Genetic Diversity and Performance: Evidence From Football Data*

Michel Beine †, Silvia Peracchi, ‡ Skerdilajda Zanaj§

Department of Economics and Management

University of Luxembourg

July 7, 2021

Abstract

The theoretical impact of genetic diversity is ambiguous since it leads to costs and benefits at the collective level. In this paper, we assess empirically the connection between genetic diversity and the performance of sport teams. Focusing on football (soccer), we built a novel dataset of national teams of European countries that have participated in the European and the World Championships since 1970. Determining the genetic diversity of national teams is based on the distance between the genetic scores of every players' origins in the team. Genetic endowments for each player are recovered using a matching algorithm based on family names. Performance is measured at both the unilateral and bilateral level. Identification of the causal link relies on an instrumental variable strategy that is based on past immigration at the country level about one generation before. Our findings indicate a positive causal link between genetic diversity and teams' performance. We find a substantial effect, a one-standard increase in diversity leading to ranking changes of two to three positions after each stage of a championship.

Keywords: Genetic diversity, Football, Sports team, Performance, Family names, Migration

JEL classification: F22, F66, O15, O47, Z22

*We thank K. Desmet, P. Poutvaara, K. Tatsiramos, P. Van Kerm, S. Scicchitano and C. Vasilakis for comments and suggestions on earlier versions. The usual disclaimers apply.

†Email: michel.beine@uni.lu

‡Corresponding author. DEM, University of Luxembourg, 6 Rue Richard Coudenhove-Kalergi, 1359 Luxembourg, Email: silvia.peracchi@uni.lu

§Email: skerdilajda.zanaj@uni.lu

1 Introduction

Over the last decades, international human mobility has been on the rise, involving millions of people moving to another country. Today, there are more than 240 million people living in a country other than the one in which they were born. This process has led to significant changes in the cultural and genetic landscapes of the host countries, with important consequences for the size and the composition of their labor force. Migrants bring with them deep-seated social values, human capital, institutions, history, and traditions. As a consequence, countries that have experienced large immigration flows in the past are characterized today by greater diversity in their populations.

National teams in international sport competitions also reflect the increased level of diversity brought about by immigration flows. In football, the most popular sport worldwide, national teams in immigration countries have become more diverse because the teams attract players from the larger and more diversified talent pool that is available in the country. At the 2018 FIFA Men's World Cup in Russia, 84 football players competed for national teams of countries other than their country of birth. It was the second-highest absolute number of foreign-born footballers in the history of the World Cup (van Campenhout et al., 2019). More significantly, in immigration countries, a high proportion of players on national teams are second-generation migrants, bringing with them some genetic endowment different from the one found in the native population of the country they play for.

Genetic or ethnic diversity is a key dimension of diversity, exerting a potential effect on productivity and collective performance. Previous work on ethnic diversity suggests that higher diversity exerts a positive effect on global productivity (Alesina et al., 2016; Alesina and La Ferrara, 2005). Regarding the genetic dimension, Ashraf and Galor (2013) argue that there is an optimal level of genetic diversity in terms of productivity. On the one hand, genetic diversity brings complementarity in skills, which results in a higher level of productivity. On the other hand, genetic information of a population is a proxy for its history, culture, and social values. Genetic diversity is an excellent summary statistic capturing divergence in the whole set of implicit beliefs, biases, conventions, and norms transmitted across generations—biologically and culturally—with high persistence (Spolaore and Wacziarg, 2009; 2018). Besides, ancestry affects culture even after several generations (Guiso et al., 2006) not only because culture is transmitted to an enormous degree intergenerationally, but also because genetic differences among individuals with different ancestries are related to differences in their values and preferences (Desmet et al., 2017). These divergences associated with diversity might mitigate or offset diversity's positive impact on productivity.

In this paper, we investigate the role of genetic diversity in the performances of national football teams. One interesting aspect of this sports activity is the fact that performances are measured precisely and are much less subject to measurement errors than are other economic activities. The case of football is interesting, beyond the fact that it is the most popular sport worldwide, since the performance of a team relies on the interaction of players who need to have very different skills, depending on their position on the pitch. This clearly refers to the complementarity of skills channel mentioned above. It is empirically unclear in football to what extent the cultural channel and the divergence-in-beliefs channel associated with higher diversity are substantial and might offset the positive effect of the skill complementarity. Anecdotal evidence suggests, however, that there is some belief that diversity does affect football performance positively. In 2012, Belgium succeeded to a 2–0 away win over Scotland during the World Cup qualification process. Commenting on this result, Scotland assistant manager Mark McGhee described the Belgian team's skill pool as follows: [1]

> They are choosing from a pool that is different from us. They have the advantage of an African connection and can bring in real athleticism...We can hope, of course, that out of the gene pool that is East Dunbartonshire, Lanarkshire and South Ayrshire we produce a group of players that will one day be as good as them. But they have a much broader base, and I think that is a huge advantage.

Former U.S. President Barack Obama, in his tribute speech to commemorate Nelson Mandela's birthday in 2018, praised the diversity of French football team, stating that

> [diversity] delivers practical benefits since it ensures that a society can draw upon the energy and skills of all... people. And if you doubt that, just ask the French football team that just won the World Cup because not all these folks look like Gauls to me....[2]

As of February 18, 2021, Belgium and France were ranked first and second worldwide respectively, according to the World Rankings provided by the Fédération Internationale de Football Association (henceforth FIFA).[3] One of the goals of this paper is to check whether these perceptions are supported by some sound statistical analysis.

To establish a causal link between the sportive teams' genetic diversity and performance, we develop specific measures of the key dimensions, i.e., performances and genetic diversity. Performance data are collected at the match and tournament level for European teams based on their results at the World Cup and the European championship competitions from 1970 onward. At the tournament level, our benchmark performance indicator is the Elo score ranking of the national team that gives a synthetic value of the recent performances of any national team. At the match level, we use the goal difference as the benchmark but show that our results are robust to alternative measures. The genetic diversity of each team is based on the bilateral *distance* of genetic scores called the "expected heterozygosity" among every two players in the team. "Expected heterozygosity" measures the probability that two randomly selected individuals from the same population differ genetically from one another for a given spectrum of traits. We follow the approach of using family names to capture the ethnic background of individuals adopted in different fields such as the patents literature (Kerr and Kerr, 2018) or the study of intergenerational mobility (Clark, 2015).[4] Our measure of ethnic diversity at the national level suggests that diversity has changed significantly over the period of investigation, especially in countries of past intensive immigration.

[1] Mark Wilson, "Brilliant Belgians just incomparable insists Scotland assistant coach McGhee,"

[2] France24, "In Mandela address, Obama cites French World Cup model as champs of diversity,"

[3] FIFA.com. "Men's ranking: Belgium, Royal Belgian Football Association." https://www.fifa.com/fifa-world-ranking/associations/association/BEL/men/

[4] This surname-based idea was previously adopted in the patents literature (Kerr and Kerr, 2018) and in the study of intergenerational mobility, as in Clark (2015). An alternative predictor of player origins would be, for instance, the birth country, as used in van Campenhout et al. (2019) for their players' diversity index. This measure would likely be a good match for players who undergo naturalization, but it would fail to capture second-generation aspects of immigration. This last is critical for our setting, as we focus on the vertical-transmission mechanisms related to group-dynamics, focus on national teams, and base our identification strategy on previous-generation migration patterns.

The econometric analysis of the causal link between genetic diversity and performance of national teams is likely to be affected by a set of confounding factors that can bias the assessed impact of diversity. Our identification strategy relies on an instrumental variable (IV) approach that makes use of the structure of past immigration flows at the population level. More specifically, we instrument the genetic diversity of football's national teams with a measure of genetic diversity for the immigration flows about one generation before (20 years). The idea is that higher diversity in immigration yesterday increases the diversity of second-generation migrants who today play for the national team of their parents' adopted country. The strict rules of eligibility for participation on a national team in football prevent the implementation of a strategy in which diversity could be manipulated by national federations. This lowers the concern that this instrument does not comply with the exclusion restriction. Our IV results therefore allow one to uncover an overlooked benefit of immigration, namely, its long-run benefit in terms of performance in collective sports. We hypothesize, and then show empirically, that genetic diversity implies significant complementarities (tactical and physical) among players, affecting performance positively. It is important to note that we do not, of course, address the direct effect of genes on sports performance. In contrast, our analysis addresses the benefits and drawbacks of genetic diversity on performance measured at a collective level. We expect genetic diversity in sports to affect performance through a variety of channels. These channels include (i) the ability to play as a team, conveyed by norms of cooperation belonging to different nationalities; (ii) the creativity of novel ways to play sports; and (iii) the improved complementarities among players in view of the different skills required for different roles in the game. We find a positive net benefit on team's performance.

Our results hold at both the tournament and match levels. At the tournament level, along with our measure based on the Elo scores, a one-standard-deviation increase in diversity would lead to a scaling upward of about 2 to 3 positions after each tournament. At the match level, a one-standard-deviation increase in diversity yields an increase of one point in the goal difference. These findings are robust to a set of robustness checks and to some invalidation exercises. The results are also robust to whether passive players are included or not, to alternative measures of ethnic distance, to the way bilateral performances are captured, and to the fact that hosting teams usually have an advantage in football. In addition, we control for coaching quality that could confound the identification of the causal impact of diversity. The results are also robust to the number of years that past immigration flows are expected to impact genetic diversity of national teams in the first stage of the IV analysis. Finally, we perform a placebo test using performances in athletics, i.e., a sport in which diversity should not play any role, given that competitions do not involve any collective cooperation. We do not find any role of genetic diversity in explaining performances in athletics.

While our paper is clearly connected with the literature on the role of ethnic and birthplace diversity, our analysis is also related to a large empirical literature looking at the role of immigration in football. This literature is reviewed in the next section. Our paper deviates from the existing papers in that we focus on the performances of national teams, not on football clubs. In the context of this investigation, a similar analysis at the club level would be more subject to endogeneity issues. Through transfers of players, a club could explicitly implement a strategy to boost diversity in order to improve the team's performances. Given the strict rules governing the composition of national teams in football, such a strategy would hardly be possible. While some naturalization strategies have sometimes been implemented, they remain more an exception than the rule.

The paper is organized as follows. Section 2 briefly reviews the relevant literature. In Section 3, we describe the data used in our analysis. Section 4 introduces the empirical analysis. Section 5 presents the main results, discusses identification issues, and Section 6 exposes the robustness checks. Our placebo analysis is detailed in Section 7. Section 8 concludes. Results tables are displayed in Section 9.

2 Literature review

The economic implications of diversity have produced a very extensive literature. Prior studies investigate the effects of ethnic diversity on growth (Easterly and Levine, 1997; Docquier et al., 2019); on economic prosperity (Alesina et al., 2016); on trade (Alesina et al., 2000); on public policy (Collier, 2001); on polarization (Bove and Elia, 2017); on individuals' preferences (Alesina and La Ferrara, 2005); and on the provision of public goods (Spolaore and Wacziarg, 2009). Prior studies also relate diversity to the performance of collective organizations. The seminal model of Lazear (1999) illustrates global organizations as multicultural teams. To offset the costs of cross-cultural interaction, the complementarities among different workers must, however, be substantial. Delis et al. (2017) use a panel of U.K. and U.S. firms listed on the stock market and track the genetic diversity of the board of directors, finding positive effects on the firm's performance as measured by risk-adjusted returns and the Tobin's Q. Delis et al. (2021) apply a similar analysis to the movie industry, finding an optimal degree of genetic diversity of actors and directors on the box office figures of attendance. In Prat (2002), diversity of team members results in diverse decision-making processes, which bring benefits in the case of actions' submodularity.

The literature on genetic diversity is more limited. Spolaore and Wacziarg (2018), Ashraf and Galor (2013) and Delis et al. (2017) are seminal contributions that relate genetic diversity and performance. This distinction between cultural and genetic distance is relevant because these diversities may present different patterns (Alesina et al., 2016). To the best of our knowledge, our paper is the first study to explore the effects of genetic diversity on sports performance.

Focusing on sports, Kahane et al. (2013) provide evidence from hockey and generally find a positive effect of cultural diversity captured by a discrete HHI index. Parshakov et al. (2018) use e-sport data to investigate the impact of cultural, language, and experience heterogeneity on performance. Cultural diversity correlates positively with tournaments performance, while language and experience diversity are found to affect performance negatively. Gould and Winter (2009) build a panel of baseball players from 1970 to 2003 and observe that workers' (players') efforts and interactions depend on the complementarities in the production technology. A recent contribution by Tovar (2020) explores the link among diversity, national identity, and performance at the player and team level, analyzing data from the Spanish and English leagues. The study found a non-linear relationship between the team's and the players' performance.[5] Also concentrating on club-level performance, Brox and Krieger (2019) provide evidence from German

[5]Another related paper using clubs and not national teams, is Haas and Nüesch (2012). This study uses match-level, panel data (ranging from 1999 to 2005) from the German Bundesliga, employing the nationality of team members. It documents a negative effect on the number of points received given the game outcome, the goal-difference, and an average of individual players' performance evaluations made by experts. In addition, Vasilakis (2017) examines how the increase in mobility has reshaped the players' market among clubs and produced

men's football, finding that an intermediate level of birthplace diversity maximizes team performance. Ingersoll et al. (2017) enlarge the set of countries and investigate the effect of cultural diversity on the club teams' performances in the top leagues in the UEFA Champions League (2003–2012) for Germany, England, Italy, France, and Spain. In their findings, culturally heterogeneous teams outperform homogeneous ones, cultural diversity being proxied by linguistic diversity data based on players' nationality.

We contribute to the sports literature in various areas. We use genetic diversity to capture deeply rooted differences in values related to culture, language, and other diversity dimensions. This diversity helps to attenuate any endogeneity concern. The dataset we build for that purpose includes a much larger number of countries and tournaments than do previous studies. We establish a causal link, not just a correlation, between performance and diversity. Finally, our perspective is innovative as we tackle the importance of an intergenerational aspect of diversity in sports teams. In doing so, we can better assess the causality of the relationship among past immigration, diversity, and sports performance.

3 Data

To analyze the impact of genetic diversity on the performance of national football teams, we collect and build indicators of diversity and performance as well as other variables. We start by explaining how key data are built, namely, genetic diversity at the team level and the performance. We then present other variables that enter into the subsequent econometric analysis.

3.1 Measuring genetic diversity at the team level

Our key indicator of interest to explain the performance of a given national football team is its genetic diversity. To capture this relationship, we gather information on the team composition. From this, we then establish a measurement for the genetic characteristics of each team member and relate how the individual genetic endowments are combined to yield an indicator of diversity.

