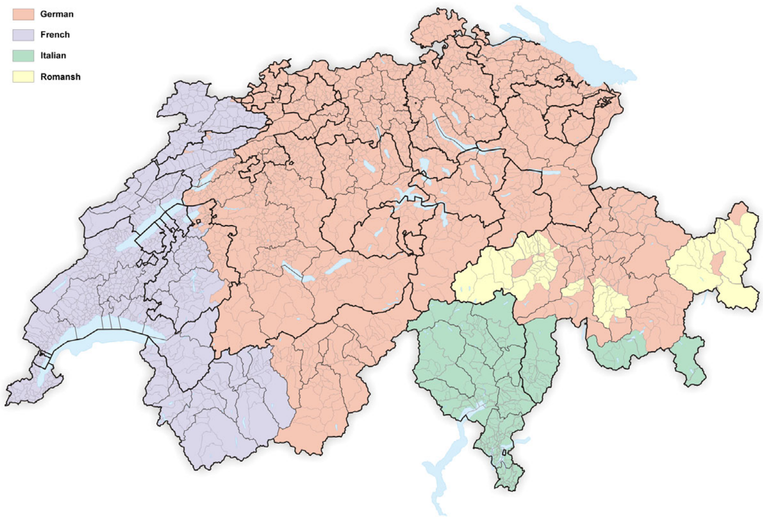
Finally, our study also relates to the literature on fairness in economic environments, whose importance was discussed by none other than Adam Smith (Ashraf, Camerer and Loewenstein, 2005). More recently, Gill and Stone (2010) theoretically analyzed the role of fairness in tournament settings. Finally, in his comprehensive review, Konow (2003) suggested that "differences owing to birth, luck and choice are all unfair and that only differences attributable to effort are fair" (p. 1207). Therefore, our study suggests that the decision making of Swiss referees is biased because it is driven by linguistic preferences and, as such, it violates the principles of fair play in soccer and, in some sense, the antidiscrimination clause of the Swiss Constitution.

3 Background

3.1 Languages

Multilingualism is a defining characteristic of Switzerland, reflecting its unique geographical positioning between France, Germany, Austria, Liechtenstein, and Italy. As mentioned above, the Swiss Constitution defines four national languages although the federal government's working languages only comprise German, French, and, to a lesser extent, Italian. The languages are geographically distributed to roughly match the country's immediate neighbors, with German to the north, French to the west, and Italian to the south-east (see Figure 3.1).

Figure 3.1: Map of languages in Switzerland.



Source: wikipedia.org

Table 3.1 shows the distribution of languages in Switzerland based on a survey conducted by the Swiss Federal Statistical Office (2019), where responders listed up to three of their main languages. German is the dominant language, with around two-thirds of the population declaring it as one of their main languages, with French second, and Italian third. Romansh is spoken by only very few people and comes in a distant fourth place.

Despite the skewed distribution of its languages, Switzerland puts tremendous effort into creating harmony between the different linguistic communities. The seven-member Federal Council, which serves as the head of the government, is elected to "ensure that the various geographical and language regions of the country are appropriately represented" (Art. 175 SR 101). In addition, children learn a second official language at school, such that 73 percent of residents in the French-speaking region and 92 percent of residents in the German-speaking region speak a second national language (Werlen, Baumgartner and Rosenberger, 2011 in Eugster et al., 2017). Despite the efforts to increase unity, there remain significant differences in culture between the linguistic regions. For example, the aforementioned study by Eugster et al. (2017) exploits the sharp change in attitudes toward work that occur along the language border between the German-speaking regions and the

Romance (French and Italian)-speaking ones to show how cultural differences may affect unemployment duration. In addition, the divides are particularly apparent in the voting behavior of the population, which often differ markedly across linguistic areas.[7]

Table 3.1: General data on cantons and population.

Language	Cantons	Population %
German	21	62.63
French	7	22.86
Italian	2	8.18
Romansh	1	0.53
English	N/A	5.41
Others	N/A	18.92

Notes: The Cantons column represents the number of cantons in which the language is prominent/official. Survey respondents listed up to three of their main languages.

3.2 Institutional settings of Swiss soccer

For the purposes of our study, it is important to have a clear understanding of the rules that govern the game as well as the tournaments in question. According to the Swiss Football Association's (2017) Rules of the Game (hereinafter "SFA.RG"), the top two leagues are both managed by the Swiss Football League, which should ensure some degree of homogeneity between them (Art. 9 SFA.RG). The in-game rules are defined by the "Laws of the Game" (hereinafter "IFAB.LG") set out by the International Football Association Board (2017) (Art. 15 SFA.RG). The most relevant provisions concern the referees' authority (Law 5 IFAB.LG) and the definition of punishable behavior (Law 12 IFAB.LG). Law 5.1 IFAB.LG states that there is a single referee, who is the main subject of our study.

[7] See https://www.swissinfo.ch/eng/society/german-vs-french_the--roesti-divide---a-barrier-that-binds-the-swiss/41193552. Last accessed on 16/07/2019.

He has full authority in the match and Law 5.3 IFAB.LG states that this referee operates in cooperation with or on the advice of other match officials. While this means that the referees should take the advice of assistant referees into account, they still have the final say for any in-game decision, which justifies this article's focus on the head referee.

