INSTITUTO DE ECONOMÍA

MAGÍSTER en ECONOMÍA

**2020**

# Scoring for Remain: The Political Impact of Premier League on the Brexit Referendum

**Diego Fuenzalida F.**

www.economia.uc.cl

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA

# TESIS DE GRADO
# MAGISTER EN ECONOMIA

**Fuenzalida, Fernández, Diego**

**Agosto, 2020**

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA

# Scoring For Remain: The Political Impact of Premier League on the Brexit Referendum

**Diego Fuenzalida Fernández**

Comisión

Francisco Gallego Yañez
Emilio Depetris-Chauvin

**Santiago, Agosto, 2020**

# Scoring for Remain: The Political Impact of Premier League on the Brexit Referendum*

Diego Fuenzalida F.†
Instituto de Economía
Pontificia Universidad Católica de Chile

August 7, 2020

**Abstract** *This paper explores the role of emotions driven by sports as political drivers. To do so, I study whether the nationality compositions, player performance, and geographic distribution of the fan base of Premier League teams on the eve of the Brexit referendum. I find that increases in the relative performance of foreign players in the 2015-16 season produces economic and statistically significant increases in the Remain vote share by approximately 1 percentage point at the district level. I do not find evidence that the relative number of foreign players explains cross-district differences in voting. I also find that expectations are important and that the effects on voter preferences seem to display persistence over time. I confirm that these results are causal by implementing a series of falsification and robustness checks. This set of results implies that voter's priors about immigrants change not by considering exposure to foreign players (an extensive margin effect) but rather by considering their relative performance (an intensive margin effect).*

*Thesis written as a master's student at the Department of Economics, Pontificia Universidad Católica de Chile. I would like to thank my thesis advisors Francisco Gallego and Emilio Depetris-Chauvin for their continuous guidance, support and comments. I am also indebted and grateful to God and my family and friends. I especially thank Santiago Cruz and Luis Hernán Cubillos for their help in obtaining the data and Carmen Luz Edwards, Gonzalo de la Carrera, Daniel Correa, Sebastián Figari and Miguel De Iruarrizaga for their support, ideas, and continuous proofreading. I am also grateful to Miguel De Iruarrizaga specifically for his partnership in starting this work a year ago for the Economía Institucional course taught by Francisco Gallego. All errors or omissions are my own responsibility.

†Comments at dfuenzalida1@uc.cl.

## Contents

# 1 Introduction

Few things ignite the passions and emotions like sports, and emotions, in turn, are recognized as drivers of people's behavior. Emotions have also been shown to have an impact on political processes[1]. A case in point is association football (henceforth, football) in the UK, which since its emergence has gone through a rapid transformation. Starting as an elite activity, it rapidly became a competition that is popular with all social classes (and has spread all over the world), awakening local fervor and zeal for its teams and players. The enthusiasm for and relevance of UK football has led its main league, the Premier League, to become one of the most important football leagues in the world in terms of the number of fans and followers, the amounts of money invested, and the high level of its in-game competition[2]. This latter characteristic has motivated many teams to look for the best players available, regardless of their origins, religion, race, or political affiliation, as a strategy to maximize the chances of a good result[3]. For the same reason, today, it could be reasonably stated that regardless of origins, religion, or race, players are admired indiscriminately by their teams' supporters, as long as the players are perceived to be important to the team. Thus, the following question naturally arises: can Premier League teams and players affect their supporters' attitudes? Specifically, can the fact that players have different nationalities change voters' attitudes? If so, is this due to a change in voters' perceptions or knowledge of foreigners? Are voters changing their behavior as a result of being exposed to foreign stars? Or are voter perceptions of foreign players simply correlated with ex ante policy or ideological views that ultimately guide voter behavior? If football players do indeed have an impact on voters, is it driven solely by the level of exposure to foreigners? Or is the driver the performance payoff that voters perceive foreigners to deliver to their favorite teams?

To help answer these questions, I take advantage of the context and timing provided by the Brexit referendum and the contemporaneous Premier League season in the UK. The Brexit referendum was held on June 23, 2016, while the 2015-16 Premier League season ended on May 25, 2016. In addition, specific characteristics of this election and the Premier League are leveraged in the empirical approach.

The Brexit referendum is particularly interesting because people voted directly for a specific public policy. Thus, the referendum is much closer to a form of direct rather than representative democracy[4]. In addition, among the main arguments of the pro-Leave campaign was that the

[1] See, for instance, Depetris-Chauvin et al. (2020), Campante et al. (2020), and/or Brader (2005).

[2] Some facts to illustrate this point: The average Premier League attendance in 2015-16 was 36,450. The revenue for the same season was 4,865 M euros. It was seconded in Europe by the German Bundesliga with 2,712M Euros.

[3] This is not necessarily true for all teams. However, discriminatory attitudes have decreased significantly by teams in the UK since the 1970s and 1980s. See, for instance, Soccerconomics, Chapter 6 by Kuper and Szymanski (2010).

[4] Although the final decision on whether to leave the European Union or not belongs to the UK parliament, the referendum allowed people to vote directly on that policy. Eventually, the decision was made by the MPs after

UK lacked control over immigration and that this had generated pressures on the local economy. Thus, the context of the Brexit referendum was highly informed by the immigration debate. Under the slogan \Take back control", parties such as UKIP (UK Independence Party) promised that leaving the EU would solve the immigration problem (Gietel-Basten, 2016). The option to leave the EU (\Leave") prevailed with 51.9% of the vote over the option to stay (\Remain"). This was indeed a narrow outcome, as was anticipated by preelection polls.

My focus on the Premier League is motivated by three other considerations in addition to the political context. First, Premier League fans are strongly attached to their local teams. However, local teams also have fans all over the UK, which is fundamental for the empirical strategy employed in this paper, since the variation in spatial distribution is what allows the correct identi cation of the estimators. Second, football is the most popular sport in the UK[5], impacting a large part of the population. Third, the level of competition in the league helps mitigate concerns that could threaten the correct estimation of a causal e ect.

Motivated by the literature that studies the impact of emotions on people's attitudes, and arguing that voters can update their beliefs by observing football players, this paper studies whether voter exposure to and the performance of foreign football players can a ect a political election. Focusing on the impact of this exposure and Premier League players' performance on votes in the 2016 Brexit referendum, I nd evidence con rming the intuition on the power of sports to change attitudes. This paper shows causal evidence that outstanding foreign football players did indeed have a signi cant e ect on the electoral performance of Remain (improving it) in the referendum. The work shows that the results are not driven by the number of foreigners on the preferred team but exclusively by the relative performance of foreign players, which I interpret to be perceived as a payo by the people of each district.

The research design exploits the geographical variation in the distribution of football fans across the UK. As an example, in the 2015-16 season, there were teams from several regions of England and one team from Wales in the Premier League. However, each team's fandom is not exclusive to people living in that team's district, as each one has fans in virtually every district in the UK. This, added to the fact that there is heterogeneity in the team-identi cation percentages across districts, provides a rich source of geographic variation in the data. Interacting the spatial distribution of fans with the nationality composition of the teams and the performance of local and foreign players provides a potentially exogenous source of variation that allows us to identify, under certain assumptions, the causal e ects of the presence and performance of foreign football players on voters.

I nd that a one-standard-deviation increase in the relative performance of foreign players increased the Remain vote share by 0.917 percentage points at the district level. This corresponds

the 2019 general election. However, it is the closest example of a direct democratic election within the British political system.

[5]With attendance in 2017 of 47.6 million, followed in second place by horse racing with 7.5 million (Statista).

to approximately 1/2 of the vote share the Remain campaign would have needed to win the referendum. Conversely, I find no evidence that the relative number of foreign players explains cross-district differences in voting. This set of results suggests that what matters is not mere exposure to foreign players but rather their relative performance, with voters using local players as the point of comparison within the season.

I next turn my attention to whether there is persistence in this effect and find that the relative performance of foreign players in the previous season also had an impact on the election. This suggests that voters update their beliefs, learning from what they observe year after year. Additionally, I find that expectations regarding players at the beginning of the season are also relevant, slightly increasing the identified effect.

To validate the identification strategy, I conduct different falsification tests with former elections. If the treatment is capturing omitted variables affecting the estimate, then the identified effect should also be present in tests based on previous elections. I find that for no former election is the treatment effect significantly different from zero. This offers evidence to support the identifying assumption of exogeneity on the constructed treatment.

This work relates and contributes to several strands of literature. First, I show that emotional responses driven by sports are relevant and can have impacts in many areas. Several papers have studied the impact of sports on the attitudes and beliefs of the population. The sports-related literature covers a wide variety of topics, ranging from how emotions affect violence to how sports can reduce prejudice or impact political processes. The link between sports and emotional responses has been studied in the context, among others, of the NFL[6] (Gantz, Bradley and Wang 2006, Rees and Schnepel 2009, Card and Dahl 2011). Card and Dahl 2011, for instance, studied how family violence increases when local American football teams suffer an unexpected defeat. Despite this, emotions driven by sports are not only linked to negative responses. One of the recent works in this topic is from Alrababa'h, et al. (2019)[7], who attempts to answer whether exposure to celebrities can reduce prejudice using the case of Mohamed Salah's arrival to Liverpool Football Club and its effect on Islamophobia and finding a reduction in hate crimes and negative beliefs about Muslim people. On this topic, an interpretation of the results is that being exposed to foreigners who are perceived positively by society may reduce immigration prejudices.

Second, I provide evidence that sports and shared experiences can have an impact on political processes. An example is found in Depetris-Chauvin, Durante, and Campante (2020). They find evidence that football can help in the formation of national identities by reducing ethnic conflict. They study the impact on these variables when there are victories in World Cup qualifiers by teams from sub-Saharan African countries, which are characterized by high ethnic fragmentation. On the same topic, several works have shown how shared collective experiences

[6] National Football League. Although the name of the sport is the same, it refers to the sport that encompasses neither foot nor ball, known as American Football, the most popular in the USA.

[7] This is, in fact, one of the main inspirations for this work.

can a ect individual attitudes in politically relevant ways (Clingingsmith, Khwaja, and Kremer 2009; Madestam and Yanagizawa-Drott 2011; Kaplan and Mukand 2014). This work contributes to the literature on these topics, showing that sports and the signals given by elite football players and observed by voters can change attitudes, even those related to electoral processes. This is also in line with the study by Alrababa'h, et al. (2019), who used the contact hypothesis and the construct of role models as a framework; their ndings suggested that the population's exposure to relevant gures or celebrities from di erent backgrounds positively changes attitudes or beliefs towards foreigners.