National team composition.

We collect data on the composition of national squads from the website *worldfootball.net*, with some comparisons and checks using *soccerway.com* and Wikipedia. Squad data on Turkey was absent for two periods in the main source, and the desired information was obtained through the source *https://www.national-football-teams.com*. For every European team that entered either tournament $\in$ {Euros, World Cup} over the period 1970 to 2018, we obtained information on players' names, their ages, and their minutes/appearances in the competition at each stage $\in$ {Qualification, Finals}.[6]

distributional effects in terms of performance and wages.

[6]Given the full name lists, we proceeded with a splitting to separate the father name information. The web source *soccerway.com* presents players' profiles with names and surnames separated. Whenever we could match the player in our sample to his profile on *soccerway.com*, we used the surname as presented in the source. In the other occurrences, name splitting was performed according the following decision rule: we extracted the last part of the full name instance by taking into account particular nominal particles, such as "De," "Van," "Van Der," "Von," "Di," etc. With Spanish and Portuguese teams, the splitting followed the typical country's customs: for Spain, the first

In our baseline specifications, we include each player from the squad list in our diversity measure, regardless of his appearance time. Ingersoll et al. (2017) focus on football clubs and identify that cultural diversity on the pitch matters positively for performance. Yet they find an insignificant effect for off-the-pitch interactions. To accommodate this possible heterogeneity, we include minutes played as weights in our diversity calculations in one of our sensitivity checks.

Ethnicity of players.

For societies with patrilineal surnames customs, surnames are known indicators of population structure and relatedness in the genetic literature (Piazza et al., 1987; Jobling, 2001), and are not new to the economic literature. For instance, works by Kerr and Kerr (2018), Clark (2015) and Buonanno and Vanin (2017) in different fields of economics use surnames to predict ethnicity and community relatedness. We follow this global approach in order to characterize the genetic diversity of each national team. We obtain data on each surname's geographical distribution from the web source *forebears.io*, which presents a set of country-level statistics for a great variety of surnames.

More specifically, for each unique surname in the full list of players in our dataset, this source provides the three countries ($country_1$, $country_2$, $country_3$) displaying the highest incidences (i.e., number of people having that surname in a particular country) and the highest frequencies (i.e., percentage of people having that surname in a particular country) of that specific surname. We then identify the best predicted country $i*$ for a surname as the country i associated with the highest value of the variable ($Incidence_i * frequency_i$, $i \in country_1, country_2, country_3$). This procedure avoids favoring very small countries, which would occur if we looked only at the frequency (e.g., virtually every surname in Monaco has very high frequencies). Further, it avoids favoring very big countries, as would happen if one relied on the incidence only (e.g., countries like the U.S. have generally higher incidences, even for rare surnames).[7] Our website of choice has the important feature of delivering accent-sensitive information, which increases precision when mapping a surname and a country of origin.[8] While measurement error concerns do arise with the choice of this proxy, this method performs quite well in capturing the second-generation of migrants who may still contribute to the team's diversity (e.g., French national Zinedine Zidane was born in Marseille and is of Algerian descent).[9] Examples of the prediction results are found in Appendix B.[10]

surname corresponds to the father's surname, and vice versa for Portugal. We focus on father surnames for cross-country comparability.

[7]A further manual cleaning was performed using a language detection algorithm in Python. While these algorithms tend to perform best for common nouns rather than surnames and for phrases rather than single words, we compared the language predicted with the country predicted and assessed and eventually corrected a minority of surnames manually.

[8]Building a small sample of 314 recent national teams' players, whose ethnicity was found through a set of online newspapers, the *forebears.io*-based technique performed better than two alternative measures considered: *www.name-prism.com/* and *http://abel.lis.illinois.edu/cgi-bin/ethnea/search.py*. The results are not reported here in the interest of space but can be obtained upon request.

[9]As a further cleaning process, we used language-predictive libraries (TextBlob, langdetect) in Python to see whether the surname prediction coming from our algorithm was in line with these library-based predictions. With this approach, in some minor cases, we corrected a minority of surnames manually. cases, we corrected a small number of surnames manually. We clarify that the main purpose of this set of libraries is to classify sentences and common names, rather than family names. Further, they predict languages, rather than ethnicity. Further, they predict language rather than ethnicity. We therefore employed this tool very conservatively.

[10]Referring to the Belgium example in Appendix B, it is obvious that the matching algorithm is efficient but not perfect. The match between the ethnicity and the surname is rather good (85 per cent of correct predictions). Two types of errors in terms of their incidence

Genetic diversity.

Diversity Div_{ist} of team i at time $t \in \{1970, ..., 2018\}$ and at competition stage s is given by :

$$Div_{ist} = \frac{1}{S_t} \sum_{j=1}^{N_t} \sum_{k=1}^{N_t} (p_{jt} p_{kt} d_{jk}), \; j \neq k \tag{1}$$

where p_{jt} and p_{kt} are the shares of players on the team (predicted to be from origin j and k respectively) belonging to the set of origins $\{1,...N_t\}$ in team i for stage s of championship t. The fraction $\frac{1}{S_t}$ operates as a normalization factor for different squad sizes reported on the web source for the qualification stages. $d_{jk}$is the genetic distance between origin j and origin k, belonging to the set of surname-predicted origins in the squad. We use genetic distances in a fashion similar to Alesina et al. (2016), implying that our indicator can be seen as a weighted average of genetic distances over all origin pairs in the team. This approach is comparable to Ingersoll et al. (2017)'s linguistic diversity and does not profoundly differ from linguistic diversity indicators proposed by the seminal work of Greenberg (1956) and re-elaborated in Fearon (2003). Data on bilateral genetic distance d_{jk} come from Spolaore and Wacziarg (2009) who adapt distance matrices from the genetic literature (Cavalli-Sforza et al., 1994). This specific modification, key to our framework, allows more weight to be given to more genetically distant origins.[11]

As a snapshot example, we report in Figure 1 the cross-country variation of diversity in the EUROs 2016. A general pattern appears with Eastern Europe teams presenting lower diversity levels, whereas in Western Europe teams show higher levels of diversity, likely reflecting accumulated migration inflows over the recent decades.[12]

occur. The most detrimental error is the case of the striker Batshuayi that is spuriously attributed to the Belgian ethnicity (rather than to the Democratic Republic of the Congo). This error is due to the fact that this surname is rare and/or the coverage of surnames incidence in the DRC is rather poor. Most of the other errors have little if no impact on the diversity level. The reason is that surnames have either some French or Dutch connotations. This leads to spurious predictions in the case of Courtois, Lambert, and Meunier on the French side and in the case of Van Der Linden or Thissen in the Dutch case. Nevertheless, when attributed to an ethnicity of a neighboring country, there is no impact on the diversity measure since the genetic distance between Belgium and these countries is zero. The errors outlined in the Belgian case are also due to the particular linguistic situation of the country that has official languages (French, Dutch, and German) that originate in the neighboring countries.

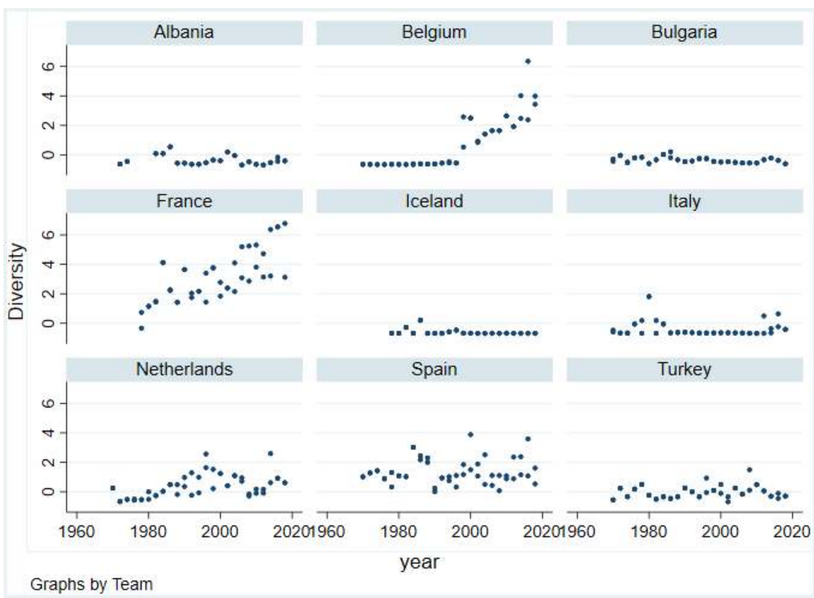
[11] This source led us to exclude two national teams from our sample, Andorra and Liechtenstein, as they are not part of the Spolaore and Wacziarg's dataset. All other countries were included.

[12] Kazakhstan's exception likely reflects the high ethnic diversity of the country: https://www.britannica.com/place/Kazakhstan/People

Figure 1. Diversity of national teams, EURO 2016, qualifications

Notes: In Figure 1, we plot a cross-sectional example for our diversity index, taking the 2016 EURO qualifications as the tournament of reference. As a general pattern, we observe higher levels of diversity in the Western area.

Figure 2. Genetic diversity over time, selected national teams



Notes: In Figure 2, we present the time variation of our index of genetic diversity for a subset of teams. While for some countries, like Belgium, we can identify a sudden change in the compositional diversity in the most recent decades, some other countries like France and the Netherlands display a smoother evolution pattern. This contrast might be explained by the different patterns of past immigration. Countries such as Portugal show higher, yet noisier team diversity levels. Italy, Albania, and Bulgaria are examples of countries with lower and relatively stable index values. These countries are, at least up to a recent period, mainly emigration rather than immigration countries. Iceland is a typical example with almost no genetic diversity in its national team due to the relative isolation of the country in terms of human mobility.

3.2 Measuring performances of national teams

We use two different dimensions to characterize the performances of national football teams. First, we use an absolute measure of performance of team i based on its ranking. This refers to the unilateral dimension of the performance data. Second, as a relative measure of performance, we use results at the match level. This measure is dyadic in nature, as the performance depends on the considered team i but also on the performance of the opponent j.

In the unilateral setting, our performance indicator is the Elo score of a team.[13] Updated after each game, the Elo score of a team is a function of its previous score, the realized and the expected results (given the opponent's relative strength) and the importance of the tournament. A complete description and formula are found in Appendix A. Based on match-level information, we constructed Elo ratings relative to the results of the EURO and World Cup qualifications and final stages for our whole sample. Our preferred measure would be the change in the score from the beginning to the end of the championship stage. For team i, performing in stage s, at Championship t, our baseline performance measure for the unilateral setting is therefore

$$Performance_{ist} = Elo\ score_{End,ist} - Elo\ score_{Beginning,ist} \tag{2}$$

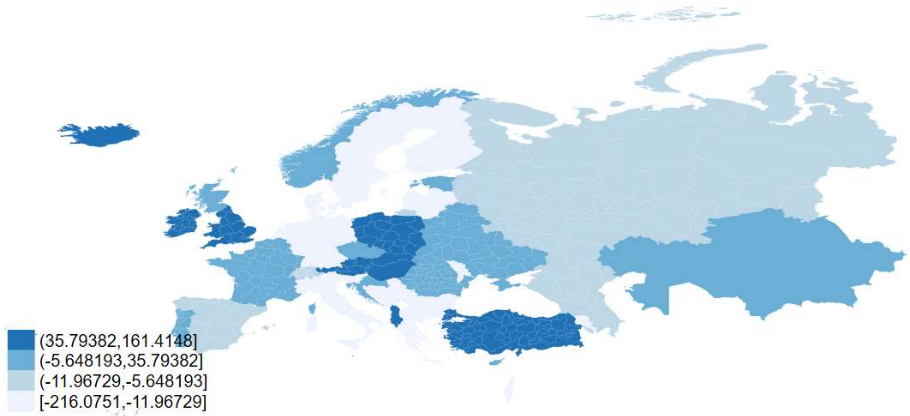
The Elo score measurement is based on an updating process, where a new value at each match replaces the old value, according to the match result and its expectation. If a team is new in the sample, this computation requires an initial value. To provide reasonable starting values, we calibrate these instances with Elo score data available for every championship and stage at *eloratings.net*. As part of our battery of robustness checks, we also employ the Elo measures proposed on the website. Significantly, our computed outcomes differ from the website's in that *eloratings.net* includes all matches with all opponents (including those non-European Teams in the World Cup final stage).[14]

We show in Figure 3 a snapshot of the score change, taking the example of the 2016 EURO Championship qualification stage. As a benchmark, France (the tournament host) had a score change of zero. In 2016, countries like Iceland and Albania qualified for the final stage for the first time in the event's history. As Figure 3 shows, the Elo score updates give more weight to unexpected results. The worst performers in terms of score changes were the 2004 Champion Greece and World Cup 2014 third-place finisher The Netherlands. The two teams did not qualify for the final stage. To complement with a time series example, Figure 4 plots the score change. The change in the score follows a stationary process, which rules out concerns related to the presence of unit roots in the outcome variables.

[13]Named after its inventor Arpad Elo, the Elo system was first introduced for comparing chess players' relative performances and was brought to football by Bob Runyan in 1997 (Langville and Meyer, 2012).

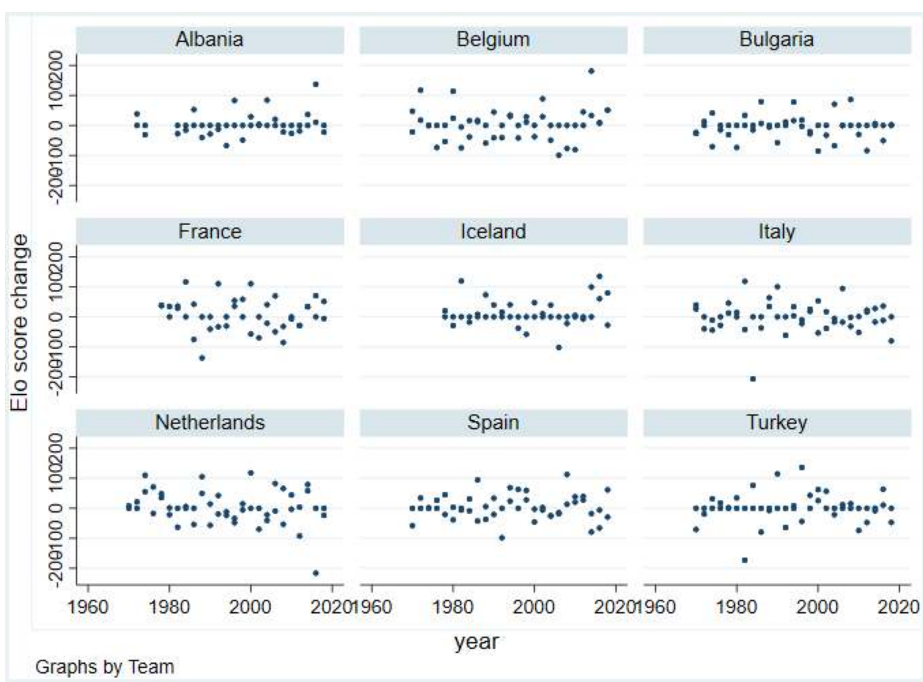
[14]It is worth mentioning that our Elo scores have a raw correlation of 97.5% with the website eloratings.net's index. In terms of the score change, the statistic is slightly lower (81.7%) but still very high.