In general, fouls can lead to a free kick or penalty being awarded to the opposing team, depending on the location of the foul (Laws 12.1, 12.2, 13, 14 IFAB.LG). Some of these fouls can also lead to a caution (yellow card) or a sending-off from the field (red card), depending on the severity of the foul (Law 12.3 IFAB.LG). In addition, a yellow card can be attributed to a variety of reasons, including certain vague conditions such as "dissent by word or action", "persistent offenses", and "unsporting behavior". The latter term encompasses a broad range of unacceptable behavior, including "show[ing] a lack of respect for the game". The offenses meriting a red card are generally less ambiguous, but notably include receiving a second yellow card within the same match. A sent-off player cannot be replaced (Law 3.6 IFAB.LG). In addition, the Swiss Football League automatically suspends players who are given a red card, until further review by a disciplinary judge, who may set a suspension of up to four games.[8] Altogether, these rules show that referees must make decisions amidst considerable ambiguity and that their decisions can have real impacts on the game being played as well as future games in the tournament.

The rules governing referee selection and payment are not as transparent. In particular, it is not clear how referees are selected for any given match. The Swiss Football League's Competition Rules only states that referees are designated by the Referee Commission (Art. 19). The same article states that referees are employed and paid by the Swiss Football Association. The salaries and payments for each match were generally low, with top salaries of 24,000 CHF per year and 1,250 CHF per match, such that referees in Swiss football are

---

[8] See http://www.sfl.ch/fr/sfl/droit-licences/droit/droit-disciplinaire/. Last accessed on 16/07/2019.

generally semi-professional. However, starting in 2018, a few of the most elite referees have had a drastic increase in their salaries.[9]

4 Data and descriptive results

4.1 Data base

The raw data covers all the games played in the two highest soccer leagues in Switzerland – the Swiss Super League and the Challenge League – in the period from the 2005/06 season (13th of July 2005) up to the first half of the 2017/18 season (17th of December 2017), which consists of 5,095 games. The first season was chosen, since prior to the 2005/06 season, there is no data on players' values from the popular website transfermarkt.com (see Sub-section 4.3 for more information). Additional sources include the sfl.ch and oddsportal.com websites. Certain information on stadium capacities and referee characteristics were collected from different open sources.

We define a team's language according to its location. This does not imply that the players themselves speak the team's language, but it does have implications for the team's social identity. In addition, the home team's language is likely to be spoken by the majority of the spectators. As shown in Eugster et al. (2017), language borders are sharply defined, with the proportion of Romance-language-speakers jumping from 20 percent in the German speaking regions to 85 percent within 5 km of the border.

To estimate the effect of having a referee from the same linguistic area, we only selected games where a referee was from the same linguistic region as only one of the teams, whereas the other team was from a different linguistic area. In total, we have 1,404 such

[9] See https://www.tdg.ch/sports/actu/arbitrage-premier-professionnalisme/story/20793516 . Last accessed on 16/07/2019.

games, 681 involving a referee from the same linguistic area as the away team and 723 where the referee was from the same linguistic area as the home team.

The languages of the referees are similarly defined according to location, which is again justified by the strong influence of geography on language. Birthplace is preferentially used to define the language; when it is not available, the current residence is used instead. More specifically, information on birthplace was used for eight out of the 54 referees in the sample, 43 were matched using residence, two were matched based on the name (and phonebook searches), and one on nationality (Austrian). The majority of the birthplace and residence information was collected from transfermarkt.com. For the missing cases, we performed individual web searches. Table 4.1 summarizes the distribution of languages over the teams and the referees. The distribution of teams is roughly in line with the figures for the general population, as presented in Table 3.1. However, we see a higher share of German-speaking referees than in the overall population.[10]

Table 4.1: Games refereed and number of referees, by language

Language	Teams	Teams (share)	Referees	Referees (share)	Language of the referee	Language of the referee (share)
German	23	60.5%	44	81.5%	1,205	85.8%
French	11	28.9%	9	16.7%	192	13.7%
Italian	4	10.5%	1	1.8%	7	0.5%

[10] Note that the data on German-speaking referees include nine Austrian referees, who refereed a total of 14 games.

4.2 Descriptive statistics

To estimate the possible effect of having a referee from the same linguistic area, we have a set of possible outcome variables on the level of a single game between a home and away teams. Table 4.2 shows that, in line with the home advantage phenomenon, the home teams obtain more points, on average, than the away teams, regardless of the linguistic area of the referee. However, we see that the average number of points of home teams in games where the referee is from the same linguistic region is higher than where the referee shares the same linguistic area with the away team.

Another dimension that may be of interest is the distribution of yellow and red cards between the teams. We can see that, in general, the difference between cards (defined as home minus away) is negative, suggesting that the away team receives more cards than the home team, which is in line with previous research.[11] We can also see that the gap between cards is wider when the referee shares the same linguistic region as the home team. This means that there are more cards assigned to the away team compared to the home team when a referee shares the same linguistic area as the home team. We also have data on distribution of yellow and red cards separately, as well as the data on cards for different positions of the players. All of these will serve as outcome variables.

4.3 Variables

To estimate the effect of having a referee from the same linguistic area, we coded a dummy variable that is one if a home team and the referee share the same linguistic area and zero otherwise (that is, if the away team and the referee share the same linguistic area). We

[11] For example, using laboratory settings, Nevill, Balmer and Williams (2002) determined that crowd noise had a significant effect on the probability of issuing a yellow card. Downward and Jones (2007) showed a positive relationship between the size of the crowd and the likelihood of getting a yellow card in the English FA Cup. Finally, Ponzo and Scoppa (2018) showed a significantly larger number of cards against away teams in games between the teams from the same city that shared the same stadium in Italian Serie A.

also use a very rich set of variables that characterizes teams' values, standings before the respective game, distribution of players' values within teams, players' age and ability, game attendance, etc. In the following, we present some of the most important measures (a more comprehensive list of variables appears in Appendix A).