Last but not least, this work is related to the literature that has tried to identify and understand the determinants and causes of the result of the Brexit referendum. Becker, Fetzner and Novy (2017) conducted an extensive, comprehensive analysis on which covariates were critical to the election outcome at the district level. Fetzner (2019) studied the causal e ect of austerity on the Brexit vote. Viskanic (2017) tried to establish causality between immigration, speci cally the arrival of Polish immigrants to the UK, and the referendum results. The results of this work suggested that signals given by foreigners and attitudes towards immigrants did, in fact, matter at the district level and that, had these signals been better, the outcome might have been di erent. These results could also indicate that voters generalize the characteristics of foreign stars to other foreigners.

The remainder of the paper is organized as follows: Section 2 outlines the context and background of the 2016 Brexit referendum and discusses the theoretical framework. Section 3 introduces the data. Sections 4 and 5 present and discuss the empirical strategy and results for the district-level analysis, respectively. Section 6 discuss and presents some robustness checks, and section 7 presents the concluding remarks.

# 2 Background

## 2.1 The 2016 Brexit Referendum

The 2016 Brexit referendum was held on June 23, 2016, in the United Kingdom. British voters were called upon to vote on whether the UK should remain as a member of the European Union (henceforth, EU) or leave it. Voting took place in 41,000 polling stations across 382 voting areas. The voting areas, which were similar to those in the 2014 elections for Members of European Parliament, were called local authorities { districts (LADs, and henceforth, districts)[8]. According to the British government's open data portal, nationwide turnout, computed as the ratio of total ballots cast to eligible voters, was 72.21%. Turnout in this referendum was the highest since the 1992 general election. The majority voted for the \Leave" option, which garnered 51.89% of the share of valid votes, while the \Remain" option achieved a vote share of just

[8]In this election, Northern Ireland encompassed one large voting area, making it an outlier in terms of size and number of voters

48.11%. Of the 382 districts (including Northern Ireland), the majority of voters in a total of 270 of them favored the Leave option.

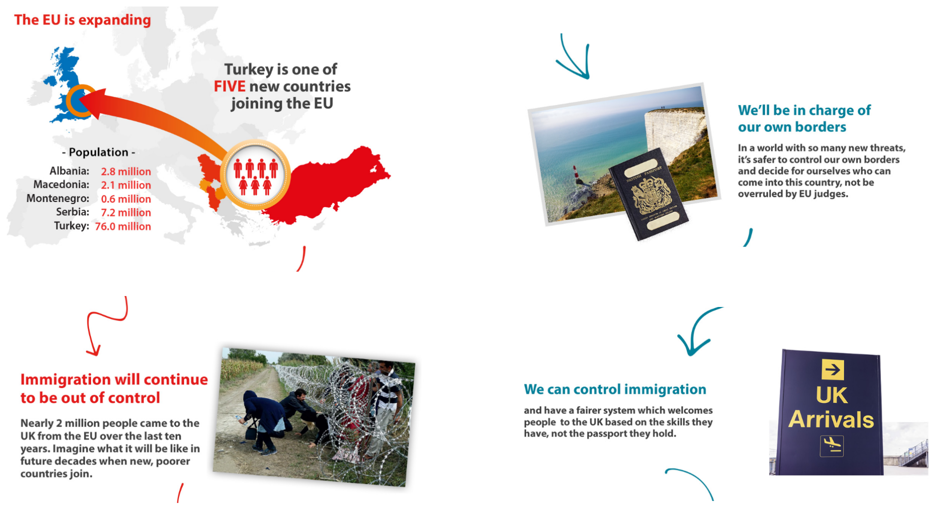


Figure 1: Example of Take Back Control Campaign

The call for the referendum was made in 2015 by Conservative prime minister David Cameron, making good on his promise to do so if the Conservative Party won the majority of votes in that year's general election campaign. Of all the major parties, the only party not to take an o cial position on the referendum was the Conservative Party. The Labour Party, the Liberal-Democrats, the Green Party and the Scottish National Party, for their part, all campaigned for Remain, while UKIP (UK Independence Party) campaigned for Leave. Among the pro-Leave arguments, an important one was that the UK lacked control over migration and that this generated pressures on the local economy. Under the slogan of \Take Back Control"[9], parties such as UKIP and a clutch of Eurosceptic Conservatives promised that leaving the EU would solve the immigration problem (Gietel-Basten, 2016). Figure 1 shows pictures taken from the \Take Back Control" campaign: on the left, in red, are images of what the campaign argued would happen if the UK stayed in the EU; on the right, in blue, the images shows some bene ts of leaving the EU.

Interestingly, the pro-Remain campaign barely addressed the migration issue. Instead, the campaign mainly focused on social security, trade, and economic issues (LSE, The Brexit Collection). This is important because if voters (or at least, the marginal voter) perceived this election as a referendum on migratory policy, the messaging of the Remain campaign might not have had much in uence on the marginal voter. Dustmann and Preston (2007) distinguish three channels that determine attitudes towards immigration: labor, welfare, and racial or cultural concerns. Dustmann, Fabbri, and Preston (2005) and Wads-worth, Dhingra, Ottaviano, and Van Reenen (2016) proposed that immigration|and increases in immigration{had had no negative e ect on the economic conditions of the local population, which suggests that a policy decision on this issue may have been a ected by cultural considerations. Moreover, Card et al. (2012) showed that cultural perceptions of immigrants drive overall opinion towards immigration policies, while economic evaluations of immigrants' contributions are less relevant. In this way, it becomes interesting to evaluate how perceptions of \what is good" can a ect a political election.

## 2.2 Conceptual Framework

As we have seen above, emotions can be a strong driver conditioning the political decisions of voters. On the other hand, it has been shown that sports strongly a ect people's attitudes and beliefs. Unlike previous works on this topic, which use the results of teams as a whole and study how these a ect agents' attitudes or beliefs, in this work, I study whether variation in the characteristics of those who cause these emotions (i.e., the origin of the players, who are ultimately the ones who achieve results for teams) can have e ects on attitudes or beliefs, speci cally in political elections. To understand the mechanisms that operate on voters, I consider a simple conceptual framework of signals. Here, voters have a utility function that depends on their beliefs and the political option chosen by the majority. Some voters also look at the number of immigrant players on their preferred football teams and how these players perform relative to

[9]www.voteleavetakecontrol.org

voters' expectations. Depending on this performance, voters may or may not update their beliefs.

For simplicity, I assume that the referendum was focused only on the issue of immigration policies[10]. As voter perceptions seem to be the overall driver, I now explore potential empirical implications related to how providing information about immigrants may affect an election. I base this framework on a simple learning model developed by Arias, Larreguy, Marshall, and Querubin (2018). I consider a simple decision model in which voters update their beliefs about immigration based on informative signals. Voters can choose between two options: Remain or Leave. For simplicity, I assume that the marginal cost of voting is equal to zero. The Remain option refers to the public policy to continue the acceptance of immigrants into the country, while Leave refers to the closure of borders to immigration. For analytical simplicity, the expected utility that voter $i$ associates with each policy $p \in \{R; L\}$ is a function of fixed ideological preferences and the elected policy, and therefore:

$$U_i^p = E[F({}_i; {}_p)]$$

where ${}_i \in \mathbb{R}$ is a positive or negative bias in ideological preferences distributed equally across the entire mass of voters and ${}_p$ is the chosen policy. Therefore, $i$ will vote for Remain $R$ if $U_i^R > U_i^L$ or for Leave $L$ if $U_i^R < U_i^L$.

Voters can learn about either immigration or immigrants from a signal given by foreigners. In particular, I assume that all voters share the same distribution of prior beliefs about immigration. Voters observe signals from their local football clubs and their players[11]. Voters can observe two kinds of signals. The first signal $s_E$ is an extensive-margin signal, which refers to how many foreign players they observe. The higher the number of foreigners who play on a football team, the greater is $s_E$. The second signal $s_I$ is an intensive-margin signal, referring to the quality of each foreign player that is observed. The better the performance of foreign players, the higher is $s_I$. The empirical question of this paper is whether $s_E$ and/or $s_I$ can affect voting decisions.

The timing of the model is as follows. First, voters have an expected utility of ${}_p$ for each policy given their prior preferences ${}_i$. Second, foreign players can send signals on the extensive $s_E$ and/or intensive margin $s_I$, which voters may or may not observe. Whether they do depends on whether the voter is a fan of the team where the foreign player plays. Third, voters who have observed the signals update their preferences ${}_i$; therefore, $U_i^p = E[F({}_i; {}_p)|s]$. Finally, the election is held, and voters reveal their preferences.

As I do not know whether increased knowledge about foreigners is perceived as good or bad among voters, I begin by being skeptical about the sign of the effect that $s_E$ may have on vot-

[10] Although this is not necessarily true, it is an assumption that does not change the results and only simplifies the analysis

[11] I refer as a local football club to the one the voter is a follower of.

ers. Nevertheless, I expect a positive signal at the intensive margin to improve the expected utility of the Remain option, thereby improving the Remain campaign's vote share. Additionally, both signals may be relevant at the same time in the process of updating voter preferences. Therefore, I formulate two hypotheses that may operate either separately or together.

**Hypothesis 1—Exposure hypothesis** Voters update their beliefs by observing $s_E$.

$$\mathbb{E}\left[F(\delta_i, \theta_p)\right] \neq \mathbb{E}\left[F(\delta_i, \theta_p) | \, s_E\right]$$

Note that the policy in question does not matter in this proposition. If exposure to foreign players is important for changing agents' priors, regardless of the quality of the players or the payoff that they bring, we should observe an effect different from zero when there are more foreign players on a team. I call this the exposure or familiarity hypothesis, referring to the increase in voter knowledge of foreigners.

**Hypothesis 2—Payoff hypothesis** Voters update their beliefs by observing $s_I$.

$$\mathbb{E}\left[F(\delta_i, \theta_R)\right] < \mathbb{E}\left[F(\delta_i, \theta_R) | \, s_I\right]$$

or

$$\mathbb{E}\left[F(\delta_i, \theta_L)\right] > \mathbb{E}\left[F(\delta_i, \theta_L) | \, s_I\right]$$

Note that the policy in question does matter in this proposition. Therefore, if this hypothesis is correct, we should expect a positive (negative) effect on the Remain (Leave) share of votes. If foreign players' performance is important for changing posterior beliefs, regardless of the players' absolute number or share in the squad, we should observe a positive effect when foreign players contribute to the team. I call this the payoff hypothesis, referring to the performance payoff that voters perceive from the presence of foreign players.