Figure 3. Change in Elo ratings of national teams, EURO 2016, qualifications



Notes: In Figure 3, we plot the cross-sectional example for our performance measure for the unilateral specifications, taking the 2016 EURO qualifications as the tournament of reference. The variation reflects the relative performances of teams that improved on or worsened their Elo scores, based on their expected vs. realized match results. (The details are in Appendix A)) As France was the host, the team accessed the final stage directly, therefore having a score change of zero.

Figure 4. Elo score changes over time, selected teams



Notes: In Figure 4, we present the time variation of our Elo score change measure for a subset of teams. This picture reflects the stationary nature of the score.

In the bilateral specification, the performance indicator is the goal difference. Data at the match level come from the collection *International Football Results from 1872 to 2020* assembled by Mart Jürisoo. It includes a complete and updated men's football international matches dataset.[15]

Figure 5 provides a summary of the key components of the bilateral measure, i.e., scored and received goals, broken down between home (left panel) and away (right panel) matches. The figures confirm that, on average, teams perform better at home than abroad, a well-known feature in football competitions. We will account for this feature in the econometric specification involving the bilateral dimension of performances.

Figure 5. All-time goals scored and received, all national teams

Notes: In Figure 5, we present the all-time averages for the teams' bilateral performances, key outcome in our baseline estimations. Blue bars represent the average goals scored, whereas red bars represent the average goals received. On the left we list results for the teams listed as *home teams* in our dataset; on the right, we depict the same statistics for the teams when listed as *away teams*.

Tables 1 and 2 in the Section 8 provide summary statistics for the main variables in the unilateral and bilateral data. The full list of countries included in the sample is given in Table 30.

3.3 Other variables

We include various covariates affecting the performances of national teams. These variables are observed at either the team or country level. In our benchmark estimates, at the team level, we include the average age in its quadratic form and the players' appearance time variation for the team. We also include the standard deviation in the team members' minutes to better disentangle possible turnover decisions or other strategic concerns that may reflect the distribution of talent within the team. Country-level controls involve population (in millions), (the log of) GDP per capita, and past immigration stocks. Population data are retrieved from the Centre d'Études Prospectives et d'Informations Internationales (CEPII) for the period up to 2014 and then completed using World Bank data for the most recent

[15] Mart Jürisoo, International Football Results from 1872 to 2020. Retrieved on January 2020. https://www.kaggle.com/martj42/international-football-results-from-1872-to-2017/tasks (version 4).

values. GDP data (at constant 2015 prices) are extracted from the United Nations data office;[16] immigrant stocks are retrieved from the World Bank and start in 1960. As we lag this information, estimates that include this covariate will reduce the sample size to more recent years (beginning in 1978). We provide extensive information on all variables in our regressions in Appendix F.

3.4 Instrument

Our goal is to estimate a causal relationship between the football teams' genetic diversity and their performance. As we include a set of controls at team and national levels, together with Team level fixed effects and country dummies, concerns regarding the endogeneity of our variable of interest are mitigated. Still, it is possible that a set of current political, cultural, economic or institutional conditions that are not considered in our framework will fall into the error term, resulting in a potential omitted variable bias. As an example, naturalized players and, more generally, players who possess more than one nationality may be able to choose which national team to play for. They may have incentives to play for countries offering favorable conditions. These conditions may reflect financial/cultural/institutional and/or football-related resources that may correlate as well with the team performance. The squad selection process may also reflect cultural and/or institutional characteristics of the countries. If this selection is carried out to favor native players over second-generation migrants, this could cause inefficiencies in the talent selection, thus undermining the teams' performance. While part of these issues may be fixed over time, we allow for time variation in these characteristics and carry out an instrumental variable approach to ensure causality under these circumstances.[17]

To play for national teams, players need to comply with strict conditions of eligibility and, in particular, need to be nationals of the represented country.[18] Eligible players would therefore be either naturalized immigrants, or children of natives or second-/third-generation immigrants in their adopted country.[19] National teams' diversity is therefore driven by the immigration history of the previous generation of their representing country. Countries with low immigration rates will therefore result, everything else being equal, in a low diversity, driven mainly by the genetic endowment of the native population. This would also be true in countries with high immigration rates but with a concentrated origin of the immigrants. High diversity will be in countries with significant immigrant flows originating from diverse areas. As past immigration to a destination country translates into current variety in its nationals, we build a historical measure of country diversity that should predict how diverse the national team will be years later.

[16] National Accounts Section of the United Nations Statistics Division: National Accounts Main Aggregates Database. https://unstats.un.org/unsd/snaama/Basic

[17] It should also be noted that we build our diversity measure from ancestry information as proxied by surnames, which we argue captures the genetic diversity well. We believe it is a suitable alternative to indices built on the country of birth or nationality. However, our diversity formula is a quantization process that involves measurement error concerns from at least two sources: our surname-to-country prediction, and the corresponding genetic distance measures obtained from the Spolaore and Wacziarg (2009) dataset. We also rely on an IV strategy to account for this type of the endogeneity concerns.

[18] FIFA added eligibility restrictions for players representing national teams in 1962: 1. Players must be naturalized citizens of the country they represent. 2. If a player is in a national team, he is ineligible to represent another nation. 3. Exceptions only matter if geopolitical changes in the countries occurred. See Hall (2012).

[19] This would have some variation on citizenship granting process that follows from the destination countries' law.

To construct our instrument, we use data on the ethnic composition of countries provided by the University of Illinois Cline Center for Advanced Social Research. The Composition of Religious and Ethnic Groups (CREG)[20] is a time-varying (since post-WWII) measure that involves country-specific information on 165 large countries. In the sample, ethnic groups are given narrow definitions (e.g. *Russian*, *Romanian*, *Scottish*), which we converted to a reference country. The classification "others" is used by the data provider to group information on one or more unknown ethnic minorities.

We build a measure of lagged country diversity, following the same diversity formula described above. We produce the following country-level index IV_{it} that we use for the country's team:

$$IV_{it} = \sum_{j=1}^{N_{t-18}} \sum_{k=1}^{N_{t-18}} (p_{jt-18} p_{kt-18} d_{jk}),\ j \neq k \tag{3}$$

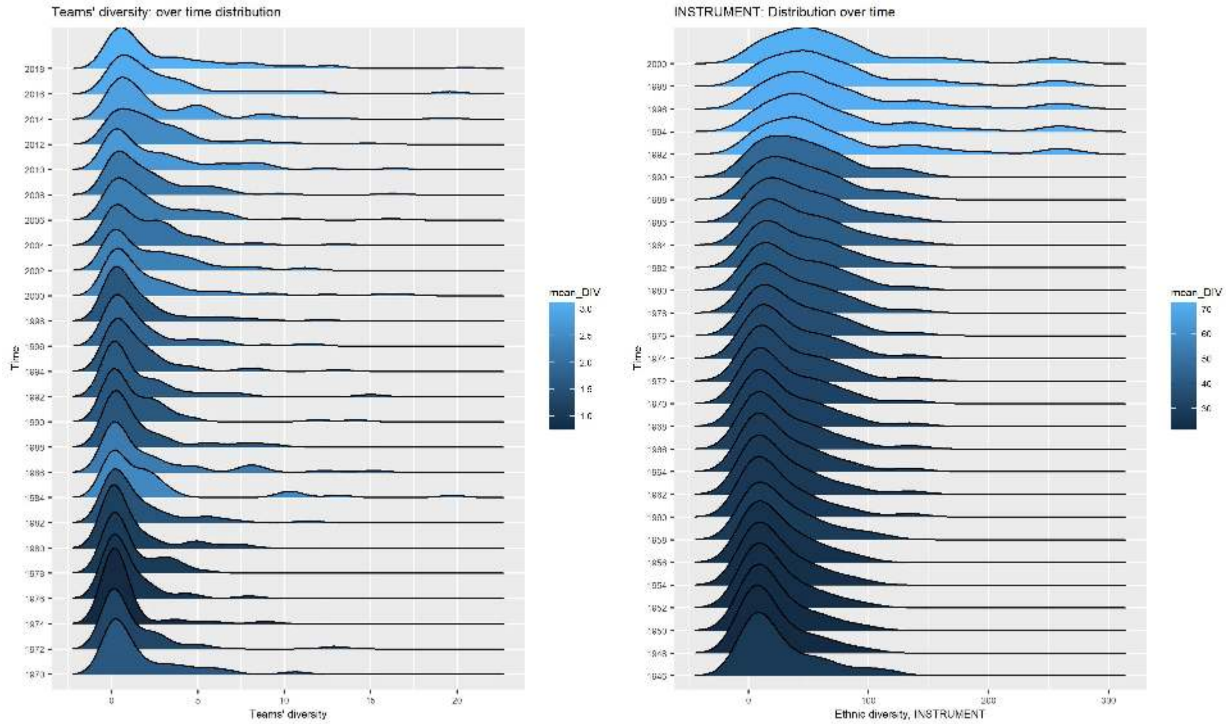
where p_{jt-18} and p_{kt-18} are shares of origins j and k immigration stocks, belonging to the set origins in country i at time $t - 18$. The instrument is used for the qualification of the final phase.

As a decision rule, the group "others" in country i was assigned a median distant country j from the Spolaore and Wacziarg (2009) dominant groups distance measure. The resulting variable was lagged to account for second-generation migration effects. While the lag choice is arbitrary, a higher gap would increase the data loss. For this reason, we use in our benchmark analysis an 18-year lag to limit the reduction in the final sample size, but 20-year and 22-year lags are also considered for sensitivity checking (see Section 6 below).

An inconvenience of the CREG dataset is that there are no data for a set of small countries (Kosovo, Malta, San Marino, Luxembourg, Montenegro, Faroe Islands), plus France and Iceland. To account for this issue, we complement the data with the World Bank's Global Bilateral Migration Database. For the years 1960-2000, this data source aggregates census and population register records, providing information at 10-year intervals. We interpolate these measures linearly for the missing countries to obtain two-yearly complementary information on our instrument. The resulting distributions are presented on the right in Figure 6 and are compared with the team diversity measure at left. The overall picture suggests a general increase in countries' ethnic diversity over time in the European continent (as displayed in the growing average values). However, this growth has been uneven across countries (as shown by the longer right tails). Although we formally assess the relevance of our instrument in the following sections, the patterns in the plots of Figure 6 seem broadly similar in the national teams' diversity and the ethnic diversity of the whole population.

[20]Cline Center for Advanced Social Research.

Figure 6. IV diversity over time



Notes: In Figure 6, we present the evolution of the distribution of diversity over time for our diversity index (on the left) and our IV index (on the right). Lighter colors represent higher yearly averages. This picture points to a positive evolution of national teams' diversity that is matched visually with a positive evolution in the lagged mean national diversity of our baseline instrument. This pattern is broadly in line with van Campenhout et al. (2019) who also suggest a growing trend in diversity occurring over time for the World Cup teams as a result of the countries' migratory histories and citizenship regimes.

4 Empirical analysis

We first carry out OLS estimations applied to the unilateral and bilateral settings in order to obtain the association between diversity and football performances. Since the estimations in these naive OLS regressions are likely confounded by some factors, we then move to the instrumental variables estimations to uncover a causal link between diversity and sports performance.

4.1 Benchmark estimations

Our benchmark unilateral estimation is as follows:

$$Performance_{ist} = \alpha + \alpha_s + \alpha_i + \alpha_t + \beta Div_{ist} + X_{it}'\Gamma + \epsilon_{ist} \tag{4}$$

where national team i performs in either or both stages s ={qualification, finals} of the two types of international tournaments, i.e., the FIFA World Cup and the UEFA Euro Cup in t $\in$ {1970, 1972, 1974, ...2016, 2018}.[21] We include stage, time, and team dummies $\alpha_s, \alpha_t, \alpha_i$ in all our specifications. Our regressor of interest is the level of genetic diversity Div_{ist}, computed as detailed in Equation (1). Vector X'_{it} includes the set of controls as explained in the previous section.

A non-negligible issue is that teams do not play the same number of matches and competitions due to the selection of teams participating in the final rounds. This is due to the specificity of the selection process of each competition for the final stage. First, by definition, teams not qualifying for the final round play a lower number of matches and competition. Second, some teams are or were automatically selected for the final stage. The host(s) of a tournament have always been exempted from the qualification stage in both types of competitions. Furthermore, up to a recent period, the title holder was also exempted from the qualification stage in the World Cup competitions.[22] In the first case, this out-selection process is directly linked with performance. To overcome such out-selection issues, our sample comprises the final scores of teams in both stages, whether they played or not in that stage. It follows that, if the team did not qualify for the next step or was the host of the competition, their scores will stay unchanged in those instances.

While fixed effects capture the effect of unobserved factors that are either constant over time or across countries, the set of covariates X_{it} arguably accounts for other unobserved factors. For instance, a country's financial resources may positively correlate with its national team's performance. At the same time, these resources may have acted as a pull effect for immigration, which would result in a higher level of diversity. We therefore include the log of GDP per capita and lagged immigration in our controls.[23] The demographic size of a country could also be linked to its diversity and the probability of having talented eligible players in every cohort.

In a separate specification, we allow for the inclusion of two further controls reconstructed from the match-level information, namely, the average diversity level and the average strength of the opponents. These two indicators permit us to better identify the effect of interest. First, we test whether the diversity of the adversaries was detrimental to the players' performance at the end of the championship. Second, we define the adversary's strength as the starting Elo score levels of the adversaries' pool, averaged across components. As the Elo scores capture the adversary's strength, a loss against a stronger team will be mitigated compared with a loss against a weaker opponent. While, for the sake of the competition, facing a more robust team may increase the chance of being eliminated, it also, in terms of score changes, is an opportunity to update the Elo score positively. These controls therefore allow a better establishing of the competition hierarchy by accounting for the variation in the Elo score due to a stronger opposition.