Following many previous studies, we used teams' monetary values obtained from a popular soccer website, *Transfermarkt*, as teams' proxies of ability.[12] The players' values were used to create additional measures such as the distribution of values between and within teams.[13] We also used data on betting odds, and teams' dummies when they play at home and separately when they play away.

As has been found previously in the literature, home advantage is affected by the attendance level. Therefore, we created a measure to reflect the attendance in a match. We used both attendance and attendance as share of the capacity of the stadium, which is the ratio between the number of viewers in a match and the maximal possible capacity of the respective stadium. Table 4.2 demonstrates that the attendance and share of capacity are unlikely to be the same between the two cases. The same can be said about teams' values that may represent teams' abilities. Therefore, selection is likely to be an issue in our settings. We discuss it in detail in the following section.

---

[12] See Peeters (2018), who showed that forecasts of international soccer results based on the *Transfermarkt's* valuations are more accurate than those based on standard predictors, such as the FIFA ranking and the ELO rating. See also Krumer and Lechner (2018) for a detailed discussion on *Transfermarkt* values.

[13] See Coates, Frick and Jewell (2016) for a discussion on the relationship between players' inequality in salaries and teams' performance.

Table 4.2: Descriptive statistics of the selected variables

	Home team from a different language area than the referee		Home team from the same language area as the referee	
Variable	*Mean*	*Standard Deviation*	*Mean*	*Standard Deviation*
Game Outcomes				
Home points	1.453	1.325	1.691	1.308
Away points	1.300	1.311	1.057	1.245
Home score	1.543	1.363	1.733	1.411
Away score	1.375	1.251	1.201	1.132
Difference in cards	-0.077	1.731	-0.524	1.812
Difference in cards defense	-0.055	1.306	-0.398	1.367
Difference in cards – midfield	-0.074	1.172	-0.080	1.241
Difference in cards – offense	0.052	0.879	-0.040	0.878
Difference in red cards	0.000	0.488	-0.041	0.471
Difference in red cards – defense	0.004	0.353	-0.047	0.349
Difference in red cards – midfield	-0.004	0.289	0.004	0.285
Difference in red cards – offense	0.000	0.162	0.001	0.200
Difference in yellow cards	-0.077	1.637	-0.482	1.719
Difference in yellow cards – defense	-0.060	1.208	-0.351	1.283
Difference in yellow cards – midfield	-0.070	1.119	-0.084	1.176
Difference in yellow cards – offense	0.052	0.842	-0.041	0.826
Game Characteristics				
Attendance	4,062	4,255	5,527	6,642
Stadium capacity	13,996	7,888	14,198	9,616
Share of capacity of the stadium	0.275	0.212	0.318	0.221
Super League (dummy variable)	0.443		0.427	
Round	17.42	9.94	17.03	10.18
Teams Characteristics				
Home team total value	10,231,299	8,938,050	11,850,712	11,988,144
Difference in total value	-2,060,080	10,722,945	2,061,376	9,962,904
Home team pre-game tournament points	19.99	15.07	22.21	16.85
Difference in pre-game tournament points	-3.270	12.789	2.265	12.719
Home team pre-game tournament ranking	7.214	3.749	6.114	3.806
Difference in pre-game tournament ranking	1.005	5.345	-1.107	5.339
Home team betting odds of winning	2.564	1.045	2.122	0.805
Difference in betting odds of winning	-0.739	2.494	-2.081	2.807
Home team HHI index of players' values	0.068	0.036	0.063	0.015
Difference in HHI index of players' values	0.005	0.036	-0.005	0.037
Home team ratio of top 3 to ranked 9–11 most valuable players	2.705	1.588	2.407	0.897
Difference in ratio of top 3 to ranked 9–11 most valuable players	0.346	1.699	-0.353	1.747
Home team average age	23.32	1.21	23.23	1.04
Away team average age	23.26	1.06	23.37	1.23
Observations	681		723	

Notes: This table presents average values and standard deviations (for non-binary variables) for selected variables. Differences are computed as home team values minus away team values.

5 Empirical Strategy

5.1 The causal question

We are interested in learning the effect of having a referee from the same linguistic area on different aspects of the outcome of a Swiss soccer game, such as allocation of yellow and red cards, as well as the difference in points between the teams as described in the previous section. If the allocation of different referees to games was entirely random, we would compare the means of these variables for home teams with referees from the same linguistic region to the means obtained for home teams with referees from a different linguistic area. The difference would be a consistent estimate of the desired effect. However, the distribution of the characteristics shown in Table 4.2 already points to deviations from randomness. Such deviations need to be taken into account in any estimation strategy if they are correlated with the outcomes of interest (e.g., Imbens and Wooldridge, 2009). Crucially, if some other factors differ between the two types of games and also plausibly explain some or all of the differences in the outcome variables of interest (so-called "confounders"), these must be controlled for in the estimation, usually by including observed measures of these factors or by the use of dummy variables.