Last but not least, it is important to highlight two points. The first refers to the fact that as the effect of the signals is on the difference between prior and posterior beliefs, expectations are important. The second point is that if it is true that there is an update in preferences caused by the observed signals, this informational learning may persist over time. Therefore, if after updating their preferences, voters receive new signals, the new priors will be the former posteriors. In this case, there may be persistence in the effect over time. Thus, I now focus on describing the data and empirical strategy used first to test whether there are exposure and/or payoff effects and second to analyze the role of expectations and whether there is persistence over time in the learning process occasioned by the signals observed by voters.

# 3 Data

## 3.1 Premier League: teams, players and fans

To identify the distribution of Premier League fans within the UK, I use two sources of data. The rst is the number (percentage) of fans each team has at the district level, which comes from Twitter Analytics data, available from the platform's o cial website. Since the available data are for the 2014-15 season, the data for three small teams are missing, namely, the data for the teams (Bournemouth, Watford, and Norwich City) promoted in the 2014-15 season from the English Football League Championship (the second division of professional football in the UK) to the Premier League for the 2015-16 season. To cover the lack of data for these teams, I collect data from Google Trends, using the volume of associated searches and then normalizing the data for each team[12]. I collect data for each county (a political division smaller than the unit of analysis) and then collapse by district.

Table A1 in the Appendix shows the descriptive statistics for variables that represent fans by district (in percentage points) and region (stadium location) of each team. Figure 2 shows the distribution of the teams in the UK. As we can observe, there is geographical variation in the locations of teams, which are scattered across most of the UK. More importantly, Figure 3 shows, as an example, the distributions of the followers of the top four teams of the season: Leicester City, Arsenal, Tottenham Hotspur, and Manchester City. Here, it can be observed that each team's fans, although generally concentrated near the team's hometown, are distributed throughout the UK, and there is heterogeneity in the team a liations in each district. This can also be deduced from Table A1. This variation will be crucial for the identi cation strategy, as explained below.

[12] For this reason, the sum of the percentage of followers of all teams does not equal exactly 100%, but comes very close. I can con dently say that this does not bias the results

Figure 2: Location in the UK of each Football Club of the 2015-16 Premier League

For the analysis of the impact of players on voting, it is necessary to obtain data on all teams and each of their players for the 2015-16 and 2014-15 seasons. For this purpose, I use historical statistics available from the Fantasy League, Fantasy Football[13], Statbunker, Transfermarkt and API-Football websites. These websites provide information about each team and player for the season, such as the number of goals, red and yellow cards, assists and points, the latter being the main performance measure used for players in this work[14]. The nationality and ethnic origin of the players were collected manually using mainly Transfermarkt and complemented only minimally with information from the British or specialized press. These data are used to

[13] On these websites, users can pretend to be the team manager, and their performance depends on the performance of their players. The results of the players in the simulation correspond to the real results of the Premier League. These sites, in turn, use as their source www.optasports.com and www.statsperform.com, respectively

[14] Although this is a measure that could be biased towards forwards, it is the only one available to compare all players. Below, we will see that it does not a ect our results.

Figure 3: Example of the distribution of the Fans of a Football Club

determine players' origin and differentiate locals from foreigners. In addition, I obtain data on the league positions and points obtained by the different teams during the 2015-16 and 2014-15 seasons[15]. These data are used to perform robustness checks and test alternative hypotheses.

## 3.2 Electoral results

For the analysis of the impact of football on voting, I use district-level[16] data on vote shares in the Brexit referendum. I collect electoral administrative data from the website of the UK

[15] For the teams that were promoted from the Championship to the Premier League were attributed the last score and position before their descent.

[16] The actual name is Local Authority Districts.

government. I obtain information on the number of votes, electoral roll, and turnout for each of the 382 districts. These 382 districts are disaggregated as follows: 22 from Wales, 32 from Scotland, 1 from Northern Ireland, and 327 from England, with the latter being divided among nine regions. Following Becker, Fetzner, and Novy (2017) and Fetzner (2019), I drop the outliers, keeping 379 observations. The outliers are Northern Ireland, which, for the purposes of the referendum, was a single large district, being by far the largest in terms of size and population, as well as the Isle of Scilly and Gibraltar, which are very small and part of overseas territories. To control for previous trends in political preferences, I use, from the same source, the results of the 2014 European parliamentary elections. I show in Table A2 in the Appendix that these elections, despite not being directly comparable with the Brexit poll, can explain almost 92% of the variation in the referendum results. I also use data from the European parliamentary elections of 2004 and 2009 and from the 1975 UK European Communities membership referendum (obtained from the replication files of Becker, Fetzner and Novy, 2017) to conduct falsification tests.

### 3.3 Other variables

In all of the specifications, I also control for a wide range of variables at the district level. The full set of district-level covariates and their descriptive statistics are shown in Tables A3 and A4, while the controls chosen are shown in Table A5, all in Appendix A. I use the full set of covariates to conduct my robustness checks, using LASSO-chosen controls to ensure that there is no cherry-picking of the covariates employed (see Appendix B). The chosen controls are classified into 4 categories: (i) exposure to EU immigration, trade, and structural funds; (ii) level of local public service provision and amount of fiscal consolidation funds; (iii) demographic and education variables; and (iv) economic structure, wages, and unemployment. These categories are taken from Becker, Fetzner, and Novy (2017), who study the factors associated with the referendum results (albeit only in terms of correlations). The source of the data employed is again the replication files of these authors. Moreover, to construct different kinds of standard errors, I take the georeferenced coordinates of each district from the replication files of Fetzner (2019).

## 4 Empirical Strategy

Section II argues that voters may learn from foreign football players, which in turn affects their voting decisions. To unpack these political effects, I test my exposure and payoff hypotheses by running the following reduced-form regression:

$$y_{ir} = \beta Exposure_{ir} + \gamma Performance_{ir} + \delta' X_{ir} + \lambda_r + \epsilon_{ir}$$

where $i$ is each district and $y$ is the outcome variable. $Exposure$ and $Performance$ are the main regressors. The first is a measure of the level of exposure to foreign players in each district. The second is a measure of these players' performance. I am interested in the coefficients $\beta$ and $\gamma$, which indicate the impact of the exposure and performance hypotheses.

Estimating this reduced-form regression may not be enough, since the coe cients of interest may still be biased for multiple reasons even after I control for other variables. First, football clubs are not randomly assigned across the UK. If immigrant players are more expensive than local ones, whether a football club has them will depend on each team's purchasing power. This is presumably related to the economic characteristics of the district in which the team is located, in which case the error will be correlated with the treatment due to the presence of an omitted variable. Second, Premier League team fans are also not randomly assigned: one could argue that if the people of a district are more pro-immigration in their beliefs, they are more likely to support football clubs that have more immigrants in their squad. In this case, there would be a reverse causality problem, which would bias the estimation upwards. The reader will note that this could be solved by controlling for the distribution of the fans of each team in each district. However, this is not initially possible given the perfect collinearity that would arise with the main regressors. Furthermore, I still do not have a proper way to measure both exposure to and the payo of foreign players for each fan and to associate it with each voting district.

I thus propose the construction of exposure and performance treatment variables that may solve these issues. To do so, I explore the potentially exogenous source of variation given by the interaction between player/team characteristics with the spatial distribution of their fans across the UK. Therefore, I construct the main treatment variables by interacting (i) the spatial distribution of the fans of each football club and (ii) the nationality composition or performance of the players of that particular team for the corresponding season. Given the nature of the data, the treatment variables are constructed as follows.

To study the exposure e ect, I de ne:

$$ShareNonUK_i = \sum_{j=1}^{20} PctFans_{ij} \frac{\sum_{k_j=1}^{K_j} [1(k_j \in NonUK)]}{K_j}$$

where $ShareNonUK_i$ is the treatment variable, which captures the exposure to foreign players in district $i$; $PctFans_{ij}$ is the percentage of fans of football team $j$ [17] in district $i$; $k_j$ is player $k$ in team $j$; $K_j$ is the total number of players on team $j$; and $1(k_j \in NonUK)]$ is an indicator function that takes the value of 1 if a player is foreign and 0 if he is not[18].

[17] J = 20 because there are 20 teams in the Premier League

[18] For this work, I will consider a player to be local if at least one of his parents is a native of the UK. In the case that his parents are foreign, and the player is a UK citizen, I will consider the parents' original nationality. The main reason for this is that, for perception purposes, I believe said player represents more of an immigrant rather than a local. The idea of using players' origins comes from three considerations. First, it is a common phenomenon for players to opt for a second nationality in order to play for national teams. Second, Card et al. (2012) suggest that cultural and racial considerations drive perceptions of immigration. Thirdly, as the main argument is what the locals update belief on what they perceive as the payo for foreigners, I consider the arrival of a family already with citizenship is considered an immigrant family.

In sum, $\frac{\sum_{k_j=1}^{K_j}[\mathbf{1}(k_j \in NonUK)]}{K_j}$ represents the share of foreign players on team $j$. Hence, the treatment variable $ShareNonUK$ is the interaction between the share of foreign players on the team and the team's percentage of fans by district. This variable could be interpreted as the percentage of foreign players in a random voter's favorite football club in district $i$.

To study the possible payoff effect, I similarly define two other variables:

$$PointsUK_i = \sum_{j=1}^{20} PctFans_{ij} \times \frac{\sum_{k_j=1}^{K_j}[Points_{kj} \times \mathbf{1}(k_j \in UK)]}{\sum_{k_j=1}^{K_j}[\mathbf{1}(k_j \in UK)]}$$

$$PointsNonUK_i = \sum_{j=1}^{20} PctFans_{ij} \times \frac{\sum_{k_j=1}^{K_j}[Points_{kj} \times \mathbf{1}(k_j \in NonUK)]}{\sum_{k_j=1}^{K_j}[\mathbf{1}(k_j \in NonUK)]}$$

where $PointsUK_i$ and $PointsNonUK_i$ are treatment variables that capture the average performance of the players on team $j$ for district $i$ for locals (UK) and foreign players (Non-UK), respectively. $PctFans_{ij}$ is the percentage of fans of football team $j$ in district $i$, $k_j$ is player $k$ on team $j$, $K_j$ is the total number of players on team $j$, $\mathbf{1}(k_j \in UK)]$ is an indicator function that takes the value of 1 when a player is local and 0 if he is not, and $\mathbf{1}(k_j \in NonUK)]$ is an indicator function that takes the value of 1 when a player is a foreigner and 0 if he is not.