In the bilateral framework, we adopt the following specification:

$$Performance_{ijst} = \alpha_i + \alpha_j + \alpha_s + \alpha_t + \beta(Diversity_{ist} - Diversity_{jst}) + \Gamma X_{ijst} + \epsilon_{ijst} \quad (5)$$

[21] Note: The year itself of the event reveals which tournament is played, so there is no need for a tournament fixed effect.

[22] Before the 2006 competition in Germany, the title-holding country was exempted from the qualification stage. In the European championship, the title holder has always been required to play the qualification games.

[23] Note that this covariate allows one to isolate the role of diversity in past immigration flows in the instrumental variable from its direct impact on performance by, for instance, increasing the talent pool.

where the baseline performance indicator is the goal difference between team i and team j facing one another at stage s of championship t.

4.2 Endogeneity concerns

As explained above, specifications (4) and (5) are subject to potential endogeneity issues from omitted variables, affecting both the genetic diversity of the national squad and its performance. To account for these concerns, we adopt an instrumental variable strategy to yield a consistent estimate of the causal link. Equations (6) and (7) will therefore represent the first-stage regressions for the unilateral and bilateral specification respectively:

$$Div_{ist} = \alpha + \alpha_s + \alpha_i + \alpha_t + \beta IV_{ist} + X_{it}'\Gamma + \epsilon_{ist} \tag{6}$$

$$Diversity_{ijst} = \alpha_i + \alpha_j + \alpha_s + \alpha_t + \beta(IV_{ist} - IV_{jst}) + \Gamma X_{ijst} + \epsilon_{ijst} \tag{7}$$

5 Results

5.1 Unilateral estimations

The baseline findings from the unilateral specification are reported in Table 3, and all include heteroskedastic robust standard errors. The dependent variable for this set of outcomes is the Elo score change from the beginning to the end of the championship stage. Columns (1) to (4) gradually include covariates and reproduce panel model results without considering possible endogeneity concerns. Columns (5) to (8) show the IV results, where the instrument is the one-generation-lagged ethnic diversity of the population. Starting from the simple model that includes only age covariates, we add deviation in the team minute appearances as well as the log of GDP per capita, population, and lagged immigration stocks.

Our estimate of the effect of diversity is positive in all our specifications. Its significance varies between 5% and 10% in the OLS columns, whereas results from the IV specifications indicate a positive coefficient, significant at the 5% level.

While it is impossible to look at overidentification concerns with a single IV, the LM test and the Kleibergen-Paap Wald rk F test both suggest that our instrument is strong. As for the size of the effect, while OLS estimates present a coefficient of just below 3, the IV model indicates a coefficient ranging from $\sim$20 (in Column 6), to $\sim$ 32.2 (in Column 8). In terms of economic magnitude, a one standard deviation increase of the diversity measure translates into an increase in the Elo score change between 20 to 32.2. Given that the in-sample standard deviation of the Elo score change is about 40, the IV results suggest a change of approximately one-half to three-quarters of a standard deviation in this outcome for a one-standard increase in the deviation of genetic diversity. To illustrate the size of our results, let us consider a couple of examples. At the end of the 2018 World Cup finals, Portugal's Elo score was

1940, Croatia's 1943, Germany's 1964, and Spain's 2010. A change of 32 points in the Elo score would make Portugal outrank Germany, climbing two positions in this ranking.

The deviation in minutes appearances is positive, suggesting that the players' strategic turnovers seem to matter for the teams' performance. This might reflect the fact that teams with a broader pool of good players perform better. Demographic aspects, such as past immigration and population, are positive but not significant factors, while GDP per capita appears to be a significant positive driver of performance in the IV specifications, suggesting that countries with more resources perform better.

In all our OLS estimations, we report results with heteroskedasticity robust standard errors, clustered at team level; in our IV specifications, we display standard errors robust to arbitrary autocorrelation of order 1 and arbitrary heteroskedasticity. Finally, we report the sample size, together with the under-identification Kleibergen-Paap rk LM test statistic (idstat) and the weak identification Kleibergen-Paap Wald rk F test statistic (widstat). This second is the equivalent of the Cragg-Donald Wald F statistic for the case in which robust standard errors are used. As an alternative to our benchmark measure, the outcome of interest would involve taking the Elo score levels at the end of the championship stage (instead of the changes) and controlling for the initial score level. We perform this exercise in Table 4, and results are virtually unchanged.

5.2 Bilateral estimations

The baseline findings concerning the bilateral specification are in Table 5. They include robust standard errors, clustered at the match level. Team i is referred to as the home team and team j to the away team. (Note that, in the final stages, only hosting countries may play at home.) The dependent variable for this framework is the goal difference as we perform the analysis at match level. Similar to the previous section on unilateral estimations, the Table 5 presents panel results in the left panel (columns 1 to 5) where potential endogeneity concerns arise, and the IV results in the right panel (columns 6 to 10). Starting with the simplest specification that considers age covariates, results gradually control for variation in appearances, per capita GDP, population, and lagged immigrant stocks. Finally, Columns 5 and 10 add three gravity covariates at the bilateral level, namely, (current or historical) contiguity, sharing a common language, and belonging to the same country at some stage in time.[24] The significance of the coefficients is in line with those of the unilateral framework. Diversity is positive but not always significant in the OLS specifications (columns 1 to 5), while it becomes significantly positive at 5% level in all IV specifications. As we would expect, home team controls have either opposite signs compared with their away team counterpart or no significant role. Past immigration stocks, when significant, increase the relative team performance, suggesting an effect related to the enlargement of the talent pool. Although its significance drops in some specifications, GDP per capita is a positive determinant of performance, reflecting that teams from richer countries can benefit from better resources, which in turn improve performance.

Concerning the economic magnitude of our coefficient of interest, in the IV specifications, a one-standard-deviation

[24]Note that, due to a historical agreement in the early phase of international football, the four main regions of the U.K. (England, Scotland, Northern Ireland, and Wales) compete as separate teams.

increase in the diversity measure leads to an increase in the goal difference of between 0.7 to 1.4 units. While we address some specification concerns in the next paragraph, the evidence from the baseline results seems much in line with the unilateral framework.

6 Robustness checks

In the following sections, we conduct a number of sensitivity exercises to assess the impact of our methodological options in the benchmark estimations. We first consider the robustness checks in the unilateral setting and then move to the bilateral framework.

6.1 Sensitivity checks in the unilateral analysis

To evaluate the sensitivity of our unilateral results, we conduct a set of robustness checks. We first introduce further controls in the unilateral regressions. We then check the robustness of the results obtained with our benchmark diversity measure. We further analyze how much our findings change if we highlight the coach's role by including controls at the level of the team's manager. Finally, since our principal analysis focuses on European teams, we assess the internal and external validity. We adjust the Elo score to also consider intercontinental matches in the unilateral analysis in order to exclude the influence of matches with non-European teams.

6.1.1 Additional Controls

In the baseline estimation, we introduce two additional covariates of interest measured at the match level. The results are in Table 6. Specifically, we add the average adversary diversity and the average adversary strength measured by their average Elo score levels. In the regressions, we gradually add controls from left to right. In Column 5, we include these two covariates jointly. The IV results are in line with those in the benchmark regressions. The adversary's diversity is, in general, negatively correlated with the team's performance. Adversary's strength appears to impact the Elo score change positively. Nevertheless, this result likely comes from the score construction, which specifically gives weight to the strength of the adversary.

6.1.2 Checks on the diversity measure and IV

In Table 7, we perform a series of sensitivity checks regarding the diversity measure. The first three columns report the same results of Table 3 using an alternative diversity measure weighted by each player's minute appearance. The alternative diversity index, denoted $Divalt_{ist}$ takes the following form:

$$Divalt_{ist} = \frac{1}{S_t}\sum_{j=1}^{N_t}\sum_{k=1}^{N_t} pAPP_{jt}pAPP_{kt}d_{jk},\ j \neq k$$

where $pAPP_{jt}, pAPP_{kt}$ are the shares of minute appearances of origin j and k respectively, belonging to the set of origins $\{1,...N_t\}$ in team i for stage s of championship t. As for our baseline index, we normalize this expression by a team size factor S_t and include genetic distance d_{jk}. By giving more weight to the most active players, this alternative measure allows us to harmonize the size of a team when computing its diversity, excluding players who were listed but never called on the pitch. A discrepancy between these results and those from the benchmark regressions would possibly indicate whether players' diversity matters in the training stages rather than at the competition level itself. As in Table 7, the outcomes are virtually unchanged: the diversity coefficients are positive and in line with the previous results. This suggests that our findings are relevant at the competition level.

Column 5 of Table 7 checks sensitivity to the use of Spolaore and Wacziarg (2009)'s baseline genetic distance, based on the majority ethnic groups, against their alternative indicator. The latter adjusts genetic distances with (time-invariant) data from Alesina et al. (2003) on ethnic group proportions. Crucially, this adjustment results in some missing values for a minority of country pairs ij, absent in Alesina et al. (2003)'s dataset. The two measures are highly correlated, and Spolaore and Wacziarg (2009)'s results are not sensitive to this alternative measurement. Important for our diversity computation, completed cases in the country pairs ij are key to our diversity index construction (we have fewer missing values for our computation of the diversity measure). In Column 5, however, we present results from using this alternative distance indicator. In terms of controls, this specification is comparable to the results in columns (3) and (7) of Table 3. These columns represent a suitable comparison in that they display the most conservative results in Table 3. Results in Column (5) indicate a coefficient of about 1, significant at the 10 % level. While this suggests some sensitivity with respect to the chosen genetic distance measure, the coefficient of diversity remains significantly positive.

Finally, the last two columns increase the lag of the instrument from 18 years to 20 (Column 6) and 22 (Column 7). The results are again comparable to columns (3) and (7) in Table 3. While the data availability is reduced for the instrument once the year gap is increased, neither the decrease in the sample size nor the different time gap affect the main results. The coefficient on diversity is slightly smaller than in the baseline (the coefficient of interest is 18.6 when considering a 20-year lag, and 14.8 with a 22-year lag). Here again, Kleibergen-Paap Wald rk F test and the LM test statistic support the relevance instrument in terms of strength.

6.1.3 Controlling for coach quality

We further test the robustness of genetic diversity's positive effect on football performance by adding control variables that involve information on the team managers. It could be argued that coaches of high quality would also favor higher genetic diversity because they anticipate its benefit on the performance. Failure to account for coach quality could, at least theoretically, confound the effect of genetic diversity on team's performance. To account for such an effect, we supplement our set of controls with variables capturing the quality of the manager of national teams.

We retrieve the information on the team manager from the same source used to capture the national teams' squad composition. More specifically, we collect three pieces of information on the person reported as *Manager* in the squad list: age; nationality; and a measure of previous experience, defined as how many prior UEFA/FIFA Championships are listed in the coach's career details. We approximate the age by the difference between the Cup year and the birth year. Furthermore, we create a dummy capturing whether the coach's nationality is different from that of the national team or not. The use of a foreign coach is clearly a measure of quality since countries have a natural bias to choose a native coach for managing their national squad. In a small number of cases where information is missing from the source, some information was added manually if available via other sources.[25]

We also construct a measure of coach quality based on coach awards. We consider two awards: the European Football Coach of the Season, and the European Football Coach of the Year. These awards are annual prizes organized by European Press or technical entities (depending on the year, the European Union of Sports Press, Association of European Journalists, UEFA, Technical Commission of Torneo di Viareggio). We extract the winner's name information from Wikipedia[26] and set a dummy equaling 1 whenever (after spelling checks) the winner of these awards was a manager included in our sample.

We add this set of controls in the sensitivity checks above and present the results of these estimations in Table 8. The significance of the diversity variable (first row of the table results) remains unchanged, and the point estimate of the coefficient is strongly comparable to the benchmark estimates. We find coach variables to be weak predictors of the Elo score-based performance. The coach's past experience is significant in only two specifications, and its effect sign is negative. The foreign coach dummy is associated with a positive coefficient, albeit not always significant in a subset of specifications.

6.1.4 The European Tournament: internal and external validity

Our analysis focuses on teams affiliated with the UEFA, the European authority of football. This choice was motivated mainly by the availability of data concerning the composition of squads. This implies some restrictions on the sample of countries that we consider. Thränhardt (1992) documents how Europe has become an immigration continent in the recent decades and details how these flows present across time as well as cross-country variations. This motivates our focus on Europe, as well as our long panel construction. In the genetic diversity and conflict literature, Arbatlı et al. (2020) limit the geographic baseline coverage of their study to Europe, Asia and Africa to maintain low levels of admixture in distant national populations. Arbatlı et al. (2020) and previously Ashraf and Galor (2013) identify the distance from Addis Ababa in East Africa as a strong predictor of the historical degree of genetic diversity of a national population.

Therefore, the restrictions in our choices are also motivated by comparability concerns in terms of the genetic diversity of the native population. We build our diversity measures so that two team members from the same surname-predicted origin do not contribute to the overall team's genetic diversity. This assumption is likely to be restrictive in the New World if surnames' admixture follows genetic admixture. Further, European countries' geographic proximity allowed us to maintain greater comparability of the indigenous groups' levels of genetic diversity.[27] This likely implies that the national teams across countries' genetic diversity levels mostly result from immigration patterns rather than

[25] For coaches whose age was not found, we approximate the age as the year average. The exact information is missing for the following coaches: Andreas Lazaridis, Guentcho Dobrev, Ilia Shuke, José Gomes da Silva, Takis Charalambous, and Tony Formosa.

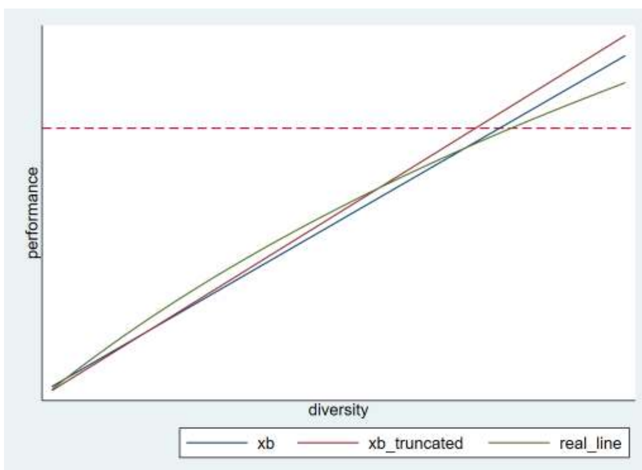
[26] Wikipedia, "European Football Coach of the Season," "European Football Coach of the Year." https://en.wikipedia.org/wiki/European Football Coach of the Season, https://en.wikipedia.org/wiki/European Football Coach of the Year

[27] Note the low levels of genetic distances for the European Continent table in Cavalli-Sforza et al. (1994).

genetic variation in the indigenous population. This said, we acknowledge that extending the same analysis to other Continents may produce technical difficulties, and we cannot conclude that external validity concerns do not arise.