Additionally, linguistic differences are likely to be connected to cultural differences. As Eugster et al. (2017) demonstrated, the cultural differences between the German-speaking and Romance-speaking regions have tangible effects on labor market outcomes. Therefore, such cultural differences may explain differences in the allocation of yellow and red cards. For example, it is possible that the playing styles of the teams, as well as the refereeing styles of the referees, would vary between different linguistic regions. These differences could then affect the number of cards awarded and the final scores. In addition, because of the imbalance in the proportions of teams and referees from each region, German-speaking teams are far more likely to share the referee's language than their French- and Italian-speaking counterparts. More specifically, home teams from the German-speaking area share the same language area with referee in 87 percent of the cases. The corresponding numbers for French and Italian teams are only 22 percent and 2 percent, respectively. As for referees, both French- and German-speaking referees have a roughly 50 percent chance of sharing the home team's language (the sole Italian referee accounts for only seven games). If not controlled for, these language-specific effects could act as confounders. For example, the differences in means in the outcome variables that we observe in Table 4.2 could be explained by the French- and Italian-speaking teams playing more aggressively than the German-speaking teams. Fortunately, any language-specific differences are effectively captured by the team-specific and referee-language-specific dummy variables, under the reasonable assumption that the cultural differences remain constant across the period of observation.

In addition to the issues mentioned above, there is the more general question of the referee selection process. As explained in Section 3.2, we only know which referees are chosen by a committee. However, the selection criteria are not publicly known. Assuming the committee's interests are to ensure refereeing fairness and quality, the relevant criteria would be the skills and experience of a referee, and avoidance of potential conflicts of interest. It is not clear whether language itself is taken into account as a source of conflict of interest, but the selection process should ensure that the linguistic discrimination observed in the data is not simply due to outright corruption. If, on the other hand, teams try to bribe referees, it could be that they would choose referees from their own linguistic region. This would then be the main driver of the observed effects. However, given the absence of any evidence for bribing the referees on a linguistic basis in Swiss soccer, we assume such a scenario to be very unlikely.

5.2 Estimator

In this paper, we utilize the recently upcoming causal machine learning literature (for an only slightly outdated survey of this literature, see Athey and Imbens, 2017, see also Athey, 2017, for a more general motivation of this literature). It combines the prediction power of the machine and statistical learning literature (for an overview see, e.g., Hastie, Tibshirani, and Friedman, 2009) with the microeconometric literature on defining and identifying causal effects (e.g., Imbens and Wooldridge, 2009). Recently, this literature has seen a surge of proposed methods, particularly in epidemiology and econometrics. Knaus, Lechner, and Strittmatter (2018) compared many of those methods systematically with respect to their theoretically properties as well as their performance in a simulation exercise. One conclusion from that paper is that random forest-based estimation approaches seem to outperform many alternative estimators.

The starting point of the causal forest literature is the causal tree introduced in a paper by Athey and Imbens (2016). In a causal tree, the sample is split sequentially into smaller and smaller strata, in which the values of the covariates X become more and more homogenous, in order to mitigate selection effects and to uncover effect heterogeneity. Once the splitting is terminated based on some stopping criterion, the treatment effect is computed within each stratum (called a "leaf") by computing the difference of the mean outcomes of treated and controls (possibly weighted by the conditional on X probabilities of being a treated or control observation). However, the literature on regression trees acknowledges that the sample may be rather unstable because of its sequential nature (if the first split is different, the full tree will likely lead to different final strata). A solution to this problem is so-called random forests. Their key idea is to induce some randomness into the tree-building process, build many trees, and then average the predictions of the many trees. The induced randomness is generated by using randomly generated subsamples (or bootstrap samples) and by considering for each splitting decision only a random selection of the covariates.

Wager and Athey (2018) used this idea to propose causal forests, which are based on a collection of causal trees with small final leaves.[14] Lechner (2018) developed these ideas further by improving on the splitting rule for the individual trees and by providing methods to estimate heterogeneous effects for a limited number of discrete policy variables (**G**roup **A**verage **T**reatment **E**ffects, GATE) at low computational costs, in addition to the highly disaggregated effects the literature has focused on so far (**I**ndividualized **A**verage **T**reatment **E**ffects, IATE). Furthermore, his paper suggested a way of performing unified inference for all aggregation levels. Finally, the approach is applicable to a multiple, discrete treatment framework. Since many of these advantages are important in the empirical analysis of this paper, this approach is used below. For all further technical details of the estimator, the reader is referred to Lechner (2018).

6 Results

Table 6.1 presents the main results of this paper. We see that when a referee does not share the same linguistic area with the home team, the difference in the total amount of cards (yellow and red) between home and away teams is -0.13 (untreated potential outcome). The negative sign suggests that the away team receives more cards than the home team, which is in line with the literature on referee bias in favor of the home teams in sports (Nevill, Balmer and Williams, 2002; Downward and Jones, 2007; Ponzo and Scoppa, 2018). However, this difference approaches -0.4 when a referee and the home team share the same linguistic area (treated potential outcome). This means that having a referee from the same linguistic region as the home team increases the gap between the numbers of cards in favor of a home team by about 0.27 cards, which is significant at the 6.1 percent level.

[14] Athey, Tibshirani, and Wager (2018) generalized this idea to many different econometric estimation problems.

When looking into different positions, we observe that defenders from the home team receive roughly the same number of cards as defenders from the away team when a referee shares the same linguistic area with the away team. However, when the referee and the home team are from the same linguistic area, the gap between the teams becomes -0.26, suggesting that the effect of having a referee from the same linguistic area when playing at home is -0.22 cards. This result is significant at the 2.9 percent level. A similar result, albeit with a lower magnitude, is observed for offensive players, whereas no significant result was found for midfielders.

When looking separately into yellow and red cards, we again find that having a referee from the same linguistic region as the home team increases the gap in red cards for defenders in favor of the home team. We find no significant effect for other positions. Regarding yellow cards, we see a significant effect for forwards. The effect for defenders is in the same direction as with red cards and even somewhat larger, but is not significant at conventional levels (p-val=0.13).