In sum, $\frac{\sum_{k_j=1}^{K_j}[Points_{kj} \times \mathbf{1}(k_j \in UK)]}{\sum_{k_j=1}^{K_j}[\mathbf{1}(k_j \in UK)]}$ represents the average performance of the local players on team $j$, while $\frac{\sum_{k_j=1}^{K_j}[Points_{kj} \times \mathbf{1}(k_j \in NonUK)]}{\sum_{k_j=1}^{K_j}[\mathbf{1}(k_j \in NonUK)]}$ represents the average performance of the foreign players on team $j$. Hence, the treatment variables $PointsUK$ and $PointsNonUK$ are interactions between the average performance of the players on a team and the percentage of fans of that team by district. This variable can be interpreted as representing the average performance of both local and foreign players on a random voter's favorite football club in district $i$. Another, similar interpretation of this variable is as the average payoff that fans perceive the football players on their teams to deliver.

It is important to note that there are several performance measures, but in the interest of selecting a representative measure for all players on a team, I use points. This measurement is chosen because (i) it is a variable that incorporates the most general characteristics and (ii) it is a better way to compare players across positions. Points are a measure of performance constructed as a linear combination of other observable measures of performance, such as goals, assists, fouls, and yellow and red cards, among others. Appendix C explains how this variable is constructed for each player position in the squad.

## 4.1 Main Specification Measuring Performance in Absolute Values

A first approach to the questions of whether voters are indeed changing their attitudes or behavior as a result of exposure to their local football players and, if players do indeed have such an impact, whether it is driven by an exposure or a payoff effect is as follows:

$$
\begin{aligned}
y_{ir} = \; & ShareNonUK_{ir} + \;_{UK} PointsUK_{ir} \\
& + \;_{NonUK} PointsNonUK_{ir} + X_{ir}^{0} \; + \;_{r} + \;_{ir}
\end{aligned}
$$

where $y_{ir}$ is the share of Remain votes in district $i$ in region $r$, $ShareNonUK_{ir}$ is the treatment variable measuring the impact of exposure to foreign players, and $PointsUK_{ir}$ and $PointsNonUK_{ir}$ are the treatment variables measuring the impact of the performance of local and foreign players, respectively. $X_{ir}^{0}$ is a vector of district-level covariates, $_{r}$ represents region fixed effects and, finally, $_{ir}$ is a heteroskedasticity- and spatial autocorrelation-robust error term calculated assuming spatial autocorrelation within 50 km, following Conley (1999) and Colella et al. (2019)[19]. Thus, I identify the effect of exposure and performance through within-region variation.

For the answer to the empirical question, the coefficients of interest are , $_{UK}$, and $_{NonUK}$. represents the impact of the share of or exposure to foreign players. $_{UK}$ and $_{NonUK}$ represent the impact of players' performance on their fans. If there is an exposure effect, then the coefficient should be different from zero. The same logic applies to the analysis of the payoff hypothesis. The coefficients of interest are $_{UK}$ and $_{NonUK}$, representing the impact of the performance of players, both local and foreign.

For the identification of the effect of performance, what matters is the coefficients' statistical significance, sign, the difference between them and significance of that difference. That is, if both coefficients are not significantly different and positive, then the effect captured does not necessarily refer to an update of information through the observation of foreigners but rather to another effect, which could be the happiness associated with having good players on the team a person is a fan of or to the fact that the team is doing well in a given season.

## 4.2 Main Specification Measuring Performance in Relative Values

Following the same logic, if voters update their beliefs because of foreign players' performance relative to that of local players, it is to be expected that $_{UK}$ and $_{NonUK}$ have opposite signs. If this is the case and $_{UK}$ + $_{UK}$ = 0, i.e., are statistically equal in absolute value, it is possible to measure the impact of relative performance or payoff as follows:

$$Points_{ir} = PointsNonUK_{ir} \quad PointsUK_{ir}$$

[19] All results are robust to assume correlation within 25, 50 and 100 km, and heteroskedasticity-robust as well.

Thus, a second approach to studying whether voters are indeed changing their attitudes or behavior because of their local football players and, if players do indeed have such an impact, whether it is driven by exposure or a perceived payoff is as follows:

$$y_{ir} = \text{ShareNonUK}_{ir} + \quad \text{Points}_{ir} + X_{ir}^{0} \quad + \quad {}_{r} + \quad {}_{ir}$$

To correctly identify causal effects, the key identifying assumption of the strategy is that $corr(T_{ir};\ {}_{ir}) = 0$, where $T$ represents the constructed treatments. From the construction of the treatment variables, it is clear that the source of variation comes from the interaction between the baseline distribution of fans across the UK and the results and characteristics of each team player in the season. Thus, the first source of variation for the treatment variables is at the district level, given by the distribution of fans in district $i$. The second source is at the level of team $j$, given by its characteristics, and at the level of player $k$.

The key to the plausibility of the exogeneity of the treatment is in the interaction between these sources of variation. The treatment variables are constructed so that the interaction represents how strongly the treatment affects each district. Since the variation in the treatments should be unrelated to events in a particular district, there is no apparent reason for it to be correlated with the error term, making it possible to identify causal effects.

This approach is based on the empirical strategy used by Acemoglu and Robinson (2006, 2007). In this work, they construct an instrument that correctly identifies the effect of improvements in life expectancy on economic development. Even though they construct their instrument for use with an IV design, I use the same argument that they use to satisfy the exclusion restriction in the construction of the independent variable for the reduced-form regression. While a shock may be common to everyone, how it is received differs from district to district, conditional on cross-district covariates. Acemoglu and Robinson (2006, 2007) also use time fixed-effects, but as I have cross-district data only, I cannot control for variation over time. However, I do have data variation at the team level.

For the identifying assumption to be met, some requirements must be fulfilled. For $\text{ShareNonUK}$, the distribution of fans must not be correlated with team characteristics, with the identifying assumption being slightly stronger if I do not control for the distribution of each team's fans in each district[20]. However, for the identification of the relative performance effect, the key is that the interaction between the difference in average performance between foreigners and locals with the spatial distribution of fans is as good as random conditional on the district-level covariates. Since there is no apparent reason to relate the difference in performance between foreigners and locals in each team for a given season to particular events and since the interaction with the

[20] The reader will remember that this is not initially possible given the perfect collinearity that will be present with the main regressors. However, in the discussion section, I show that under certain conditions, I can run the main specification controlling by fans-fixed-effect and that the results do not change.

fan distribution captures the intensity with which each district perceives this difference, there is no apparent reason for $Points$ to be correlated with the error term, making it possible to identify causal effects.

Therefore, if we are willing to accept these identifying assumptions, the results can be properly interpreted as causal effects on the electoral outcome. The interpretation of the results corresponds to a reduced-form estimation, capturing all the effects of the interaction. To validate the identifying assumptions, I conduct a falsification test with data from previous elections, as discussed in detail in the discussion section.

# 5 Results

## 5.1 Unpacking Effects of Exposure and Performance Signals

I begin in Table 1 by testing only the effect of exposure to foreign players ($ShareNonUK$) on the Remain share of votes. To simplify interpretation, all independent variables are standardized. The purpose of this table is to reveal how the set of controls affects the treatment and what happens if we omit them. To start, I only use $ShareNonUK$ since its interpretation is easier to understand than that of the set of covariates. This variable could be interpreted as the percentage of foreign players in a random voter's favorite football club in district $i$.

I start in column (1) with the naive regression without controls. Even though there is a positive correlation between exposure and the share of Remain votes, it is not statistically significant. Then, I control for prepolitical trends, i.e., the UKIP vote share in the European parliamentary elections of 2014. Including this variable allows me to study the effect on voter behavior conditional on 2014 political preferences. Column (2) shows that this improves the estimator efficiency, and the estimation's $R^2$ increases significantly. This is to be expected since this prior election explains most of the variation in the data[21].

Across columns (3) to (6), I continue to add different controls that, if omitted, could bias the result (all of which are described in Table A5 in the Appendix). In column (3), I control for EU exposure and immigration in each district, and we can observe that the coefficient drops markedly in value. Hence, omitting these variables would bias the estimator. This could be because there is a potential correlation between EU exposure and immigration and the football teams' fan bases that the estimator was previously capturing. The same occurs when we incorporate local and public financial services in column (4).

[21] For the variance decomposition, see Table A2 in the Appendix, where we can observe that the votes cast for UKIP towards votes cast for Leave are near 1 to 1.

Table 1: Main Speci cation - The E ect of Exposure To Foreign Players on Brexit

Share non-UK is sum of the foreign share of players by team, interacted with the percentage of fans by team at district-level. Share non-UK is standardized. UKIP Share of Votes is the 2014 EP Election.

| | Dependent Variable: Remain Share of Votes | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Share Non-UK | 1.440 | 1.923 | 1.288 | 0.613 | 0.892 | 0.772 | 0.442 |
| | (1.084) | (0.443) | (0.411) | (0.266) | (0.259) | (0.251) | (0.219) |
| Controls | | | | | | | |
| UKIP 2014 Vote Share | | -0.974 | -0.906 | -0.898 | -0.581 | -0.578 | -0.633 |
| | | (0.0467) | (0.0461) | (0.0470) | (0.0602) | (0.0571) | (0.0501) |
| EU exposure and immigration | No | No | Yes | Yes | Yes | Yes | Yes |
| Local public and  scal services | No | No | No | Yes | Yes | Yes | Yes |
| Demographics and education | No | No | No | No | Yes | Yes | Yes |
| Economic structure | No | No | No | No | No | Yes | Yes |
| Region F.E. | No | No | No | No | No | No | Yes |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 | 47.087 | 47.087 | 47.087 | 47.087 |
| Observations | 378 | 378 | 378 | 375 | 375 | 374 | 374 |
| $R^2$ | 0.019 | 0.774 | 0.837 | 0.872 | 0.922 | 0.933 | 0.945 |

Conley (1999) standard errors in parentheses, assuming correlation within 50 km.

In columns (5) and (6), I add variables on demographics and education and economic structures. The demographic and education controls are important because if we think that voters might be updating their signals, the effect should differ depending on how close their priors were to their posteriors. On the other hand, with the economic structure controls, I make sure that the treatment is not capturing each district's economic wealth through the composition of its team's fan base.

Finally, in column (7), I add region fixed effects, which allow a within-region interpretation of the effects. This is important since within the UK, both political preferences and the fan bases of football teams are rather local. Because of this, I also use spatial correlation-robust standard errors[22]. For this table, column (7) is the preferred specification. I estimate that the political impact of an increase of one standard deviation in exposure to foreign players corresponds to an increase of 0.443 percentage points, which is statistically significant at 10%. However, we must be cautious in interpreting this result causally, mainly because the variables that measure player performance are omitted. If these variables are correlated with $ShareNonUK$, the specification will be capturing other effects.