In the World Cup final stage, qualified UEFA teams meet finalists from UEFA and from other continents. These intercontinental matches are not considered in the baseline specifications, as we do not have data about the opponents for those games. Conceptually, this sample exclusion is likely to be positively correlated with performance, in the sense that this circumstance would arise only with World Cup finalists. On the other hand, diversity coefficients would be biased by this exclusion if, given the controls, diversity was somehow correlated with the probability of excluding the team match. For instance, this could happen in case of a misspecification of the linearity of the effect. Suppose there is a marginally decreasing effect of diversity on performance (green curve of Figure 7), for which the linear model that we use constitutes a linear approximation (blue line). The truncation of high performing teams' matches in the World Cup finals can lead to an upward bias (see the higher slope from the red line) in the estimated effect of diversity.

Figure 7. Upward bias in the effect of diversity from misspecification and sample composition.



Notes: Simulation example: 2000 observations. Data generating process for y: 0.9*log(0.6*x+0.2)+1.1. The xb line represent the slope from a regression of x on y. The xb_truncated line represents the slope from a regression of x on y, in a subset of observations for which y is below .8 (approximately 83% of the initial sample observations). This graph shows a higher slope resulting from this top-truncation.

We do not find any evidence of a possible quadratic effect of diversity in our OLS specifications.[28] In our Unilateral specification, we can further test whether our outcomes are preserved if the performance measure also includes matches against inter-continental teams. We reproduce the unilateral results of Table 3, but we replace the outcome variable Elo' score change that we compute, with the equivalent measure directly obtained from *eloratings.net*. Crucially, in this alternative outcome variable, all championship matches are included. Concerns would arise if our results of this alternative set of regressions were not in line with the baseline ones, as this could possibly indicate an upward bias in the bilateral setting. As Table 9 reflects, results are virtually unchanged by this modification. The coefficient for diversity ranges from roughly 17 to 27 in the IV, while it is roughly 3.5 in the OLS.

[28] Results are not presented here in the interest of space but are available upon request.

6.2 Sensitivity checks in the bilateral estimations

Similar to the unilateral analysis, we perform a series of robustness checks in our specifications used at the match level.

6.2.1 Additional controls

To better assess the match-level dimension of this result, we add the relative teams' initial Elo score levels as a control. In Table 10, we present results that complement the previous outcomes with this additional control. As one could expect, initial scores of the teams are positive predictors of their relative performance. Nevertheless, the effect of diversity remains significantly positive in all the IV-based results, as in the benchmark. This suggests that diversity has a distinctive role in performance during the match, and that positive skill complementarity is involved in the team's coordination.

6.2.2 Alternative regression methods

We assess the robustness of several methodological choices made in the regression analysis within the bilateral framework. We first carry out some sensitivity checks with respect to the way performance is measured. We follow a comparable approach to the unilateral specifications. Columns (1) to (4) of Table 11 report the results of, respectively, a specification where diversity is replaced with its appearance's re-weighted measure; a regression where diversity is computed with the alternative genetic distance measure as proposed in Spolaore and Wacziarg (2009); and of two specifications with alternative lags for our instrument (20 years and 22 years). The resulting coefficients are very comparable to the baseline evidence.

A second check concerns the use of linear regression models. Since the goal difference is a discrete variable (ranging between –13 and 12), the linear models may become less appropriate as they assume a continuous variable. We address this concern in two different ways. First, we perform an inverse hyperbolic sine (IHS) transform to the variable. This type of procedure has been proposed in the literature by Burbidge et al. (1988) as an alternative to the log transformation. Indeed, such a transformation allows for the inclusion of variables that take zeroes and negative values, while maintaining approximatively the same interpretation of the coefficient as the log form. Second, we conduct a Poisson-based regression with scores as our outcome of interest. Results from Column (5) suggest that a hyperbolic sine transformation does not lead to different outcomes in the results of interest: a positive coefficient for the diversity measure of roughly 0.57 is significant at 1% level. Column (6) of Table 11 estimates a linear model on the number of goals at home as dependent variable. The coefficient on diversity is significant at 5% level and approximately 0.56.

Finally, an alternative specification on the bilateral diversity is proposed. Instead of the benchmark bilateral diversity measure corresponding to the difference ($Diversity_{home} - Diversity_{away}$) we allow the two terms to enter separately, allowing for the presence of a different effect for the home team and away team. Each term is instrumented (resulting in two first stages). Results indicate coefficients with opposite signs: the goal difference is, as expected, impacted positively by the home team and negatively by the adversary, with significance at a 5% level.

6.2.3 Number of goals as performance indicator

We further check the robustness of our results by using the number of goals scored or taken as alternative measures of the team's performance. We accommodate the discrete and non-negative nature of such an outcome by using a count data model estimated by Poisson. To account for endogeneity concerns, we use a control function approach (see for instance Lin and Wooldridge (2019) for a discussion of the relevance of this approach). Table 12 presents average marginal effects of diversity on the two teams' outcomes considered separately. In Appendix E, we also present the estimated coefficients from the structural equation and the results of the first stages obtained through this approach.

Our dependent variable is the number of goals made by the home team in one set of regressions and by the away team in a second set of regressions. Such an outcome is a discrete and non-negative count variable, which encourages our choice for a Poisson second stage. We follow the procedure suggested by Lin and Wooldridge (2019) and use the two diversity measures separately, as the outcome variables are also team specific.

We have two separate first stages, one for each of the two variables of interest (diversity_i and diversity_j). We include team_i plus team_j fixed effects and time fixed effects. We bootstrap standard errors in both stages with 2000 repetitions, and cluster them at the ij pair (unordered, i.e., ij=ji). The control function approach plugs the residuals of the first stage into the second, rather than the fitted values. This conveniently avoids inserting estimated fixed effects in the second equation, which is of exponential form for our specification.

We standardize our regressors of interest to simplify the interpretation of the partial effects and present average marginal effects (AME) in Table 12. The full table of coefficients is found in Appendix E for the sake of completeness (Table 23 and 24). We maintain the same five different sets of controls to compare with the benchmark.

The AME results suggest an effect broadly in line with our previous findings. As the top part of Table 12 displays, the diversity of the home team (respectively, away team) when the effect is significant and positively (respectively, negatively) affects its performance. The diversity of the opponent negatively affects it. The expected goal count increases from 0.43 to 0.52 (columns 1 to 3) for a given increase of a standard deviation increase in the home team diversity, while it decreases by roughly the same amounts, from 0.375 to 0.63 (columns 3, 4, and 5) for a given increase of the away team diversity. Results are broadly similar in the away score specifications, shown at the bottom of Table 12. In this specification, however, it is only the relative team's diversity that significantly (and positively) affects its performance.

6.2.4 Controlling for coach information

Akin to the sensitivity checks performed in the unilateral regressions, we further test the robustness of our bilateral results by adding variables capturing information about the teams' managers. Table 13 documents the results. As in the unilateral framework, we control for age, tenure, a foreign nationality dummy, and a measure of coach quality for both the home and the away team. In this set of regressions, the away team's foreign coach dummy is positively associated with the away team performance, as is the coach quality measure (based on awards). Significantly, a positive coefficient associated with our diversity measure is maintained as significant in our IV regressions.

7 Placebo analysis

As a final analysis assessing the validity of our results, we perform a placebo analysis using national performances from athletics as the outcome variable. Since the main channel explaining the positive impact of genetic diversity goes through the complementarity of skills at the team level, we should expect that genetic diversity does not play any role in explaining the performances at the individual level. Athletics is an accessible and mostly individual sport. We therefore assume the national pool of talent that athletics federations can rely on is comparable to that of football. If the placebo analysis returned significant coefficients of the football team's diversity index on athletics performance, we might have concerns that some omitted variable—such as the presence of a particular set of origins—would positively affect the national talent pool and our performance outcome. This mechanism may go beyond the size of lagged immigration, which we control for.

For the sake of this analysis, we extract information from Wikipedia about the total number of medals and gold medals won by each nation in the European Athletics Indoor Championships[29] and the European Athletics Outdoor Championships,[30] The European Athletics Outdoor Championships is an athletics event that started in 1934 with a quadrennial frequency until 2010 when it switched to a biennial frequency.[31]
The number of athletes that each national federation can enroll in any of these championships is based on their performance and is capped from above for each nation and discipline.[32]

As noted above, we collect information on the number of medals each nation won in each championship. To match these data with our original biennial data of football events, we consider athletics championships held in year t (if t is an odd year) as having been held in $t+1$. Whenever we have more than one event in the same year, we average the total medals won by a nation by year. We therefore obtain two indicators of athletic performance at national level: the number of total medals obtained by the national representatives, and the number of gold medals. The results of the placebo exercise are reported in tables 27, 28 and 29. Specifically, Table 27 serves as a direct comparison and presents results from our regressions on our benchmark outcome. In these tables, different from our baseline, we control for population, as we deem it an essential covariate for our athletics-based tables. Table 28 shows placebo results when the dependent variable is the number of total medals; Table 29 presents results when the dependent variable is the number of gold medals.

Coefficients of our diversity score in tables 28 and 29 turn out to be insignificant, suggesting that diversity in football teams does not impact the performances of an individual sport such as athletics. All in all, this strengthens the case of a positive impact of genetic diversity through its impact in terms of complementarity of skills.

[29] Wikipedia, "European Athletics Indoor Championships." https://en.wikipedia.org/wiki/European Athletics Indoor Championships.

[30] Wikipedia, "European Athletics Championships." https://en.wikipedia.org/wiki/European Athletics Championships.

[31] It is organized by the European Athletics Association (EAA), which is the continental committee of the worldwide International Association of Athletic Federations (IAAF). EAA is based in Switzerland (as are the UEFA and FIFA) and comprises 51 national associations (or members). EAA also organizes the European Athletics Indoor Championships, now a biennial event, but its frequency was yearly until 1990. A gap of three years passed between 2002 and 2005's tournaments.

[32] European Athletics, "Competition regulations," https://european-athletics.com/competition-regulations/.

8 Conclusion

Diversity is a double-edged sword. Greater diversity is beneficial in teamwork since teams can draw on a larger variety of skills and knowledge from a diverse group of people. However, diversity might also lead to decreased team performance and team effectiveness if more diversity brings lack of coordination and increased conflict. In this paper, we assess the effect of genetic diversity, due to past migration flows, on sport performance. To do so, we have built a new dataset that brings together information about the genetic diversity of European national football teams playing in the World Cup or European Cup, qualifications and finals, and several time-varying performance indicators for each national team. Genetic diversity of players may lead to a lack of team spirit on the one hand but, on the other hand, may lead to innovative ways to play. In addition, it is well known that some football-specific skills (e.g., endurance capacity, muscle performance, height, or technical skills) are related to genetic background (see Lippi et al. (2010)). Therefore, genetic diversity boosts complementarities among players holding different positions on the football team. Hence, overall, we expect genetic diversity to improve sportive performance. This is confirmed in our analysis. We establish a positive causal relationship between this measure of team diversity and both a team's Elo score and the probability of winning a match. We also prove that this diversity benefits teams beyond any effect stemming from population size, GDP per capita, coach experience, and other factors. The result is quite large and not negligible. Analyzed using a variety of perspectives, and taking into account endogeneity and measurement error concerns via an instrumentation method, the overall evidence produced in our specifications strongly suggests that diversity enhances performance at match level—as proxied by the goal difference—and translates into higher overall team (Elo) scores at the end of the championship.

Our findings complement the flourishing but limited literature on countries' genetic diversity and corresponding effects. Our contribution is a novel one as it focuses on the sports team. The results are robust to a large list of checks where we use variation of the diversity measure and of the instrument. We also perform a placebo test to rule out any remaining concerns about some omitted variable, such as the presence of a particular set of origins that would positively affect the national talent pool and our national team diversity. In the placebo test, we show that, as expected, genetic diversity does not affect the performance of national athletics teams because each athlete competes individually rather than on a team.

Our study is not intended to be a biological one. We examine the effect on performance today of deep-rooted values and traits shaped thousands of years ago. These characteristics and the associated information they bear, proxied by genetic scores, cannot be captured (or measured) by simple country fixed effects or other cultural and institutional characteristics formed in humanity's more recent history. It is important to stress that our results do not carry any implications in terms of superiority or inferiority of particular genetic endowments of specific origins over other ones. Rather, our interest is on the genetic diversity among the players on a team and how these differences translate into a comparative advantage at the *team level* in sportive performance and innovative play. We find, in fact, that different deep-seated factors embodied by the genetics scores do matter.

To conclude, our work highlights a less evident, yet relevant, effect of the mixing of populations worldwide due to international migration. The effects of these population movements have attracted an impressive amount of economic literature interested in the economic as well as cultural effects of migration in the destination and origin countries. Further research in this field shall extend our analysis to larger geographical areas and also to other sports played in teams.

9 Tables section

Note: additional tables are presented in appendices D and E.

Table 1. Summary statistics table, unilateral framework

	Mean	Standard Deviation	N	Min	Max
Elo score changes, computed	.2977382	40.12878	1900	-216.0751	181.1425
Elo score, computed	1671.196	234.233	1900	873.0499	2157.986
Performance measures					
Elo score	1692.84	232.8866	1900	852	2223
Elo score changes	.8968421	45.75327	1900	-233	217
Diversity measures					
Diversity	.0327621	1.027647	1900	-.6808419	6.777219
Diversity, appearance	.0307232	1.02827	1900	-.6285349	8.371243
Diversity, SW	2.257027	2.557093	1900	4.07e-09	19.57592
Team level variables					
Adversary's diversity	.0071005	1.008427	1900	-1.125619	10.04668
Adversary's strength	1700.037	117.0431	1900	1411.171	2140.289
Foreign coach	.1763158	.3811889	1900	0	1
Coach age	50.66891	8.08903	1900	28	74
Coach award	.0826316	.275397	1900	0	1
Stand. dev. appearances	191.6779	65.7022	1900	59.59386	451.4403
Stand. dev. squad age	3.610854	.6359852	1900	1.953301	6.213847
Squad age	27.55415	1.004444	1900	24.28572	31.04545
Squad age, squared	760.2397	55.36567	1900	589.796	963.8202
Squad size	30.09105	8.048669	1900	15	59
Macroeconomic variables					
Log immig. stocks, 18y lag	12.44378	2.622838	1676	0	16.294
Log of GDP/capita	9.668765	1.03384	1900	6.836052	11.58409
Population (mln)	23.25316	27.97488	1900	.024766	148.336
IV					
IV, 18y lag	40.08942	40.92574	1900	0	206.7919

Notes: The unilateral specification involves a dataset of national teams appearing once for each stage of the tournament, World Cup or EURO Cup, for each of the years considered. Note that the team appears as an observation in this dataset even when it did not participate in that stage (qualification or final) to avoid dropping its information. When a country did not qualify for the finals, the levels of the explanatory variables will be equal to those at the end of the qualification stage. Similarly, its Elo score will be unchanged.