More importantly, such a bias is likely to affect the outcome of the game. Our findings suggest that having a referee from the same linguistic area when playing at home increases the difference in points between the home and away teams by 0.23 compared to the case when the away team shares the same linguistic region with the referee. The result is significant at the 3.4 percent level. To put this result into perspective, in a tight league where teams compete for the place in European tournaments with the possibility of large monetary prizes, every point is important.[15] For example, in the 2018/19 season, only three out of the four teams that had the same number of points qualified for the UEFA Europa League

[15] A win in the group stage of the UEFA Europa League is worth 360,000 Euros. Qualification from the group stage is worth an additional 800,000 Euros, not including revenue from TV rights and attendance. For more details, on the UEFA Europa League, see Krumer (2019).

tournament (FC Lugano, FC Luzern and FC Thun). The remaining team, FC St. Gallen, did not qualify because of a worse goal difference, which was the next tiebreaker criterion.

We were also interested in effect heterogeneity across different attendance levels. In Appendices B and C, we present the results for these effects across attendance and capacity shares levels, respectively. In general, the results indicate that there is no significant difference in the effects across different levels. However, we have to be cautious about these results because of a very low number of observations in each category.

Overall, the results provide empirical evidence that linguistic discrimination is a concern for Swiss society in general and for Swiss soccer in particular. Therefore, it is crucial that decision-makers are aware of such a problem. This awareness is important since proper feedback training to referees may reduce the bias (Plessner and Haar, 2006).

Table 6.1: Levels and effects of having a referee from the same language area

Variable	Untreated Potential Outcome	Treated Potential Outcome	Effect of Treatment	Standard Error of the Effect
Difference in cards	-0.13	-0.396	-0.266*	0.142
Difference in cards – defense	-0.039	-0.262	-0.223**	0.102
Difference in cards – midfield	-0.151	-0.047	0.104	0.096
Difference in cards – offense	0.06	-0.079	-0.139*	0.073
Difference in red cards	0.00967	-0.046	-0.055	0.038
Difference in red cards – defense	0.032	-0.046	-0.078***	0.025
Difference in red cards – midfield	-0.018	0.00093	0.019	0.024
Difference in red cards – offense	-0.00387	-0.00076	0.00311	0.0177
Difference in yellow cards	-0.14	-0.35	-0.211	0.133
Difference in yellow cards – defense	-0.071	-0.216	-0.145	0.096
Difference in yellow cards – midfield	-0.133	-0.048	0.085	0.09
Difference in yellow cards – offense	0.064	-0.078	-0.142**	0.069
Home team score	1.595	1.707	0.112	0.124
Away team score	1.342	1.211	-0.132	0.096
Difference in score	0.253	0.496	0.243	0.151
Tournament points earned by home team	1.472	1.706	0.234**	0.11

Notes: Average treatment effect. . *, **, *** denotes significance at the 10%, 5%, 1% level respectively. Weights-based inference used with sampling splitting as suggested by Lechner (2018).

7 Conclusion

Switzerland is a multi-lingual country with four official languages. It is also one of the richest countries in the world and is ranked sixth on the last World Happiness Ranking.[16] Despite all the wealth and happiness, there is still some tension between people from different language areas. We find evidence of in-group favoritism on the soccer pitch, which may represent real problems of the society (Elaad, Krumer and Kantor, 2018).

Our results are striking, given the high level of monitoring and training that the referees undergo. Identifying such in-group favoritism among highly trained referees calls for extra attention among Swiss decision makers from other professions, such as managers, judges, or members of hiring committees, who may have to make a decision in favor or against people from different language regions.

If differences in language is an issue in a country like Switzerland, then it is even more of a concern in less prosperous countries such as Moldova, Ukraine, and even Spain, where language-related debates have been much more violent. Governments should not underestimate the explosiveness of the linguistic differences and should act to reduce any future tension. For example, beyond compulsory school programs, governments may incentivize people studying non-native languages. This may help people (literally) better understand each other, which in turn may reduce bias and tension.

8 References

Angerer, S., Glätzle-Rützler, D., Lergetporer, P. and Sutter, M., 2016. Cooperation and discrimination within and across language borders: Evidence from children in a bilingual city. *European Economic Review*, 90, pp. 254–264.

Arrow, K. J., 1972. Models of job discrimination. In Racial discrimination in economic life. Lexington, Massachusetts: D.C. Heath, pp. 83–102.

[16] See https://countryeconomy.com/demography/world-happiness-index . Last accessed on 17/07/2019.

Ashraf, N., Camerer, C.F. and Loewenstein, G., 2005. Adam Smith, behavioral economist. *Journal of Economic Perspectives*, 19(3), pp. 131–145.

Athey, S., 2017. Beyond prediction: Using big data for policy problems, *Science* 355, pp. 483–485.

Athey, S. and Imbens, G.W., 2016. Recursive Partitioning for Heterogeneous Causal Effects, *Proceedings of the National Academy of Sciences of the United States of America*, 113 (27), pp. 7353–7360.

Athey, S. and Imbens, G.W., 2017. The State of Applied Econometrics: Causality and Policy Evaluation, *Journal of Economic Perspectives*, 31 (7), pp. 3–32.

Athey, S., Tibshirani, J. and Wager, S., 2018. Generalized Random Forests, *arXiv*: 1610.01271v4, forthcoming in the *Annals of Statistics*.

Becker, G. S., 1957. The economics of discrimination: an economic view of racial discrimination. University of Chicago.

Bertrand, M., Chugh, D. and Mullainathan, S., 2005. Implicit discrimination. *American Economic Review*, 95(2), pp. 94–98.