In Table 2, I try to unpack the effects of both exposure to and the performance of players on the referendum results. Column (1) is the same as column (7) in Table 1 and is used as a benchmark. In column (2), I run the test with a performance measurement for the first time, adding $PointsNonUK$ to the specification. Note that the $ShareNonUK$ estimator drops significantly, changes sign and loses its statistical significance. This suggests that the relation found in column (1) is also capturing the effect of foreign players' performance as an omitted variable. Next, I add $PointsUK$ in column (3). Both the non-UK share and non-UK points estimators rise, indicating that these were capturing the effects of omitting $PointsUK$.

If we move on to analyze the performance measurements for locals and foreigners, we learn several things. First, we can observe that both are economically and statistically significant. Thus, we can identify that the effect does not depend on the performance of players as a whole but on their origins. Second, the results suggest that voters compare foreign players to local players. Hence, players' performance in relative rather than absolute terms is what appears to relevant in this effect. Moreover, the coefficients of performance are not statistically different from each other in absolute value, which allows us to estimate the relative performance of foreigners over locals using $\Delta Points = PointsNonUK - PointsUK$. This makes the interpretation simple and allows us to obtain more efficient estimators.

Column (4) shows the outcome of this estimation. I estimate the political impact of exposure to foreign players as well as the impact of foreign players' better performance relative to that of locals. Using the theoretical framework discussed previously, we learn that the extensive signal $s_E$, that is, the number of foreigners on the local team, does not seem to be of relevance or

[22] All results are robust to assuming correlation within 25, 50, and 100 km and are heteroskedasticity-robust as well.

generate any change whatsoever in voters' posteriors. However, the intensive signal observed by voters does modify their posteriors. That said, this signal is not constituted simply by the absolute performance of foreigners. Voter seems to use the performance of local players as a benchmark. Thus, voters learn from foreign players' better relative performance. In sum, if we are willing to accept the identifying assumption, we nd that an improvement of one standard deviation in foreign players' relative performance explains a cross-district increase of 0.927 percentage points in the Remain vote share.

Table 2: Main Speci cation - The E ect of Exposure and Performance of Foreign Players on Brexit

Share non-UK is sum of the foreign share of players by team, interacted with the percentage of fans by team at district-level. Points non-UK and Points UK are the foreign and UK average points reached by players by team, interacted with the percentage of fans by team at district-level. Points = Points non-UK - Points UK, i.e. the relative points of foreigners to locals. All independent variables are standardized.

| | Dependent Variable: Remain Share of Votes in 2016 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share non-UK | 0.442 | -0.228 | 0.115 | -0.288 |
| | (0.219) | (0.439) | (0.426) | (0.372) |
| Points non-UK | | 0.828 | 1.465 | |
| | | (0.390) | (0.403) | |
| Points UK | | | -1.223 | |
| | | | (0.390) | |
| Points | | | | 0.927 |
| | | | | (0.317) |
| Controls | | | | |
| Full Set of Controls | Yes | Yes | Yes | Yes |
| UKIP Share of Votes | Yes | Yes | Yes | Yes |
| Region F.E. | Yes | Yes | Yes | Yes |
| [Prob > Chi $^2$] | | | | |
| Points non-UK + Point UK = 0 | | | [0.572] | |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 | 47.087 |
| Observations | 374 | 374 | 374 | 374 |
| $R^2$ | 0.945 | 0.946 | 0.948 | 0.947 |

Conley (1999) standard errors in parentheses, assuming correlation within 50 km.

p-value in brackets.

So far, we have learned several things. First, it is important to control for covariates so as not to bias the results. Second, there seems to be no evidence supporting the exposure hypothesis. This suggests that there is no learning on the part of voters from being exposed to foreign players. Conversely, there is evidence suggesting an effect generated by the performance of foreign players, perceived as the relative payoff of foreigners' contribution to a squad's results. This also suggests that voters take local players as the comparison group.

Two questions arise naturally from the above. The first is about voters' learning process triggered by the relative performance signal. If this learning occurs in every football season, then the signals given by the players in the previous season may also have impacted the election results. If there is persistence in the identified effect, the constructed treatment could be capturing the previous season's effect, biasing the results. Second, the fact that foreign players' performance is evaluated in relation to that of locals suggests that voters have expectations for their team's players. Therefore, it is important to control not only for expectations using local players as benchmarks but also for the expectations that voters have for each player at the beginning of the season. Thus, in the next subsection, I turn my attention to studying the role of expectations over time and the possible presence of persistence in the identified effect.

### 5.2 Do Expectations and Persistence Matter?

To study the relevance of expectations over time and the potential persistence of the identified effect, I first add three new variables to the main specification. The first two are the treatments, $ShareNonUK$ and $\Delta Points$, constructed for the immediately previous season, that is, 2014-15. The estimators of these variables will indicate whether there is persistence in the discussed effects. Then, I use the relative price (market value) of foreign players at the beginning of the season as a measurement of expectations. If foreign players' average price at the beginning of the season is far higher than local players' price, it is reasonable to assume that voters expect better relative performance from the former. The relevance of this variable is also as a proxy for voters' prior belief on players' future performance. For this, I create the variable $\Delta Players\ Market\ Value = NonUK\ Market\ Value - UK\ Market\ Value$, which will be a proxy for expectations over time. Like this, the main specification becomes:

$$
\begin{aligned}
y_{ir} &= \beta\, ShareNonUK_{t;ir} + \gamma\, \Delta Points_{t;ir} \\
&+ \beta_2 ShareNonUK_{t-1;ir} + \gamma_2\, \Delta Points_{t-1;ir} + \delta\, \Delta Players\ Market\ Value_{ir} \\
&+ X'_{ir}\theta + \alpha_r + \varepsilon_{ir}
\end{aligned}
$$

The results are shown in Table 3. As above, I begin in column (1) with the main specification as a benchmark.

Table 3: Alternative Specification - The Role of Inertia and Expectations on The Effect of Exposure and Performance of Foreign Players on Brexit

Share non-UK is sum of the foreign share of players by team, interacted with the percentage of fans by team at district-level. $\Delta Points =$ Points non-UK - Points UK, i.e. the relative points of foreigners to locals. Variables in $\in t-1$ are the same variables constructed for the previous season. $\Delta$ $Players\ Market\ Value$ is the relative market value of foreign towards players at the beginning of the season. All independent variables are standardized.

| | Dependent Variable: Remain Share of Votes in 2016 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Share non-UK | -0.288 | -0.358 | -0.635 | -0.511 |
| | (0.372) | (0.352) | (0.355) | (0.355) |
| $\Delta$ Points | 0.927 | 0.681 | 1.146 | 0.779 |
| | (0.317) | (0.373) | (0.336) | (0.388) |
| Lags and Expectations Controls | | | | |
| Share non-UK $\in t-1$ | | 0.683 | | 0.454 |
| | | (0.233) | | (0.238) |
| $\Delta$ Points $\in t-1$ | | 0.617 | | 0.674 |
| | | (0.384) | | (0.381) |
| $\Delta$ Players' Market Value | | | 0.509 | 0.346 |
| | | | (0.272) | (0.310) |
| Controls | | | | |
| Full Set of Controls | Yes | Yes | Yes | Yes |
| UKIP Share of Votes | Yes | Yes | Yes | Yes |
| Region F.E. | Yes | Yes | Yes | Yes |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 | 47.087 |
| Observations | 374 | 374 | 374 | 374 |
| $R^2$ | 0.947 | 0.948 | 0.948 | 0.949 |

Conley (1999) standard errors in parentheses, assuming correlation within 50 km.

In column (2), I study whether there is persistence in the effect, adding to the main specification the treatment variables constructed for the previous season. First, I find that the lagged variables for both margins are positive and statistically significant. This suggests that there is persistence and that it is considerable. Moreover, we see that the estimator of $\Delta Points$ decreases to 0.681. This suggests two things. First, there is persistence in voters' learning process since $\Delta Points_{t-1}$ is economically and statistically significant. Second, as the point estimate decreases, there is persistence at the team level. This is because most of the players on each team do not change from season to season. Therefore, the point estimate for the 2015-16 season is capturing part of the effect of the 2014-15 season, biasing the result upward.

In column (3), I study how the treatment effect changes when I control for expectations, measured as the relative market value of foreign players. We can observe that the point estimate of $\Delta Points$ slightly increases. The economic interpretation is that if the effect is larger when there is a relative performance surprise and if I do not control for voters' expectations, this surprise factor will be captured by $\Delta Points$. Omitting expectations over time biases the point estimate slightly downward, suggesting that while this variable matters, performance relative to that of locals is far more important.

Finally, in column (4), I control for both lags and the relative market value. As expected, $\Delta Points$ increases relative to the estimate in column (2). Additionally, adding players' market value hardly biases $\Delta Points_{t-1}$ at all since that is a proxy for the 2015-16 season. This confirms what we learned above: that there is persistence in the identified effect and that expectations matter but are not crucial.

Up to this point, we have learned several things. I have shown evidence that there is a causal effect of football, specifically football players' performance, on a political election. Conversely, there seems to be no evidence supporting an effect associated with exposure to foreign players on the election results. The main effect comes from the intensive margin only, that is, from the performance of foreign players benchmarked against that of locals. Voters also have certain expectations on player performance at the beginning of the season. However, these are much less important than the relative difference in performance with that of locals over the season. Last but not least, we learn that this effect persists over time, arising from the learning process of voters and the relative stability of teams over time.

Since there is persistence in the effect, I take a weighted average of the treatments to analyze the effects. To do so, I estimate the main specification but using averages. That is, I construct

$$\text{Average } ShareNonUK = \frac{ShareNonUK_t + ShareNonUK_{t-1}}{2}$$

and,

$$\text{Average } \Delta Points = \frac{\Delta Points_t + \Delta Points_{t-1}}{2}$$

as new measurements of exposure and performance, for then to estimate the main specification.