Table 2. Summary statistics table, bilateral framework

	Mean	Standard deviation	N	Min	Max
Performance measures					
Goal difference	.4818158	2.068111	3877	-8	11
Goal difference, hyperbolic sine	.2760175	1.222668	3877	-2.776472	3.093102
Diversity measures					
Bilateral diversity	.0019554	1.043691	3877	-5.382044	5.568716
Bilateral diversity, appearance	.0005331	1.041595	3877	-6.740676	5.36406
Bilateral diversity, SW	.0014641	1.038141	3877	-5.003557	5.647373
Diversity, home	.0510492	1.046821	3877	-.6975578	6.850133
Diversity, away	.0497902	1.042281	3877	-.7024601	7.061377
Team level variables					
Stand. dev. squad age, home	3.685712	.883574	3877	1.953301	13.27831
Squad age, home	27.65924	1.030211	3877	24.28572	31.04545
Stand. dev. squad age, away	3.683282	.8835517	3877	1.953301	13.27831
Squad age, away	27.67031	1.027588	3877	24.28572	31.04545
Squad age, squared, home	766.0945	56.99869	3877	589.796	963.8202
Squad age, squared, away	766.702	56.8858	3877	589.796	963.8202
Stand. dev. appearances, home	236.2873	63.79323	3877	59.59386	451.4403
Stand. dev. appearances, away	235.3479	64.33492	3877	67.74986	451.4403
Foreign coach, home	.1761671	.3810114	3877	0	1
Foreign coach, away	.1779727	.3825391	3877	0	1
Coach age, home	51.12574	8.117754	3877	28	74
Coach age, away	51.19401	8.117002	3877	28	74
Macroeconomic variables					
Population (mln), home	24.49271	29.01123	3877	.223537	148.689
Population (mln), away	24.26327	29.02497	3877	.223537	148.689
Log of GDP/capita, home	9.691932	1.028319	3877	6.836052	11.58409
Log of GDP/capita, away	9.681962	1.033324	3877	6.836052	11.58409
Log immig. stocks, 18y lag, home	12.66511	2.460724	3576	0	16.294
Log immig. stocks, 18y lag, away	12.63034	2.459263	3576	0	16.294
Adversary's strength, home	1674.52	108.3997	3877	1416.287	2117.771
Adversary's strength, away	1673.245	111.0915	3877	1400.824	2140.289
Contiguity	.0949188	.2931405	3877	0	1
Same nation	.0211504	.1439041	3877	0	1
Common language	.0502966	.2185845	3877	0	1

Notes: The Bilateral specification involves a dataset of matches held in the qualification and final stages of the EURO or World Cup, where both adversaries belong to the UEFA affiliation.

Table 3. Football performance and ethnic diversity of national teams: unilateral estimations

	Dependent variable: change in rating of national football team (Elo score)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Variable of interest								
Diversity	2.974^{**}	2.673^{**}	2.807^{**}	2.695^{*}	22.288^{**}	20.814^{**}	23.652^{**}	32.202^{**}
	(1.142)	(1.156)	(1.129)	(1.387)	(10.567)	(10.273)	(11.655)	(16.299)
Control variables								
Stand. dev. appearances		0.276^{***}	0.277^{***}	0.310^{***}		0.291^{***}	0.293^{***}	0.318^{***}
		(0.023)	(0.023)	(0.028)		(0.026)	(0.027)	(0.032)
Log of GDP/capita			8.814	9.701			18.034^{**}	16.642^{*}
			(6.494)	(7.755)			(8.061)	(8.983)
Population (mln)			0.322				0.140	
			(0.208)				(0.414)	
Log immig. stocks, 18y lag				-0.496				1.356
				(1.371)				(1.781)
Observations	1900	1900	1900	1676	1900	1900	1900	1676
Kleibergen-Paap LM test					0.00	0.00	0.00	0.00
Kleibergen-Paap F test					20.34	20.31	17.47	11.83
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Baseline estimates for the unilateral framework. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 5–7) / years 1978–2018 (in columns 4 and 8). Dependent variable: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 4. Football performance and ethnic diversity of national teams: alternative measure of rating

	Dependent variable: Ending rating of national football team (Elo score)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Variable of interest								
Diversity	2.906**	3.372**	3.445**	3.869**	26.615**	25.199**	23.073**	38.367**
	(1.238)	(1.089)	(1.078)	(1.185)	(10.643)	(10.404)	(11.130)	(16.987)
Control variables								
Elo's inital levels, computed	0.869***	0.872***	0.871***	0.854***	0.871***	0.869***	0.868***	0.852***
	(0.013)	(0.012)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)	(0.016)
Stand. dev. appearances		0.139***	0.140***	0.134***		0.295***	0.297***	0.315***
		(0.016)	(0.016)	(0.017)		(0.026)	(0.025)	(0.032)
Log of GDP/capita			8.002	3.930			17.064**	12.863
			(6.424)	(7.929)			(7.740)	(9.251)
Population (mln)			0.607*				0.502	
			(0.349)				(0.399)	
Log immig. stocks, 18y lag				1.380				3.425*
				(1.637)				(1.842)
Observations	1900	1900	1900	1676	1900	1900	1900	1676
Kleibergen-Paap LM test					0.00	0.00	0.00	0.00
Kleibergen-Paap F test					20.54	20.51	17.44	11.89
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates for the unilateral framework. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 5–7) / years 1978–2018 (in columns 4 and 8). Dependent variable: Elo score levels of the national team (end of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 5. Goal difference and genetic diversity: bilateral estimations

	Dependent variable: goal difference									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV
Variable of interest										
Bilateral diversity	0.091*	0.034	0.031	0.035	0.035	0.720**	0.969**	0.942**	1.414**	1.408**
	(0.047)	(0.044)	(0.044)	(0.045)	(0.045)	(0.297)	(0.342)	(0.299)	(0.512)	(0.511)
Control variables										
Log of GDP/capita, home		0.136	0.137	-0.010	-0.013		0.465*	0.459*	0.145	0.142
		(0.195)	(0.195)	(0.208)	(0.208)		(0.243)	(0.235)	(0.258)	(0.257)
Log of GDP/capita, away		-0.729***	-0.727***	-0.471**	-0.478**		-1.111***	-1.097***	-0.758**	-0.763**
		(0.198)	(0.198)	(0.214)	(0.214)		(0.252)	(0.242)	(0.285)	(0.285)
Population (mln), home			-0.005*					-0.002		
			(0.003)					(0.003)		
Population (mln), away			0.000					-0.003		
			(0.003)					(0.003)		
Log immig. stocks, 18y lag, home				0.068	0.067				0.154**	0.153**
				(0.044)	(0.044)				(0.062)	(0.061)
Log immig. stocks, 18y lag, away				-0.097**	-0.097**				-0.158**	-0.158**
				(0.046)	(0.046)				(0.057)	(0.057)
Observations	3877	3877	3877	3568	3568	3877	3877	3877	3568	3568
Kleibergen-Paap LM test						0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test						31.46	24.52	32.29	14.48	14.49
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minute appearances		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Geo-political controls					Yes					Yes

Notes: Baseline stimates for the bilateral framework (match-level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970-2018 (for columns 1-3,6-8) / years 1978-2018 (in columns 4-5 and 8-9). Dependent variable: Goal difference. In all regressions, we include Team and year fixed effects, as well as a stage dummy. Column 1 to Column 5 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at Team pair level. Column 5 to Column 8 display IV results, with heteroskedastic robust standard errors in parentheses, clustered at Team pair level. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F-test for weak instruments. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 6. Football performance and genetic diversity: further controls

	Dependent variable: change in rating of national football team (Elo score)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV
Variable of interest										
Diversity	2.974^{**}	3.247^{**}	2.288^{*}	2.136	2.219	22.288^{**}	27.755^{**}	23.339^{*}	32.172^{*}	33.295^{*}
	(1.142)	(1.067)	(1.247)	(1.532)	(1.515)	(10.567)	(12.685)	(11.931)	(17.383)	(17.595)
Control variables										
Log of GDP/capita		6.968	8.760	9.752	9.472		17.415^{**}	18.029^{**}	16.641^{*}	16.461^{*}
		(6.770)	(6.502)	(7.787)	(7.792)		(8.464)	(8.046)	(8.988)	(9.059)
Population (mln)		0.219	0.304				0.004	0.132		
		(0.232)	(0.208)				(0.422)	(0.411)		
Adversary's diversity		-1.326			-1.671		-2.664^{**}			-2.308^{*}
		(1.258)			(1.191)		(1.293)			(1.242)
Adversary's strength			0.022^{**}	0.021^{**}	0.026^{**}			0.011	0.000	0.006
			(0.008)	(0.009)	(0.009)			(0.014)	(0.020)	(0.019)
Stand. dev. appearances			0.280^{***}	0.313^{***}	0.311^{***}			0.296^{***}	0.318^{***}	0.316^{***}
			(0.023)	(0.028)	(0.028)			(0.027)	(0.034)	(0.034)
Log immig. stocks, 18y lag				-0.596	-0.516				1.353	1.529
				(1.395)	(1.397)				(1.853)	(1.882)
Observations	1900	1900	1900	1676	1676	1900	1900	1900	1676	1676
Kleibergen-Paap LM test						0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test						20.34	16.93	16.72	10.27	10.05
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates for the unilateral framework. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 6–8) / years 1978–2018 (in columns 4–5, 8–9). Dependent variable: Elo score levels of the national team (end of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 7. Football performance and genetic diversity: robustness checks

	Dependent variable: change in rating of national football team (Elo score)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV: Diversity, appearance	**IV:**Diversity, appearance	**IV:**Diversity, appearance	**IV:**Diversity, appearance	**IV:**Diversity, SW	**IV:** 20 years lag	**IV:** 22 years lag
Variable of interest							
Diversity, appearance	25.715^{**}	23.998^{*}	29.340^{*}	33.358^{*}			
	(12.887)	(12.505)	(15.426)	(18.091)			
Diversity, SW					10.975^{*}		
					(5.806)		
Diversity						18.603^{*}	14.854^{*}
						(9.675)	(8.017)
Control variables							
Stand. dev. appearances		0.294^{***}	0.296^{***}	0.324^{***}	0.292^{***}	0.306^{***}	0.309^{***}
		(0.027)	(0.028)	(0.033)	(0.028)	(0.027)	(0.027)
Log of GDP/capita			20.526^{**}	15.058^{*}	23.217^{**}	11.060	8.896
			(9.179)	(8.917)	(10.208)	(7.583)	(7.497)
Population (mln)			0.052		0.056	0.118	0.109
			(0.477)		(0.454)	(0.380)	(0.404)
Log immig. stocks, 18y lag				2.388			
				(2.282)			
Observations	1900	1900	1900	1676	1900	1784	1670
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	15.33	15.29	12.29	9.95	9.88	20.54	24.29
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates for the unilateral framework. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (columns 1–3, and 5) / years 1978–2018 (in Column 4) first year available for the instrument to 2018 (columns 6 and 7). Dependent variable: Elo score levels of the national team (end of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 8. Football performance and ethnic diversity: accounting for coach information

	Dependent variable: change in rating of national football team (Elo score)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Variable of interest								
Diversity	3.079**	2.741**	2.855**	2.800**	24.766**	23.279**	25.348**	34.327**
	(1.168)	(1.154)	(1.120)	(1.378)	(10.611)	(10.327)	(11.529)	(16.339)
Control variables								
Coach age	-0.111	-0.130	-0.119	-0.170	-0.053	-0.073	-0.043	-0.056
	(0.172)	(0.160)	(0.165)	(0.159)	(0.172)	(0.164)	(0.167)	(0.197)
Coach tenure	-0.963	-0.898	-0.990	-0.959	-1.123	-1.019	-1.162*	-1.311*
	(0.718)	(0.669)	(0.675)	(0.686)	(0.708)	(0.672)	(0.678)	(0.785)
Coach award	4.294	2.188	2.112	3.372	5.105	2.938	2.915	4.409
	(4.583)	(4.653)	(4.777)	(4.991)	(4.489)	(4.391)	(4.513)	(5.085)
Foreign_coach==1	5.592*	4.308	3.971	4.426	6.404**	4.951	4.092	5.343
	(2.817)	(2.676)	(2.646)	(2.850)	(3.254)	(3.069)	(3.087)	(3.542)
Stand. dev. appearances		0.274***	0.275***	0.307***		0.288***	0.290***	0.312***
		(0.023)	(0.023)	(0.029)		(0.027)	(0.027)	(0.033)
Log of GDP/capita			8.616	9.274			18.635**	16.656*
			(6.704)	(8.102)			(8.111)	(9.154)
Population (mln)			0.410*				0.215	
			(0.216)				(0.425)	
Log immig. stocks, 18y lag				-0.556				1.403
				(1.435)				(1.800)
Observations	1900	1900	1900	1676	1900	1900	1900	1676
Kleibergen-Paap LM test					0.00	0.00	0.00	0.00
Kleibergen-Paap F test					20.62	20.60	17.63	11.97
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Baseline estimates for the unilateral framework. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 5–7) / years 1978–2018 (in columns 4 and 8). Dependent variable: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 9. Football performance and genetic diversity: accounting for intercontinental matches