Coates, D. Frick, B. and Jewell, T., 2016. Superstar salaries and soccer success: The impact of designated players in Major League Soccer. *Journal of Sports Economics*, 17(7), pp. 716–735.

Downward, P. and Jones, M., 2007. Effects of crowd size on referee decisions: Analysis of the FA Cup. *Journal of Sports Sciences*, 25(14), pp. 1541–1545.

Efferson, C., Lalive, R. and Fehr, E., 2008. The coevolution of cultural groups and ingroup favoritism. *Science*, 321(5897), pp. 1844–1849.

Elaad, G., Krumer, A. and Kantor, J., 2018. Corruption and sensitive soccer games: cross-country evidence. *Journal of Law, Economics, and Organization*, 34(3), pp. 364–394.

Eugster, B., Lalive, R., Steinhauer, A. and Zweimüller, J., 2017. Culture, work attitudes, and job search: Evidence from the Swiss language border. *Journal of the European Economic Association*, 15(5), pp. 1056–1100.

Fidrmuc, J., Ginsburgh, V. and Weber, S., 2009. Voting on the choice of core languages in the European Union. *European Journal of Political Economy*, 25(1), pp. 56–62.

Gallo, E., Grund, T. and Reade, J., 2013. Punishing the foreigner: implicit discrimination in the Premier League based on oppositional identity. *Oxford Bulletin of Economics and Statistics*, 75(1), pp. 136–156.

Gill, D. and Stone, R., 2010. Fairness and desert in tournaments. *Games and Economic Behavior*, 69(2), pp. 346–364.

Greenwald, A.G. and Banaji, M.R., 1995. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), pp. 4–27.

Hastie, T., Tibshirani, R. and Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd edition, Springer (10th printing with corrections, 2013).

Imbens, G.W. and Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), pp. 5–86.

International Football Association Board, 2017. Laws of the Game 2017/18. Retrieved from http://resources.fifa.com/mm/document/footballdevelopment/refereeing/02/90/11/67/lawsofthegame2017-2018-en_neutral.pdf

Konow, J., 2003. Which is the fairest one of all? A positive analysis of justice theories. *Journal of Economic Literature*, 41(4), pp. 1188–1239.

Knaus, M. C., Lechner, M. and Strittmatter, A., 2018. The Performance of Machine Learning Approaches to Estimate Heterogeneous Policy Effects, *arXiv: 1810.13237v2*.

Krumer, A., 2019. Testing the effect of kick-off time in the UEFA Europa League. *European Sport Management Quarterly*, forthcoming.

Krumer, A. and Lechner, M., 2018. Midweek effect on soccer performance: Evidence from the German Bundesliga. *Economic Inquiry*, 56(1), pp. 193–207.

Lechner, M. 2018. Modified Causal Forest for Estimating Heterogeneous Causal Effects, arXiv: 1812.09487v2.

Mongeon, K. and Longley, N., 2015. Testing for ethnicity discrimination among NHL referees: A duration model approach. *Eastern Economic Journal*, 41(1), pp. 86–101.

Mongeon, K. and Longley, N., 2017. Team-Level Referee Discrimination in the National Hockey League. In Breaking the Ice (pp. 131–147). Springer, Cham.

Nevill, A.M., Balmer, N.J. and Williams, A.M., 2002. The influence of crowd noise and experience upon refereeing decisions in football. *Psychology of Sport and Exercise*, 3(4), pp. 261–272.

Peeters, T., 2018. Testing the Wisdom of Crowds in the field: Transfermarkt valuations and international soccer results. *International Journal of Forecasting*, 34(1), pp. 17–29.

Phelps, E.S., 1972. The statistical theory of racism and sexism. *American Economic Review*, 62(4), pp. 659–661.

Plessner, H. and Haar, T., 2006. Sports performance judgments from a social cognitive perspective. *Psychology of Sport and Exercise*, 7(6), pp. 555–575.

Ponzo, M. and Scoppa, V., 2018. Does the home advantage depend on crowd support? Evidence from same-stadium derbies. *Journal of Sports Economics*, 19(4), pp. 562–582.

Pope, B.R. and Pope, N.G., 2015. Own-nationality bias: evidence from UEFA Champions League football referees. *Economic Inquiry*, 53(2), pp. 1292–1304.

Price, J. and Wolfers, J., 2010. Racial discrimination among NBA referees. *The Quarterly Journal of Economics*, 125(4), pp. 1859–1887.

Sandberg, A., 2017. Competing identities: a field study of in-group bias among professional evaluators. *The Economic Journal*, 128(613), pp. 2131–2159.

Sumner, W.G. 1906. Folkways: A study of the sociological importance of usages, manners, customs, mores, and morals. Boston: Ginn and Company.

Swiss Federal Government, 1999. SR 101 Bundesverfassung der Schweizerischen Eidgenossenschaft vom 18. April 1999. Retrieved 2018-03-21, from https://www.admin.ch/opc/de/classified-compilation/19995395/index.html#a8

Swiss Federal Statistical Office, 2019. Ständige Wohnbevölkerung nach Hauptsprachen in der Schweiz, 2017. Retrieved 2019-08-14, from https://www.bfs.admin.ch/bfs/de/home/statistiken/bevoelkerung/sprachen-religionen/sprachen.assetdetail.7466556.html

Swiss Football Association, 2017. Réglement de Jeu (Rules of the Game). Retrieved from https://org.football.ch/portaldata/28/Resources/dokumente/fr/02_reglement_de_jeu_et_lois_du_jeu/2.1_Reglement_de_jeu_(juin_2017).pdf

Szymanski, S., 2000. A market test for discrimination in the English professional soccer leagues. *Journal of Political Economy*, 108(3), pp.590–603.