Table 4: Alternatives Speci cation - The E ect of Exposure and Performance of Foreign Players on Brexit Assuming Inertia in The Treatment

Average Share non-UK is the average between Share non-UK and Share non-UK t 1[2]. Points non-UK and Points UK are the foreign and UK average points reached by players by team, interacted with the percentage of fans by team at district-level. Points = Points non-UK - Points UK, i.e. the relative points of foreigners to locals. All independent variables are standardized.

| | Dependent Variable: Remain Share of Votes in 2016 | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Average Share non-UK | 0.525 | | 0.0854 | -0.218 |
| | (0.213) | | (0.257) | (0.265) |
| Average Points | | 0.797 | 0.755 | 0.917 |
| | | (0.187) | (0.249) | (0.264) |
| Players' Market Value | | | | 0.469 |
| | | | | (0.316) |
| Controls | | | | |
| Full Set of Controls | Yes | Yes | Yes | Yes |
| UKIP Share of Votes | Yes | Yes | Yes | Yes |
| Region F.E. | Yes | Yes | Yes | Yes |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 | 47.087 |
| Observations | 374 | 374 | 374 | 374 |
| $R^2$ | 0.945 | 0.948 | 0.948 | 0.948 |

Conley (1999) standard errors in parentheses, assuming correlation within 50 km.

Table 4 shows these results. In column (1), I look at the effect of $ShareNonUK$. Column (2) estimates the impact of $Points$ with the exposure variable omitted. Column (3) shows the main specification. Note that the estimation results lead us to the same conclusion as that reached previously in Table 2. All the effect that $ShareNonUK$ might be capturing is gone when we measure performance using $Points$, which, again, confirms that it is relative performance rather than simply exposure to foreign players that makes voters update their preferences.

Note in columns (2) and (3) that this exposure effect appears to be just noise, since the point estimator of $Points$ does not change when the variable is omitted and the estimator efficiency improves. This will be important because to control for the distribution of teams' fans, I will have to omit this variable due to perfect collinearity. Thus, when I further control for the latter variable, I can be confident that I am not leaving out any variables that could change the results.

Finally, in column (4), I add the players' market value. Once again, the results remain the same: omitting this variable biases the $Points$ point estimate slightly downward. In sum, a relative improvement of one standard deviation in foreign players' performance boosts the Remain vote share by 0.917 percentage points. We come to essentially the same conclusions as those drawn from Table 2 but with further details now on the process of voter learning, the persistence in time of the identified effect, and the low relevance of expectations over time.

# 6 Discussion

## 6.1 Identification Assumption and Falsification Test

Since I found that the main mechanism through which voters update their preferences operates at the intensive margin, I turn to discuss this assumption in more detail in this section. The empirical strategy relies on $corr(Points_{ir}; e_{ir} | X') = 0$. It is not possible to test this directly. However, it is a recurrent practice in the literature to perform a falsification or placebo test to validate the assumption. To do so, I argue that due to the treatment timing, I can run a placebo test with the results of prior elections. Conditional on the covariates of the main specification, if the treatment is indeed exogenous, I should find that there is no significant effect of the treatment on the results of prior elections in a reduced-form regression. This would rule out the possibility of a correlation of the treatment with potentially unobservable variables excluded from the regression.

Given the nature of the Brexit referendum, it is not easy to find a directly comparable election. The closest approximation that can be employed is the 2014 European parliamentary elections, which I use as a control for prepolitical trends. For this reason, I test the performance hypothesis using three alternatives: the 1975 EU referendum and the European parliamentary elections of 2004 and 2009.

The results are shown in Table 5. In this table, dependent and independent variables are standardized to simplify the comparison. Column (1) corresponds to the benchmark, i.e., using the Remain vote shares in the Brexit referendum as the dependent variable. Column (2) presents the main specification but uses the Leave vote shares in the 1975 referendum as the dependent variable. Columns (3) and (4) use as dependent variable the share of the votes for UKIP in the 2004 and 2009 European elections, respectively. The different specifications in the panels correspond to those used in Table 4.

The first important thing to notice is that for column (1), all results are statistically significant and larger in magnitude in each of their respective rows. Panel A presents only $Points$, omitting $ShareNonUK$ since as we discussed previously, its effect is null and only seems to add noise to the estimation. The results are encouraging, showing an effect that is statistically equal to zero for the three falsification tests. In case of further doubt, in Panel B, I control for expectations and $ShareNonUK$, as in column (4) of Table 5, and, once again, the previous result hold, even though the point estimations for the UKIP vote share show a slight increase. However, they remain statistically nonsignificant (equal to zero). Lastly, Panel C presents the same estimates as in Panel A but with controls for expectations, with the same result. Thus, it seems plausible that $corr(Points_i; e_i) = 0$. Therefore, the previous results can be interpreted causally.

One caveat is that while in Panel B, the coefficients for UKIP for 2004 and 2009 are statistically not different from zero, their magnitudes are higher than those of the other coefficients. This could be because the treatment is still capturing some omitted variables, for example, the distribution of the fans across the UK. It is not possible to test all specifications controlling for the distribution of fans. The reason is that this could generate problems since the treatments are linear combinations of this distribution. However, it remains to be shown that this is not what the full effect is capturing; I do this in the subsection below.

Table 5: Falsification Test - The Effect of Treatment on Former Elections
Average Points the average between Points and Points 2 t 1. Points non-UK and Points UK are the foreign and UK average points reached by players by team, interacted with the percentage of fans by team at district-level. Points = Points non-UK - Points UK, i.e. the relative points of foreigners to locals. All independent and dependent variables are standardized.

| | | Falsification Tests | | |
|---|---|---|---|---|
| | Remain 2016 | Leave 1975 | UKIP 2004 | UKIP 2009 |
| | (1) | (2) | (3) | (4) |
| | Panel A: Average Delta Points | | | |
| Average Points | 0.0766 | 0.00927 | 0.00608 | 0.00735 |
| | (0.0180) | (0.0376) | (0.0236) | (0.0197) |
| | Panel B: Share Non-UK and Market Value Controls | | | |
| Average Points | 0.0881 | 0.00991 | 0.0546 | 0.0329 |
| | (0.0254) | (0.0542) | (0.0343) | (0.0270) |
| | Panel C: Market Value Controls | | | |
| Average Points | 0.0774 | 0.00946 | 0.00409 | 0.00609 |
| | (0.0180) | (0.0375) | (0.0232) | (0.0200) |
| Controls | | | | |
| Full Set of Controls | Yes | Yes | Yes | Yes |
| UKIP Share of Votes | Yes | Yes | Yes | Yes |
| Region F.E. | Yes | Yes | Yes | Yes |
| Observations | 374 | 374 | 374 | 374 |

Conley (1999) standard errors in parentheses, assuming correlation within 50 km.

## 6.2 Spatial Distribution of Fans

Even though I already control for pretreatment political trends, preferences, and a full set of district-level covariates in my main specification, another valid concern could be that the characteristics of a specific team's fans could be the reason why the treatment is capturing an effect. For example, if one team had a promigrant player policy, this could result in a different effect of foreign players being better (i.e., of the $PointsNonUK$ variable) in that team than in other teams. Therefore, the treatment would be capturing other effects instead of the one studied, biasing the results. As stated above, it is not trivial to control for the distribution of fans in the main reduced-form specification due to collinearity issues. Nevertheless, it is possible to run some exercises to mitigate this concern.

In Table 6, I present four different alternatives to control for the distribution of fans in each district. First, note that I use the specification without including either $ShareNonUK$ or the market value of the players. The reason for this is that as they do not signi cantly bias P oints and omitting these variables delivers more degrees of freedom so as not to saturate the regression with multicollinearity. Therefore, column (1) presents the benchmark, and then, in columns (2) to (4), I add di erent combinations of teams. Column (5) includes all teams, taking Liverpool as the baseline since it is the team with the most fans in the data.

We can observe that none of these speci cations markedly biases the results. Hence, the main result continues to be robust, and the estimators do not seem to lose e ciency. Although spec- i cation (5) seems to be central since it allows us to implement a kind of team xed e ect, analyzing this result in the end allows us to dismiss evidence in favor of the exposure hypothesis and to analyze the e ect of players' relative market value at the beginning of the season. Obviously, if the results were di erent, the interpretation would be di erent. However, this exercise should reassure the reader that not including these controls at the outset does not bias the results.

Table 6: Robustness Checks - Controlling for Di erent Team Fans Spatial-Distribution
Average Points the average between Points and Points 2 t 1. Points non-UK and Points UK are the foreign and UK average points reached by players by team, interacted with the percentage of fans by team at district-level. P oints = Points non-UK - Points UK, i.e. the relative points of foreigners to locals. Average Points is standardized. Big 4 and Big 6 are the 4 and 6 biggest teams in fans share at district level. Top 7 are the are the best 7 of the 2015-16 season. All teams omits Liverpool due collinearity.

| | Dependent Variable: Remain Share of Votes in 2016 | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Average Points | 0.797 | 0.680 | 0.613 | 0.766 | 0.952 |
| | (0.187) | (0.182) | (0.159) | (0.176) | (0.225) |
| Spatial Distribution of Fans | No | Big 4 | Big 6 | Top 7 | All Teams |
| Controls | | | | | |
| Full Set of Controls | Yes | Yes | Yes | Yes | Yes |
| UKIP Share of Votes | Yes | Yes | Yes | Yes | Yes |
| Region F.E. | Yes | Yes | Yes | Yes | Yes |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 | 47.087 | 47.087 |
| Observations | 374 | 374 | 374 | 374 | 374 |

Conley (1999) standard errors in parentheses, assuming correlation within 50km.

## 6.3 Other Hypotheses

Another valid concern is that each team's overall results could also be correlated with the Brexit referendum results and the treatment variable. It is natural to think that if the players on a given team are better than those on other teams (i.e., on average have more points), it is more likely that said team achieved a better result that season.

This argument is also related to the shared experiences literature[23]. Suppose the foreign players on a given team are performing very well. This makes the team perform better that season, which might affect the marginal voter because the team's fans are happier and hence they have different attitudes. In this case, the treatment variables Points, especially in foreign player-intensive teams, could capture the effect of a good or bad season for a given team on the political decisions of that team's fans. If so, this will bias the point estimate of interest.

In Table 7, I control for these characteristics, specifically the absolute success of each team in the Premier League as well as its success relative to the past season. Once again, I show that the main results are robust to the use of these controls. If we look at the Points coefficient, it remains stable. In case the reader still has doubts, Appendix B tests this hypothesis directly, showing no evidence that this mechanism is operative.

[23] See, for example, Depetris-Chauvin, Durante and Campante (2019).