	Dependent variable: change in rating of national football team (Elo score); scores from websource							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Variable of interest								
Diversity	3.690^{**}	3.355^{**}	3.417^{**}	3.531^{**}	18.984^{*}	17.341^{*}	21.404^{*}	27.353^{*}
	(1.272)	(1.126)	(1.112)	(1.302)	(10.390)	(9.929)	(11.711)	(15.457)
Control variables								
Stand. dev. appearances		0.307^{***}	0.307^{***}	0.352^{***}		0.325^{***}	0.326^{***}	0.364^{***}
		(0.027)	(0.028)	(0.032)		(0.032)	(0.032)	(0.036)
Log of GDP/capita			3.483	-1.976			11.483	3.702
			(6.749)	(7.308)			(8.679)	(9.004)
Population (mln)			0.074				-0.083	
			(0.189)				(0.402)	
Log immig. stocks, 18y lag				1.379				2.872
				(1.385)				(1.805)
Observations	1900	1900	1900	1676	1900	1900	1900	1676
Kleibergen-Paap LM test					0.00	0.00	0.00	0.00
Kleibergen-Paap F test					20.34	20.31	17.47	11.83
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Baseline estimates for the unilateral framework. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 5–7) / years 1978–2018 (in columns 4 and 8). Dependent variable: changes in the Elo score of the national team (end vs. beginning of the championship stage), as from the web source eloratings.net. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 10. Goal difference and diversity: controlling for initial strength

	Dependent variable: goal difference									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV
Variable of interest										
Bilateral diversity	0.094**	0.041	0.039	0.039	0.039	0.612**	0.897**	0.861**	1.315**	1.309**
	(0.046)	(0.043)	(0.043)	(0.045)	(0.045)	(0.272)	(0.319)	(0.277)	(0.476)	(0.474)
Control variables										
Initial Elo score, home	0.002***	0.001**	0.001**	0.001**	0.001**	0.002***	0.001***	0.001***	0.001**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Initial Elo score, away	-0.002***	-0.001***	-0.001***	-0.001**	-0.001**	-0.002***	-0.001***	-0.001***	-0.001**	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log of GDP/capita, home		0.167	0.168	0.039	0.038		0.476**	0.466**	0.207	0.205
		(0.195)	(0.195)	(0.207)	(0.207)		(0.237)	(0.230)	(0.251)	(0.251)
Log of GDP/capita, away		-0.759***	-0.757***	-0.517**	-0.522**		-1.110***	-1.092***	-0.790**	-0.794**
		(0.197)	(0.197)	(0.214)	(0.214)		(0.244)	(0.235)	(0.277)	(0.276)
Population (mln), home			-0.004					-0.001		
			(0.003)					(0.003)		
Population (mln), away			-0.001					-0.004		
			(0.003)					(0.003)		
Log immig. stocks, 18y lag, home				0.051	0.050				0.121**	0.120**
				(0.044)	(0.044)				(0.057)	(0.057)
Log immig. stocks, 18y lag, away				-0.080*	-0.081*				-0.135**	-0.135**
				(0.046)	(0.046)				(0.054)	(0.054)
Observations	3877	3877	3877	3568	3568	3877	3877	3877	3568	3568
Kleibergen-Paap LM test						0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test						31.87	25.31	33.49	15.26	15.27
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minute appearances		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Geo-political controls					Yes					Yes

Notes: Estimates for the bilateral framework (match-level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970-2018 (for columns 1-3,6-8) / years 1978-2018 (in columns 4-5 and 8-9). Dependent variable: goal difference. In all regressions, we include Team and year fixed effects, as well as a stage dummy. Column 1 to Column 5 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. Column 5 to Column 8 display IV results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F-test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 11. Bilateral framework—further results

	Dependent variable: goal difference				hyperbolic sine	home score	goal difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV: Diversity, appearance	**IV:** Diversity, SW	**IV:** 20 years lag	**IV:** 22 years lag	**IV:**Goal difference, hyperbolic sine	**IV:**Outcome: home score	**IV:**Diversity, home vs. away
Variable of interest							
Bilateral diversity			0.888^{***}	0.793^{***}	0.573^{***}	0.556^{**}	
			(0.269)	(0.237)	(0.172)	(0.229)	
Bilateral diversity, appearance	1.125^{**}						
	(0.368)						
Bilateral diversity, SW		1.065^{**}					
		(0.354)					
Diversity, home							0.726^{**}
							(0.296)
Diversity, away							-0.651^{**}
							(0.278)
Control variables							
Population (mln), home	-0.004	-0.001	-0.002	-0.004	-0.000	-0.002	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Population (mln), away	-0.002	-0.004	-0.003	-0.002	-0.002	0.000	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Log of GDP/capita, home	0.577^{**}	0.596^{**}	0.445^{*}	0.411^{*}	0.272^{**}	0.292^{*}	0.475^{*}
	(0.262)	(0.267)	(0.231)	(0.221)	(0.139)	(0.163)	(0.266)
Log of GDP/capita, away	-1.124^{***}	-1.208^{***}	-1.020^{***}	-0.908^{***}	-0.422^{**}	-0.918^{***}	-1.083^{***}
	(0.251)	(0.276)	(0.232)	(0.226)	(0.130)	(0.197)	(0.256)
Finals==QUALI	0.568^{***}	0.544^{***}	0.569^{***}	0.537^{***}	0.284^{***}	0.453^{***}	0.550^{***}
	(0.113)	(0.112)	(0.106)	(0.106)	(0.066)	(0.085)	(0.163)
Observations	3877	3877	3762	3643	3877	3877	3877
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	28.13	22.76	40.42	47.09	32.29	32.29	15.47
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minute appearances	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates for the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–018 (for columns 1–2, 5–7) / the first year available for the instrument to 2018 (in columns 3–4). Dependent variable: goal difference for columns 1–4 and 6, its hyperbolic sine transformation in Column 5 and the goals scored by the home team in Column 7. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–7 display IV results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 12. Bilateral framework, goals for, goals against

	Dependent variable: home team's goals scored				
	(1) **IV:**Poisson, control function	(2) **IV:**Poisson, control function	(3) **IV:**Poisson, control function	(4) **IV:**Poisson, control function	(5) **IV:**Poisson, control function
AMEs diversity away	-0.215	-0.393	-0.375*	-0.630*	-0.631*
	(0.215)	(0.249)	(0.222)	(0.345)	(0.346)
AMEs diversity home	0.436*	0.516*	0.542**	0.596	0.592
	(0.242)	(0.280)	(0.250)	(0.454)	(0.458)

	Dependent variable: away team's goals scored				
	(1) **IV:**Poisson, control function	(2) **IV:**Poisson, control function	(3) **IV:**Poisson, control function	(4) **IV:**Poisson, control function	(5) **IV:**Poisson, control function
AMEs diversity away	0.395**	0.434**	0.354**	0.489*	0.497*
	(0.171)	(0.193)	(0.169)	(0.280)	(0.281)
AMEs diversity home	-0.181	-0.204	-0.216	-0.485	-0.471
	(0.183)	(0.215)	(0.191)	(0.376)	(0.381)
Observations	3877	3877	3877	3568	3568
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances		Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: Average marginal effects. Estimates for the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–018 (for columns 1–3) / years 1978–2018 (in columns 4–5). Dependent variable: home team's number of goals scored in the top sub-table, away team's number of goals scored in the top sub-table. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–5 display results from a Poisson, control-function regression, with bootstrapped (2000 reps) standard errors in parentheses, clustered at team pair level. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 13. Bilateral framework, controlling for coach quality

	Dependent variable: goal difference									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	OLS	OLS	OLS	IV	IV	IV	IV	IV
Variable of interest										
Bilateral diversity	0.091*	0.032	0.030	0.034	0.034	0.684**	0.960**	0.923**	1.354**	1.348**
	(0.047)	(0.044)	(0.044)	(0.045)	(0.045)	(0.291)	(0.342)	(0.295)	(0.498)	(0.497)
Control variables										
Coach tenure, home	0.014	0.004	0.004	0.012	0.012	0.007	-0.006	-0.006	-0.008	-0.008
	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)	(0.016)	(0.016)	(0.016)	(0.019)	(0.019)
Coach tenure, away	-0.018	-0.006	-0.006	0.005	0.005	-0.010	0.008	0.007	0.028	0.028
	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)	(0.019)	(0.019)
Coach age, home	-0.004	-0.005	-0.004	-0.006	-0.006	-0.002	-0.001	-0.001	0.000	0.000
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Coach age, away	0.005	0.007	0.007	0.007	0.007	0.003	0.002	0.003	0.000	0.000
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)
Coach award, home	0.186*	0.122	0.106	0.078	0.074	0.187*	0.131	0.125	0.076	0.073
	(0.099)	(0.095)	(0.095)	(0.097)	(0.097)	(0.100)	(0.100)	(0.100)	(0.110)	(0.109)
Coach award, away	-0.335***	-0.243**	-0.243**	-0.243**	-0.246**	-0.348***	-0.270**	-0.283**	-0.265**	-0.267**
	(0.090)	(0.089)	(0.090)	(0.092)	(0.092)	(0.092)	(0.097)	(0.097)	(0.107)	(0.107)
Foreign coach, home	0.041	-0.024	-0.018	0.014	0.014	0.069	0.022	0.021	0.117	0.117
	(0.077)	(0.074)	(0.074)	(0.078)	(0.077)	(0.081)	(0.083)	(0.083)	(0.099)	(0.099)
Foreign coach, away	-0.239**	-0.143*	-0.143*	-0.204**	-0.202**	-0.246**	-0.148*	-0.142*	-0.247**	-0.245**
	(0.084)	(0.082)	(0.082)	(0.083)	(0.083)	(0.085)	(0.086)	(0.086)	(0.095)	(0.094)
Log of GDP/capita, home		0.135	0.136	-0.021	-0.024		0.461*	0.451*	0.128	0.124
		(0.197)	(0.197)	(0.209)	(0.209)		(0.245)	(0.236)	(0.257)	(0.256)
Log of GDP/capita, away		-0.669***	-0.668***	-0.412*	-0.420**		-1.065***	-1.047***	-0.705**	-0.711**
		(0.199)	(0.199)	(0.214)	(0.214)		(0.257)	(0.245)	(0.284)	(0.284)
Population (mln), home			-0.004					-0.001		
			(0.003)					(0.003)		
Population (mln), away			-0.000					-0.004		
			(0.003)					(0.003)		
Log immig. stocks, 18y lag, home				0.067	0.067				0.149**	0.149**
				(0.044)	(0.044)				(0.061)	(0.061)
Log immig. stocks, 18y lag, away				-0.091**	-0.092**				-0.153**	-0.152**
				(0.046)	(0.046)				(0.056)	(0.056)
Observations	3873	3873	3873	3568	3568	3873	3873	3873	3568	3568
Kleibergen-Paap LM test						0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test						30.97	24.22	32.21	14.64	14.65
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minute appearances		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Geo-political controls					Yes					Yes

Notes: Estimates for the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 6–8) / years 1978–2018 (in columns 4–5 and 8–9). Dependent variable: goal difference. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–5 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

References

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2):155–194.

Alesina, A., Harnoss, J., and Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2):101–138.

Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3):762–800.

Alesina, A., Spolaore, E., and Wacziarg, R. (2000). Economic integration and political disintegration. *American Economic Review*, 90(5):1276–1296.

Arbatlı, C. E., Ashraf, Q. H., Galor, O., and Klemp, M. (2020). Diversity and conflict. *Econometrica*, 88(2):727–797.

Ashraf, Q. and Galor, O. (2013). The "Out of africa" hypothesis, human genetic diversity, and comparative economic development. *American Economic Review*, 103(1):1–46.

Bove, V. and Elia, L. (2017). Migration, diversity, and economic growth. *World Development*, 89(C):227–239.

Brox, E. and Krieger, T. (2019). Birthplace diversity and team performance. *ZEW Discussion Paper No. 19-020, Mannheim.*

Buonanno, P. and Vanin, P. (2017). Social closure, surnames and crime. *Journal of Economic Behavior & Organization*, 137:160–175.

Burbidge, J. B., Magee, L., and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401).

Cavalli-Sforza, L. L., Menozzi, P., and Piazza, A. (1994). *The history and geography of human genes*. Princeton, N.J.: Princeton University Press.

Clark, G. (2015). *The son also rises: Surnames and the history of social mobility*. Princeton, N.J.: Princeton University Press.

Cline Center for Advanced Social Research, (2020). Composition of religious and ethnic groups (CREG) project. `https://clinecenter.illinois.edu/project/Religious-Ethnic-Identity/composition-religious-and-ethnic-groups-creg-project`.

Collier, P. (2001). Implications of ethnic diversity. *Economic Policy*, 16(32):128–166.

Delis, M. D., Gaganis, C., Hasan, I., and Pasiouras, F. (2017). The effect of board directors from countries with different genetic diversity levels on corporate performance. *Management Science*, 63(1):231–249.

Delis, M. D., Litina, A., and Zanaj, S. (2021). Diversity of the screen. *DEM Discussion Papers Series*.

Desmet, K., Ortuño-Ortín, I., and Wacziarg, R. (2017). Culture, ethnicity, and diversity. *American Economic Review*, 107(9):2479–2513.

Docquier, F., Turati, R., Valette, J., and Vasilakis, C. (2019). Multiculturalism and growth: Skill-specific evidence from the post-World War II period. *halshs-01425462*.

Easterly, W. and Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250.

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth*, 8(2):195–222.

FIFA (2020). Men's ranking: Belgium, royal belgian football association.
`https://www.fifa.com/fifa-world-ranking/associations/association/BEL/men/`.

France24 (July 18, 2018). In Mandela address, Obama cites French World Cup model as champs of diversity. *France24*.
`https://www.france24.com/en/20180718-south-africa-world-cup-obama-mandela-france-world-cup-football-champs-model-diversity`.

Gould, E. D. and Winter, E. (2009). Interactions between workers and the technology of production: Evidence from professional baseball. *The Review of Economics and Statistics*, 91(1):188–200.

Greenberg, J. H. (1956). The measurement of linguistic diversity. *Language*, 32(1):109–115.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does culture affect economic outcomes? *Journal of Economic perspectives*, 20(2):23–48.

Haas, H. and Nüesch, S. (2012). Are multinational teams more successful? *The International Journal of Human Resource Management*, 23(15):3105–3113.

Hall, C. D. (2012). Fishing for all-stars in a time of global free agency: Understanding fifa eligibility rules and the impact on the us men's national team. *Marquette Sports Law Review*, 23:191.

Head, K., Mayer, T., and Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of international Economics*, 81(1):1–14.

Ingersoll, K., Malesky, E., and Saiegh, S. M. (2017). Heterogeneity and team performance: Evaluating the effect of cultural diversity in the world's top soccer league. *Journal of Sports Analytics*, 3(2):67–92.