Wager, S., and S. Athey (2018): Estimation and Inference of Heterogeneous Treatment Effects using Random Forests, *Journal of the American Statistical Society*, 113(523), pp.1228–1242.

Werlen, I., Baumgartner, J., & Rosenberger, L. (2011). Sprachkompetenzen der erwachsenen Bevölkerung in der Schweiz. Seismo Zürich.

Yuki, M., 2003. Intergroup comparison versus intragroup relationships: A cross-cultural examination of social identity theory in North American and East Asian cultural contexts. *Social Psychology Quarterly*, 66(2), pp.166–183.

Zitzewitz, E., 2006. Nationalism in winter sports judging and its lessons for organizational decision making. *Journal of Economics and Management Strategy*, 15(1), pp.67–99.

Appendix A: Descriptive statistics

The following table contains descriptive statistics for the variables (outcomes, treatment, and controls) relevant for this study.

Variable	Home team from a different language area than the referee N=681	Home team from the same language area as the referee N=723
Game Characteristics		
Attendance	4,062.89 (4,255.07)	5,527.58 (6,642.05)
Attendance as share of capacity of stadium	0.28 (0.21)	0.32 (0.22)
Challenge league (dummy variable)	0.56 (0.5)	0.57 (0.5)
Season	2,011.45 (3.22)	2,011.26 (3.31)
Stadium capacity	13,996.03 (7,888.69)	14,198.63 (9,616.09)
Sunday (dummy variable)	0.39 (0.49)	0.37 (0.48)
Super league (dummy variable)	0.44 (0.5)	0.43 (0.5)
Tournament round	17.43 (9.94)	17.03 (10.18)
Average odds of a draw	3.5 (0.42)	3.65 (0.61)
Home average odds of winning	2.56 (1.05)	2.12 (0.81)
Diff* in average odds of winning	-0.74 (2.49)	-2.08 (2.81)
High odds of a draw	3.81 (0.59)	4.01 (0.85)
Home high odds of winning	2.83 (1.28)	2.31 (0.99)
Diff in high odds of winning	-0.93 (3.15)	-2.58 (3.64)
Game Outcomes		
Diff in cards	-0.08 (1.73)	-0.52 (1.81)
Diff in cards – defense	-0.06 (1.31)	-0.4 (1.37)
Diff in cards – midfield	-0.07 (1.17)	-0.08 (1.24)
Diff in cards – offense	0.05 (0.88)	-0.04 (0.88)
Diff in red cards	0 (0.49)	-0.04 (0.47)
Diff in red cards – defense	0 (0.35)	-0.05 (0.35)
Diff in red cards – midfield	0 (0.29)	0 (0.29)
Diff in red cards – offense	0 (0.16)	0 (0.2)
Diff in yellow cards	-0.08 (1.64)	-0.48 (1.72)
Diff in yellow cards – defense	-0.06 (1.21)	-0.35 (1.28)
Diff in yellow cards – midfield	-0.07 (1.12)	-0.08 (1.18)
Diff in yellow cards – offense	0.05 (0.84)	-0.04 (0.83)
Home score	1.54 (1.36)	1.73 (1.41)
Away score	1.38 (1.25)	1.2 (1.13)
Diff in score	0.17 (1.91)	0.53 (1.84)
Tournament points earned by home team	1.45 (1.33)	1.69 (1.31)
Team Characteristics – Market Value (in Euros)		
Home average value	363,598.22 (269,220.74)	434,817.5 (379,433.8)
Diff in average value	-82,401.65 (338,394.62)	84,402.78 (317,321.79)
Home Herfindahl-Hirschman Index	0.07 (0.04)	0.06 (0.02)
Diff in Herfindahl-HirschmanÂ Index (HHI)	0.01 (0.04)	-0.01 (0.04)
Home median value	293,329.66 (200,650.94)	354,740.66 (281,416.91)
Diff in median value	-72,243.02 (252,372.22)	75,159.06 (237,950.39)
Home standardized HHI	0.04 (0.04)	0.03 (0.02)
Diff in standardized HHI	0.01 (0.04)	-0.01 (0.04)
Home total value	10,231,299.56 (8,938,050.78)	11,850,712.31 (11,988,144.87)
Diff in total value	-2,060,080.76 (10,722,945.58)	2,061,376.21 (9,962,904.08)
Home value coefficient of variation	0.8 (0.29)	0.72 (0.22)
Diff in value coefficient of variation	0.08 (0.35)	-0.08 (0.35)