Table 7: Robustness Checks - Controlling for Di erent Team Fans Spatial-Distribution Average Points the average between Points and Points 2 t 1. Points non-UK and Points UK are the foreign and UK average points reached by players by team, interacted with the percentage of fans by team at district-level. P oints = Points non-UK - Points UK, i.e. the relative points of foreigners to locals. Average Points is standardized. 2015-16 Season Position is the position of the teams in the season, and Di erence is relative to the past season.

| | Dependent Variable: | | | |
|---|---|---|---|---|
| | Remain Share of Votes in 2016 | | | |
| | (1) | (2) | (3) | (4) |
| Average Points | 0.797 | 0.821 | 0.828 | 0.774 |
| | (0.187) | (0.201) | (0.195) | (0.209) |
| 2015-16 Season Position | No | Yes | No | Yes |
| Di erence in Position With Last Season | No | No | Yes | Yes |
| Controls | | | | |
| Full Set of Controls | Yes | Yes | Yes | Yes |
| UKIP Share of Votes | Yes | Yes | No | No |
| Region F.E. | Yes | Yes | Yes | Yes |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 | 47.087 |
| Observations | 374 | 374 | 374 | 374 |

Conley (1999) standard errors in parentheses, calculated assuming correlation within 50 km.

## 7 Conclusion

Few things awaken emotions like sports, and emotions, in turn, have been shown to have an impact on political processes. Arguing that voters can update their beliefs by observing football players, this paper studies whether voter exposure to and the performance of foreign football players can a ect a political election. Using information on Premier League team nationality compositions, player performance, and geographic distributions of fan bases on the eve of the Brexit referendum, I provide new evidence of the e ects of sports on political processes. This paper shows that an improvement in the performance of foreign football players relative to local players makes voters update their beliefs and preferences, a ecting their voting decision.

I nd that a one-standard-deviation increase in the relative performance of foreign players boosted the within-region Remain vote share by approximately 1 percentage point at the district level. This corresponds to approximately 1/2 of the percentage points of the vote share that the Remain campaign would have needed to win the referendum. Conversely, I do not nd evidence that the relative number of foreign players explains cross-district di erences in voting. This set of results suggests that what matters is not mere exposure to foreign players but rather their relative performance, with voters seeming to use local players as the reference group within the season.

Consistent with a signal-based learning model, I also find that the identified effect persists over time. This is interpreted as the relative differences in performance from the previous season also affecting the election. In other words, updates to preferences in the past make voters' former posteriors their new priors, thus restarting the learning process.

Last but not least, this work shows the importance of people's expectations. Voters interpret foreigners' performance using local players as their benchmark. Additionally, using the market value of players at the beginning of the season, I show that voters have certain expectations on player performance at the beginning of the season. However, these are much less important than the relative difference in performance with locals during the season.

To conclude, the findings of this paper contribute to (i) the understanding of novel mechanisms whereby emotional cues operate on political processes and voter decisions, (ii) the learning processes of voters due to the reception of new signals, and (iii) the relative importance of expectations in this process.

# 8 References

[1] Acemoglu, D. S. Johnson (2006) Disease and Development: The E ect of Life Expectancy on Economic Growth. Working Paper. NBER.

[2] Acemoglu, D. S. Johnson (2007) Disease and Development: The E ect of Life Expectancy on Economic Growth. Journal of Political Economy 2007 115:6, 925-985

[3] Allport, G. (1979).The Nature of Prejudice. 25 ed. Cambridge, Massachusetts: Perseus BooksPublishing, L.L.C

[4] Alrababa'h, A., Marble, W., Mousa, S., & A.A. Siegel (2019). Can Exposure to Celebrities Reduce Prejudice? The E ect of Mohamed Salah on Islamophobic Behaviors and Attitudes. Working Paper.

[5] Aris, E., Larraguy, H., Marshall, J. & P. Querubin (2018). Priors rule: When do Malfeasance Revelations Help or Hurt Incumbent Parties? NBER.

[6] Brader, T. (2005) Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions.American Journal of Political Science.

[7] Becker, S., Fetzer, T. & D. Novy (2017). Who Voted For Brexit? A Comprehensive District-level Analysis. Economic Policy.

[8] Campante, F., Depetris-Chauvin, E., & R. Durante (2019).The Virus Of Fear: The Political Impact Of Eebola In The U.S. Working Paper. NBER.

[9] Card, D. & B. Dahl, (2011). Family Violence And Football: The E ect Of Unexpected Emotional Cues Ov Violent Behavior. The Quaterly Journal Of Economics.

[10] Card, D., Dustmann, C. & I. Preston (2012). Immigration, Wages, and Compositional Amenities. Journal Of The European Economic Association

[11] Colella, F., Lalive, R., Sakalli, S. O., & M. Thoenig. Inference with Arbitrary Clustering. Working Paper. IZA, Institute Of Labor Economics.

[12] Conley, T. (1999). Gmm estimation with cross sectional dependenceJournal Of Econometrics

[13] Clingingsmith, D., Khwaja A, I., and M. Kremer (2009). Estimating the Impact of the Hajj: Religion and Tolerance in Islam's Global Gathering. Quarterly Journal of Economics.

[14] Gantz, W., Bradley, S., and Z. Wang (2006). Televised NFL Games, the Family, and Domestic Violence. Handbook of Sports and Media

[15] Depetris-Chauvin, E., Durante, R., & F. Campante (2020). Building Nations Through Shared Experiences: Evidence from African Football. American Economic Review.

[16] Dustmann, C., Fabbri, F. & I. Preston (2005). The Impact of Immigration on the British Labour Market. The Economic Journal.

[17] Dustmann, C. & I. Preston (2007). Racial and Economic Factors in Attitudes to Immigration. The B.E. Journal Of Economic Analysis & Policy.

[18] Fetzner, T. (2019). Dis Austerity Cause Brexit? American Economic Review.

[19] Fantasy Premier. https://fantasy.premierleague.com/statistics

[20] Gauriot, R. & Page, L., (2018). Fooled by Randomness: Over-Rewarding Luck. Working Paper

[21] Gietel-Basten, S. (2016). Why Brexit? The Toxic of Immigration and Austerity. Population and Development Review.

[22] Google Trends. https://trends.google.es/trends/?geo=ES

[23] Kaplan, E., and S. Mukand (2014). The Persistence of Political Partisanship: Evidence from 9/11. Unpublished

[24] Kuper, S., and S. Szymanski (2010). Soccerconomics.

[25] Larreboure, M. & F. Gonzalez (2020). The Impact of Women's March on The U.S. House Election. Working Paper.

[26] Madestam, A., and D. Yanagizawa-Drott (2011). Shaping the Nation: The E ect of Fourth

of July on Political Preferences and Behavior in the United States.Unpublished

[27] ONS. O ce of National Statistics. https://www.ons.gov.uk/

[28] Rees, D., and K. T. Schnepel (2209). College Football Games and Crime.Journal of Sports Economics.

[29] Statbunker https://www.statbunker.com/

[30] Statista. https://www.statista.com/topics/1773/premier-league

[31] Take Back Control. https://www.takebackcontrol.com/

[32] The Brexit Colletion, LSE. https://digital.library.lse.ac.uk/collections/brexit

[33] Transfermarket. https://www.transfermarkt.com/

[34] Twitter Analitycs. https://interactive.twitter.com/premierleague/?mode=teamteamId=all

[35] Viskanic, M. (2017). Fear and Loathing on the Campaign Trail: Did Immigration Cause-Brexit? Working Paper.

[36] Wadsworth, J., Dhingra, S., Ottaviano, G. & J. Van Reenen (2016). Why Immigration Is No Reason To Leave The EU.The Magazine for Economic Performance, LSE.

# A Appendix: Other Tables

Table A.1: Descriptive statistics - 2015-2016 Premier League Teams

| Pos | Team | Region | Pct Fans | | | |
|---|---|---|---|---|---|---|
| | | | Mean | Std. Dev. | Min | Max |
| 1 | Leicester City | East Midlands | 0.0162 | 0.0216 | 0 | 0.1850 |
| 2 | Arsenal | Greater London | 0.1390 | 0.0325 | 0.0306 | 0.2524 |
| 3 | Tottenham Hotspur | Greater London | 0.0659 | 0.0203 | 0.0184 | 0.2032 |
| 4 | Manchester City | North West | 0.0626 | 0.0195 | 0.0286 | 0.2044 |
| 5 | Manchester United | North West | 0.1323 | 0.0226 | 0.0285 | 0.2044 |
| 6 | Southampton | South East | 0.0294 | 0.0176 | 0.0111 | 0.1828 |
| 7 | West Ham United | Greater London | 0.0353 | 0.0164 | 0.0122 | 0.1456 |
| 8 | Liverpool | North West England | 0.1580 | 0.0352 | 0.0842 | 0.4478 |
| 9 | Stoke City | West Midlands | 0.0203 | 0.0157 | 0.0103 | 0.2820 |
| 10 | Chelsea | Greater London | 0.1035 | 0.0226 | 0.0307 | 0.1854 |
| 11 | Everton | North West England | 0.0370 | 0.0150 | 0.0162 | 0.1714 |
| 12 | Swansea City | Wales | 0.0240 | 0.0217 | 0.0093 | 0.3432 |
| 13 | Watford | East | 0.0084 | 0.0074 | 0 | 0.0398 |
| 14 | West Bromwich Albion | West Midlands | 0.0177 | 0.0097 | 0 | 0.1073 |
| 15 | Crystal Palace | Greater London | 0.0148 | 0.0076 | 0 | 0.0824 |
| 16 | Bournemouth | South West | 0.0127 | 0.0367 | 0.0023 | 0.2660 |
| 17 | Sunderland | North East Englad | 0.0269 | 0.0209 | 0.0088 | 0.2601 |
| 18 | Newcastle United | North East Englad | 0.0455 | 0.0280 | 0.0181 | 0.2789 |
| 19 | Norwich City | East | 0.0097 | 0.0239 | 0.0024 | 0.2310 |
| 20 | Aston Villa | West Midlands | 0.0320 | 0.0165 | 0.0155 | 0.1515 |

Table A.2: OLS - MEP 2014 Election on Brexit Referendum
Regressors are the share of votes of each party in the 2014 Members of the European Parliament Election.

| | (1) |
|---|---|
| | Dependent Variable: Remain Share of Votes in 2016 |
| Conservatives | 0.208 |
| | (0.0398) |
| Labour | 0.116 |
| | (0.0416) |
| Liberal-Democrats | 0.0453 |
| | (0.0577) |
| UKIP | -0.804 |
| | (0.0313) |
| Green Party | 0.723 |
| | (0.0567) |
| British National Pary | -4.052 |
| | (0.407) |
| Mean of Dependent Var | 47.087 |
| Observations | 379 |
| Adjusted $R^2$ | 0.918 |