Home value ratio of top 11 value players over 12–22	1.94 (1.52)	1.76 (1.74)
Away value ratio of top 11 value players over 12–22	1.86 (1.65)	1.89 (1.68)
Home value ratio of top 3 value players over 12–14	3.41 (1.93)	3.07 (1.44)
Diff in value ratio of top 3 value players over 12–14	0.43 (2.23)	-0.46 (2.29)
Home value ratio of top 3 value players over 9–11	2.71 (1.59)	2.41 (0.9)
Diff in value ratio of top 3 value players over 9–11	0.35 (1.7)	-0.35 (1.75)
Home value standard deviation	311,291.59 (296,067.87)	360,942.45 (419,927.31)
Diff in value standard deviation	-55,473.4 (404,941.39)	57,886.05 (389,348.6)
Team Characteristics – Age		
Home age coefficient of variation	0.19 (0.03)	0.18 (0.03)
Diff in age coefficient of variation	0.01 (0.03)	-0.01 (0.03)
Home age ratio of top 11 value players over 12–22	1.03 (0.22)	1.06 (0.08)
Away age ratio of top 11 value players over 12–22	1.05 (0.08)	1.04 (0.2)
Home age standard deviation	4.35 (0.69)	4.17 (0.67)
Diff in age standard deviation	0.17 (0.87)	-0.17 (0.86)
Home average age	23.33 (1.21)	23.23 (1.05)
Away average age	23.27 (1.06)	23.38 (1.24)
Home average age of players over 20	25.42 (1.09)	25.23 (0.97)
Diff in average age of players over 20	0.16 (1.29)	-0.2 (1.28)
Home average age of top 11 value players	25.3 (1.48)	25.07 (1.23)
Diff in average age of top 11 value players	0.22 (1.82)	-0.3 (1.8)
Home maximum age	33.49 (2.62)	32.93 (2.54)
Diff in maximum age	0.49 (3.35)	-0.55 (3.36)
Home median age	22.68 (1.45)	22.55 (1.28)
Diff in median age	0.14 (1.78)	-0.19 (1.74)
Home minimum age	16.83 (1.12)	16.97 (1.27)
Diff in minimum age	-0.26 (1.47)	0.18 (1.64)
Team Characteristics – Tournament standing		
Home pre-game net goals	-3.09 (12.75)	0.89 (12.91)
Diff in pre-game net goals	-4.84 (18.59)	3.29 (18.27)
Home pre-game tournament points	20 (15.07)	22.21 (16.85)
Diff in pre-game tournament points	-3.27 (12.79)	2.27 (12.72)
Home pre-game tournament ranking	7.21 (3.75)	6.11 (3.81)
Diff in pre-game tournament ranking	1.01 (5.35)	-1.11 (5.34)
Starting line-up Characteristics – Value		
Home average value	438,414.76 (401,014.41)	559,333.58 (621,335.77)
Away average value	561,223.47 (573,309.18)	424,398.34 (401,562.31)
Home HHI	0.13 (0.03)	0.12 (0.03)
Diff in HHI	0 (0.04)	0 (0.04)
Home median value	382,466.96 (312,915.62)	504,910.1 (568,511.89)
Diff in median value	-126,813.51 (463,160.63)	129,128.63 (495,588.64)
Home standardized HHI	0.03 (0.03)	0.03 (0.03)
Diff in standardized HHI	0 (0.04)	0 (0.04)
Home total value	4,822,562.41 (4,411,158.48)	6,152,669.43 (6,834,693.42)
Diff in total value	-1,350,895.74 (5,699,839.9)	1,484,287.69 (6,183,756.66)
Home value ratio of top 3 value players over 9–11	4.63 (4.54)	4.74 (4.25)
Diff in value ratio of top 3 value players over 9–11	-0.34 (6)	0.19 (5.53)
Home value standard deviation	276,080.24 (361,474.6)	349,733.55 (544,434.87)
Diff in value standard deviation	-55,882.98 (462,244.04)	84,002.78 (561,676.8)
Starting line-up Characteristics – Age		
Home age standard deviation	3.91 (0.87)	3.78 (0.82)
Diff in age standard deviation	0.12 (1.16)	-0.12 (1.11)
Home average age	25.23 (1.57)	25.06 (1.45)
Away average age	25.01 (1.48)	25.27 (1.6)
Home maximum age	32.15 (2.41)	31.68 (2.39)

Diff in maximum age	0.44 (3.21)	-0.51 (3.29)
Home median age	24.83 (1.99)	24.65 (1.85)
Diff in median age	0.2 (2.69)	-0.19 (2.6)
Home minimum age	19.76 (1.49)	19.84 (1.43)
Diff in minimum age	-0.01 (2.12)	0.02 (2.05)

* Differences always computed as home team values minus away team values.

Appendix B: Test for treatment effect heterogeneity across attendance levels

Variable	Wald Statistic	p-value under $\chi^2(9)$
Difference in cards	10.49	0.3120
Difference in cards – defense	8.19	0.5147
Difference in cards – midfield	8.19	0.5146
Difference in cards – offense	5.83	0.7569
Difference in red cards	7.31	0.6054
Difference in red cards – defense	7.44	0.5910
Difference in red cards – midfield	17.19	0.0458
Difference in red cards – offense	7.31	0.6053
Difference in yellow cards	12.12	0.2065
Difference in yellow cards – defense	6.44	0.6954
Difference in yellow cards – midfield	8.74	0.4618
Difference in yellow cards – offense	6.08	0.7319
Home team score	2.67	0.9760
Away team score	2.9	0.9683
Difference in score	5.95	0.7445
Tournament points earned by home team	1.87	0.9934

Note: Split into 10 roughly equally sized categories.

Appendix C: Test for treatment effect heterogeneity across capacity share levels

Variable	**Wald Statistic**	**p-value under $\chi^2(9)$**
Difference in cards	12.55	0.1838
Difference in cards – defense	13.27	0.1507
Difference in cards – midfield	8.33	0.5013
Difference in cards – offense	4.94	0.8392
Difference in red cards	6.47	0.6919
Difference in red cards – defense	6.92	0.6451
Difference in red cards – midfield	13.23	0.1525
Difference in red cards – offense	5.36	0.8023
Difference in yellow cards	12.93	0.1660
Difference in yellow cards – defense	11.53	0.2413
Difference in yellow cards – midfield	11.09	0.2696
Difference in yellow cards – offense	4.75	0.8559
Home team score	2.92	0.9672
Away team score	7.16	0.6207
Difference in score	8.37	0.4977
Tournament points earned by home team	2.61	0.9777

Note: Split into 10 roughly equally sized categories.