Robust standard errors in parentheses

Table A.3: Descriptive Statistics - Vector of Possible Controls (1)

| Variable Name | Mean | Std. Dev. | Obs. |
|---|---|---|---|
| EU exposure through immigration, trade and structural funds | | | |
| Initial EU accession migrant resident share (2001) | 0.002 | 0.002 | 379 |
| EU accession migrant growth (2001-2011) | 0.017 | 0.016 | 379 |
| Initial EU 15 migrant resident share (2001) | 0.013 | 0.011 | 379 |
| EU 15 migrant growth (2001-2011) | 0.003 | 0.007 | 379 |
| Initial migrants from elsewhere resident share (2001) | 0.052 | 0.062 | 379 |
| Migrants from elsewhere growth (2001-2011) | 0.025 | 0.031 | 379 |
| Total economy EU dependence (2010) | 0.097 | 0.032 | 379 |
| EU Structural Funds per capita (2013) | 46.748 | 93.326 | 369 |
| 1975 referendum Leave share | 0.314 | 0.053 | 379 |
| Local public service provision and scal consolidation funds | | | |
| Share of residents commuting to London (2011) | 0.017 | 0.044 | 376 |
| Owned (outright + mortgage) share (2001) | 0.700 | 0.095 | 379 |
| Owned (outright + mortgage) share growth (2001-2011) | -0.040 | 0.028 | 379 |
| Council rented share (2001) | 0.121 | 0.079 | 379 |
| Council rented share growth (2001-2011) | -0.037 | 0.038 | 379 |
| Total scal cuts (2010-2015) | 448.032 | 122.125 | 379 |
| Share of suspected cancer patient treated within 62 Days (2015) | 82.983 | 7.224 | 378 |
| Public employment share (2009) | 0.212 | 0.070 | 379 |

Table A.4: Descriptive Statistics - Vector of Possible Controls (2)

| Variable Name | Mean | Std. Dev. | Obs. |
|---|---|---|---|
| Demography and education | | | |
| Share of res. pop. no quali cations (2001) | 0.353 | 0.069 | 379 |
| Share of res. pop. no quali cations growth (2001-2011) | -0.043 | 0.025 | 379 |
| Share of res. pop. quali cation 4+ (2001) | 0.196 | 0.073 | 379 |
| Share of res. pop. quali cation 4+ growth (2001-2011) | 0.075 | 0.015 | 379 |
| Population 60 older (2001) | 0.215 | 0.037 | 379 |
| Population 60 older growth (2001-2011) | 0.182 | 0.104 | 379 |
| Mean life satisfaction APS well-being data (2015) | 7.572 | 0.180 | 378 |
| CV life satisfaction APS well-being data (2015) | 1.141 | 0.440 | 378 |
| Economic structure, wages and unemployment | | | |
| Retail employment share (2001) | 0.1654 | 0.0216 | 379 |
| Retail employment share change (2001-2011) | -0.0071 | 0.0080 | 379 |
| Manufacturing employment share (2001) | 0.1511 | 0.0527 | 379 |
| Manufacturing employment share change (2001-2011) | -0.0584 | 0.0204 | 379 |
| Construction employment share (2001) | 0.0701 | 0.0143 | 379 |
| Construction employment share change (2001-2011) | 0.0098 | 0.0058 | 379 |
| Finance employment share (2001) | 0.0446 | 0.0264 | 379 |
| Finance employment share change (2001-2011) | -0.0041 | 0.0075 | 379 |
| Median hourly pay (2005) | 11.0051 | 1.9897 | 379 |
| Median hourly pay change (2005-2015) | 0.2369 | 0.0864 | 379 |
| Interquartile pay range (2005) | 9.9733 | 3.1095 | 371 |
| Interquartile pay range growth (2005-2015) | 0.2005 | 0.1290 | 367 |
| Unemployment rate (2015) | 5.2920 | 2.1157 | 377 |
| Self-employment rate (2015) | 10.6524 | 3.7346 | 378 |
| Participation rate (2015) | 78.6425 | 4.4908 | 379 |
| European Parliament 2014 Election | | | |
| Conservatives (%) | 25.403 | 8.608 | 379 |
| Labour Party (%) | 22.797 | 11.953 | 379 |
| Liberal Democrats (%) | 6.827 | 4.565 | 379 |
| UKIP (%) | 28.967 | 9.337 | 379 |
| Green Party (%) | 7.596 | 3.440 | 379 |
| British National Party (%) | 1.121 | 0.620 | 379 |

Table A.5: Vector of Chosen Controls for the Main Speci cation

| |
|---|
| EU exposure through immigration, trade and structural funds |
| Initial EU accession migrant resident share (2001) |
| Initial EU 15 migrant resident share (2001) |
| Total economy EU dependence (2010) |
| EU Structural Funds per capita (2013) |
| Local public service provision and scal consolidation funds |
| Share of residents commuting to London (2011) |
| Owned (outright + mortgage) share (2001) |
| Council rented share (2001) |
| Total scal cuts (2010-2015) |
| Demography and education |
| Share of res. pop. no quali cations (2001) |
| Share of res. pop. no quali cations growth (2001-2011) |
| Share of res. pop. quali cation 4+ (2001) |
| Share of res. pop. quali cation 4+ growth (2001-2011) |
| Economic structure, wages and unemployment |
| Retail employment share (2001) |
| Retail employment share change (2001-2011) |
| Construction employment share (2001) |
| Construction employment share change (2001-2011) |
| Finance employment share (2001) |
| Finance employment share change (2001-2011) |
| Median hourly pay (2005) |
| Median hourly pay change (2005-2015) |
| Unemployment rate (2015) |
| Self-employment rate (2015) |
| Participation rate (2015) |
| European Parliament 2014 Election |
| UKIP (%) |

# B Appendix: Other Results

## B.1 LASSO-Chosen Controls

One valid concern, is that given dimension of the vector of possible controls, one will be able to choose the vector control arbitrarily, manipulating the results of the main speci cation. In order to avoid cherry-picking covariates, and following a similar empiric approach as Becker, Fetzner and Novy (2017) and Larreboure and Gonzalez (2019), I will choose the vector of controls for each speci cation by machine-learning LASSO algorithm (Least Absolute Shrinkage and Selection Operator) in a robustness check. Table B1 shows the LASSO-chosen vector of controls. From a total of 47, 12 where chosen. Table B2. Shows the results. The rst column as always, is the benchmark, thereby at looking columns (2) and (2) it can be seen that the main results are robust to LASSO-Chosen controls.

Table B.1: Robustness Check - Vector of LASSO- Chosen Controls for the Main Speci cation

| LASSO-Chose Controls | Numbers of Control not Chosen |
|---|---|
| EU exposure through immigration, trade and structural funds | |
| None | 9 |
| Local public service provision and scal consolidation funds | |
| Owned (outright + mortgage) share growth (2001) | 5 |
| Council rented share growth (2001) | |
| Treated | |
| Demography and education | |
| Share of res. pop. no quali cations (2001) | 4 |
| Share of res. pop. no quali cations growth (2001-2011) | |
| Share of res. pop. quali cation 4+ (2001) | |
| Share of res. pop. quali cation 4+ growth (2001-2011) | |
| Economic structure, wages and unemployment | |
| Retail employment share (2001) | 13 |
| Construction employment share change (2001-2011) | |
| Finance employment share (2001) | |
| Economic structure, wages and unemployment | |
| UKIP (%) | 3 |
| Green Party (%) | |
| British National Party (%) | |

Table B.2: Robustness Checks - Lasso-Chosen Controls

Points non-UK and Points UK are the foreign and UK average points reached by players by team, interacted with the percentage of fans by team at district-level. Points = Points non-UK - Points UK, i.e. the relative points of foreigners to locals. Variables with 2 t 1 are the same variables calculated for the previous season. Average Points is the average between Points and Points 2 t 1. Points is standardized. Big 4 and Big 6 are the 4 and 6 biggest teams in fans share at district level. Top 7 are the are the best 7 of the 2015-16 season.

| | Dependent Variable: Remain Share of Votes in 2016 | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Average Points | 0.797 | 1.046 | 0.808 |
| | (0.187) | (0.224) | (0.234) |
| Controls | | | |
| Lasso-Chosen Controls | No | Yes | Yes |
| Full Set of Controls | Yes | No | No |
| UKIP Share of Votes | Yes | No | No |
| Region F.E. | Yes | No | Yes |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 |
| Observations | 374 | 374 | 374 |

Conley (1999) standard errors in parentheses, calculated assuming correlation within 50km..

Table B.3: Robustness Check - Alternative Hypothesis

Di erence in position with last season is the position in the season 2014-2015 minus the season 2015-2016. 15-16 Season Position is the position in the season for each team.

| | Dependent Variable: Remain Share of Votes in 2016 | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| 15-16 Season Position | -0.0738 | -0.0553 | | | -0.0551 | -0.0737 |
| | (0.109) | (0.110) | | | (0.114) | (0.124) |
| Di erence in position with last season | | | 0.197 | -0.110 | 0.154 | -0.165 |
| | | | (0.583) | (0.670) | (0.618) | (0.715) |
| Controls | | | | | | |
| Full Set for Main Speci cation | Yes | No | Yes | No | Yes | No |
| LASSO-Chosen | No | Yes | No | Yes | No | Yes |
| Mean of Dependent Var | 47.087 | 47.087 | 47.087 | 47.087 | 47.087 | 47.087 |
| Observations | 375 | 378 | 375 | 378 | 375 | 378 |
| Adjusted $R^2$ | 0.958 | 0.952 | 0.958 | 0.951 | 0.958 | 0.951 |

Conley (1999), standard errors in parentheses.

# C Appendix: Players Points

Table C.1: Points Based on Players Performance

| Action | Points |
|---|---|
| For playing up to 60 minutes | 1 |
| For playing 60 minutes or more (excluding stoppage time) | 2 |
| For each goal scored by a goalkeeper or defender | 6 |
| For each goal scored by a midfielder | 5 |
| For each goal scored by a forward | 4 |
| For each goal assist | 3 |
| For a clean sheet by a goalkeeper or defender | 4 |
| For a clean sheet by a midfielder | 1 |
| For every 3 shot saves by a goalkeeper | 1 |
| For each penalty save | 5 |
| For each penalty miss | -2 |
| Bonus points for the best players in a match | 1-3 |
| For every 2 goals conceded by a goalkeeper or defender | -1 |
| For each yellow card | -1 |
| For each red card | -3 |
| For each own goal | -2 |

https://fantasy.premierleague.com/help/rules.