Jobling, M. A. (2001). In the name of the father: surnames and genetics. *Trends in Genetics*, 17(6):353–357.

Jürisoo, M. (January 2020). International football results from 1872 to 2020. *Version 4*.
`https://www.kaggle.com/martj42/international-football-results-from-1872-to-2017/`.

Kahane, L., Longley, N., and Simmons, R. (2013). The effects of coworker heterogeneity on firm-level output: Assessing the impacts of cultural and language diversity in the national hockey league. *Review of Economics and Statistics*, 95(1):302–314.

Kerr, S. P. and Kerr, W. R. (2018). Global collaborative patents. *The Economic Journal*, 128(612):F235–F272.

Langville, A. N. and Meyer, C. D. (2012). *Who's# 1?: The science of rating and ranking.* Princeton, N.J. Princeton University Press.

Lazear, E. P. (1999). Globalisation and the market for team-mates. *Economic Journal*, 109(454).

Lin, W. and Wooldridge, J. M. (2019). Chapter 2 -Testing and correcting for endogeneity in nonlinear unobserved effects models. N.p.: Elsevier, Academic Press. In *Panel data econometrics*, ed. M. Tsionas, 21–43. N.p.: Elsevier (Academic Press).

Lippi, G., Longo, U. G., and Maffulli, N. (2010). Genetics and sports. *British medical bulletin*, 93(1):27–47.

Özden, C., Parsons, C. R., Schiff, M., and Walmsley, T. L. (2011). Where on Earth is Everybody? The evolution of global bilateral migration 1960–2000. *The World Bank Economic Review*, 25(1):12–56.

Parshakov, P., Coates, D., and Zavertiaeva, M. (2018). Is diversity good or bad? Evidence from eSports teams analysis. *Applied Economics*, 50(47):5064—5075.

Piazza, A., Rendine, S., Zei, G., Moroni, A., and Cavalli-Sforza, L. L. (1987). Migration rates of human populations from surname distributions. *Nature*, 329(6141):714–716.

Prat, A. (2002). Should a team be homogeneous? *European Economic Review*, 46(7):1187–1207.

Spolaore, E. and Wacziarg, R. (2009). The diffusion of development. *The Quarterly Journal of Economics*, 124(2):469–529.

Spolaore, E. and Wacziarg, R. (2018). Ancestry and development: New evidence. *Journal of Applied Econometrics*, 33(5):748–762.

Thränhardt, D. (1992). *Europe, a new immigration continent: Policies and politics in comparative perspective.* 2nd ed. Vol. 1. Münster: LIT Verlag.

Tovar, J. (2020). Performance, diversity and national identity evidence from association football. *Economic Inquiry*, 58(2):897–916.

van Campenhout, G., van Sterkenburg, J., and Oonk, G. (2019). Has the world cup become more migratory? A comparative history of foreign-born players in national football teams, c. 1930-2018. *Comparative Migration Studies*, 7(1):1–19.

Vasilakis, C. (2017). Does talent migration increase inequality? A quantitative assessment in football labour market. *Journal of Economic Dynamics and Control*, 85(C):150 – 166.

Wilson, M. (June 15, 2021). Brilliant belgians just incomparable insists scotland assistant coach mcghee. *Mail Online*. https://www.dailymail.co.uk/sport/football/article-2409118/Belgium-incomparable-Scotland-says-coach-Mark-McGhee-ahead-World-Cup-qualifier.html.

Appendices

Appendix A

The Elo formula

Updated at every match, the Elo score is computed as

$$R_n = R_o + K(W - W_e)$$

where R_n is the new rating, R_0 is the old (pre-match) rating. K is a constant weight for the tournament played: 60 for World Cup finals, 50 for the EUROs finals, 40 for World Cup and EURO qualifiers. K is then adjusted as follows for the goal difference in the game. It is increased by half if a game is won by 2 goals, by 3/4 if a game is won by 3 goals, and by $3/4 + (N - 3)/8$ if the game is won by 4 or more goals, where N is the goal difference.
W accounts for the game result (1 for a win, 0.5 for a draw, and 0 for a loss). W_e is the expected result:

$$W_e = \frac{1}{(10^{-\frac{dr}{400}} + 1)}$$

where dr equals the difference in ratings plus 100 points for a team playing at home. For a check on the similarity with the FIFA Ranking adopted after 2018 World Cup in Russia, please see the official FIFA publication:
https://resources.fifa.com/image/upload/revision-of-the-fifa-coca-cola-world-ranking.pdf?cloudid=fzltr4s8tz3v3vy0aqo1.

Appendix B

Our prediction algorithm: examples

	BELGIUM TEAM, 2018 World Cup Finals
Adnan_Januzaj	‘”Kosovo”’
Axel_Witsel	‘”Netherlands”’
Dedryck_Boyata	‘”DR Congo”’
Dries_Mertens	‘”Belgium”’
Eden_Hazard	‘”United States”’
Jan_Vertonghen	‘”Belgium”’
Kevin_De_Bruyne	‘”Belgium”’
Koen_Casteels	‘”Belgium”’
Leander_Dendoncker	‘”Belgium”’
Marouane_Fellaini	‘”Morocco”’
Michy_Batshuayi	‘”Belgium”’
Mousa_Dembélé	‘”Mali”’
Nacer_Chadli	‘”Morocco”’
Romelu_Lukaku	‘”DR Congo”’
Simon_Mignolet	‘”Belgium”’
Thibaut_Courtois	‘”France”’
Thomas_Meunier	‘”France”’
Thomas_Vermaelen	‘”Belgium”’
Toby_Alderweireld	‘”Belgium”’
Vincent_Kompany	‘”DR Congo”’
Yannick_Carrasco	‘”Spain”’
Youri_Tielemans	‘”Belgium”’

Origins predicted vs names of the Belgium squad members, in the 2018 World Cup finals.

	SWEDEN TEAM, 2018 World Cup Finals
Albin_Ekdal	‘”Sweden”’
Andreas_Granqvist	‘”Sweden”’
Emil_Forsberg	‘”Sweden”’
Emil_Krafth	‘”Sweden”’
Filip_Helander	‘”Finland”’
Gustav_Svensson	‘”Sweden”’
Isaac_Kiese_Thelin	‘”Sweden”’
Jimmy_Durmaz	‘”Turkey”’
Johan_Johnsson	‘”Sweden”’
John_Guidetti	‘”Italy”’
Kristoffer_Nordfeldt	‘”Sweden”’
Ludwig_Augustinsson	‘”Sweden”’
Marcus_Berg	‘”Norway”’
Marcus_Rohdén	‘”Sweden”’
Martin_Olsson	‘”Sweden”’
Mikael_Lustig	‘”Sweden”’
Ola_Toivonen	‘”Finland”’
Oscar_Hiljemark	‘”Sweden”’
Robin_Olsen	‘”Norway”’
Sebastian_Larsson	‘”Sweden”’
Victor_Lindelöf	‘”Sweden”’
Viktor_Claesson	‘”Sweden”’

Origins predicted vs names of the Sweden squad members, in the 2018 World Cup finals.

	BELGIUM TEAM, 1990 World Cup Finals
Bruno_Versavel	‘”Belgium”’
Enzo_Scifo	‘”Italy”’
Eric_Gerets	‘”Belgium”’
Filip_De_Wilde	‘”Belgium”’
Franky_Van_Der_Elst	‘”Belgium”’
François_De_Sart	‘”Belgium”’
Georges_Grün	‘”Germany”’
Gilbert_Bodart	‘”Belgium”’
Jan_Ceulemans	‘”Belgium”’
Lei_Clijsters	‘”Belgium”’
Lorenzo_Staelens	‘”Belgium”’
Marc_Degryse	‘”Belgium”’
Marc_Emmers	‘”Belgium”’
Marc_Van_Der_Linden	‘”Netherlands”’
Michel_De_Wolf	‘”Belgium”’
Michel_Preud_homme	‘”Belgium”’
Nico_Claesen	‘”Belgium”’
Pascal_Plovie	‘”Belgium”’
Patrick_Vervoort	‘”Belgium”’
Philippe_Albert	‘”Germany”’
Stéphane_Demol	‘”Belgium”’

Origins predicted vs names of the Belgium squad members, in the 1990 World Cup finals.

	SWEDEN TEAM, 1990 World Cup Finals
Anders_Limpar	‘”Hungary”’
Glenn_Hysén	‘”Sweden”’
Glenn_Strömberg	‘”Sweden”’
Joakim_Nilsson	‘”Sweden”’
Johnny_Ekström	‘”Sweden”’
Jonas_Thern	‘”Sweden”’
Klas_Ingesson	‘”Sweden”’
Lars_Eriksson	‘”Sweden”’
Leif_Engqvist	‘”Sweden”’
Mats_Gren	‘”Sweden”’
Mats_Magnusson	‘”Sweden”’
Niklas_Nyhlén	‘”Sweden”’
Peter_Larsson	‘”Sweden”’
Roger_Ljung	‘”Sweden”’
Roland_Nilsson	‘”Sweden”’
Stefan_Pettersson	‘”Sweden”’
Stefan_Schwarz	‘”Germany”’
Sven_Andersson	‘”Sweden”’
Thomas_Ravelli	‘”Italy”’
Tomas_Brolin	‘”Sweden”’
Ulrik_Jansson	‘”Sweden”’

Origins predicted vs names of the Sweden squad members, in the 1990 World Cup finals.

BELGIUM TEAM, 1970 World Cup Finals	
Alfons_Peeters	‘”Belgium”’
Erwin_Vandendaele	‘”Belgium”’
Frans_Janssens	‘”Belgium”’
Georges_Heylens	‘”Belgium”’
Jacques_Beurlet	‘”Belgium”’
Jacques_Duquesne	‘”Belgium”’
Jan_Verheyen	‘”Belgium”’
Jean_Dockx	‘”Belgium”’
Jean_Thissen	‘”Netherlands”’
Johan_Devrindt	‘”Belgium”’
Léon_Jeck	‘”Germany”’
Léon_Semmeling	‘”Belgium”’
Marie_Trappeniers	‘”Belgium”’
Maurice_Martens	‘”Belgium”’
Nicolas_Dewalque	‘”Belgium”’
Odilon_Polleunis	‘”Belgium”’
Paul_Van_Himst	‘”Belgium”’
Pierre_Carteus	‘”Belgium”’
Raoul_Lambert	‘”France”’
Wilfried_Puis	‘”Belgium”’
Wilfried_Van_Moer	‘”Belgium”’

Origins predicted vs names of the
Belgium squad members, in the 1970 World Cup finals.

SWEDEN TEAM, 1970 World Cup Finals	
Björn_Nordqvist	‘”Sweden”’
Bo_Larsson	‘”Sweden”’
Claes_Cronqvist	‘”Sweden”’
Gunnar_Larsson	‘”Sweden”’
Göran_Nicklasson	‘”Sweden”’
Hans_Selander	‘”Sweden”’
Inge_Ejderstedt	‘”Sweden”’
Jan_Olsson	‘”Sweden”’
Krister_Kristensson	‘”Sweden”’
Kurt_Axelsson	‘”Sweden”’
Leif_Eriksson	‘”Sweden”’
Ove_Grahn	‘”Sweden”’
Ove_Kindvall	‘”Sweden”’
Roland_Grip	‘”Sweden”’
Ronney_Pettersson	‘”Sweden”’
Ronnie_Hellström	‘”Sweden”’
Sten_Pålsson	‘”Sweden”’
Thomas_Nordahl	‘”Norway”’
Tom_Turesson	‘”Sweden”’
Tommy_Svensson	‘”Sweden”’
Örjan_Persson	‘”Sweden”’

Origins predicted vs names of the
Sweden squad members, in the 1970 World Cup finals.

Appendix C

FIFA and UEFA

The inauguration of the FIFA World Cup championship was held in 1930. The first tournament was held in and won by Uruguay, and it was the only tournament for which no qualification stage was set. All countries affiliated with FIFA were invited to participate, and 13 countries accepted. Since then, the playing of the World Cup was established as every four years (with the exception of World War II breaks in 1942 and 1946), and a qualification process determined the final-stage participants. Both the number of participating countries and of qualified teams increased over time. Initially set at 16, the latter would increase to 24 in 1982, then to 32 in 1998, and will reach 48 in 2026.[33]

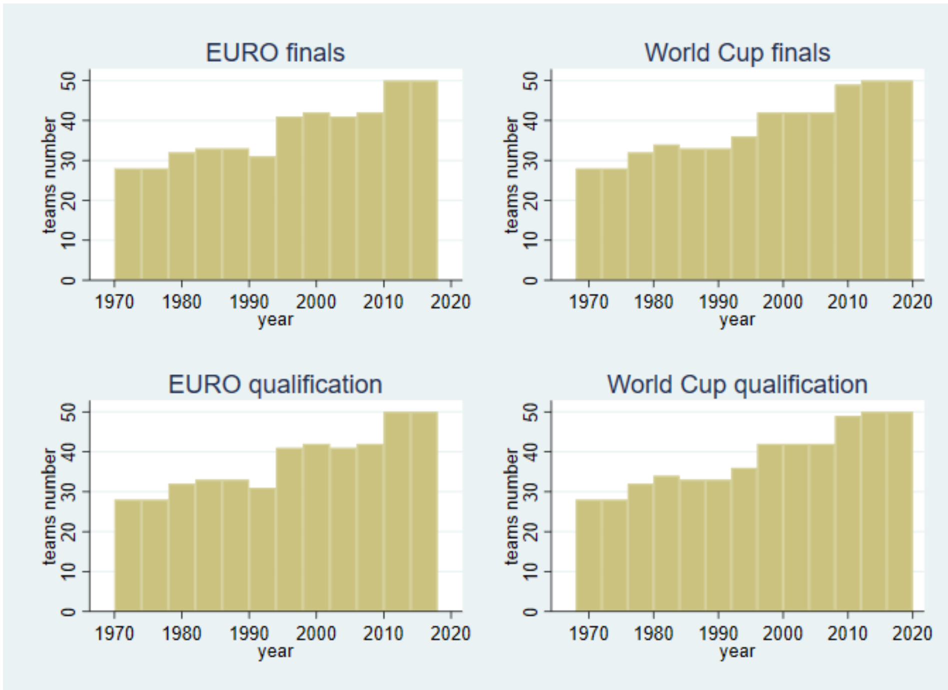
Relatively newer, the Union des Associations Européennes de Football (UEFA) was founded in 1954 and it organized the first European Nations' Cup (currently referred as to UEFA EUROs) in 1960. The Soviet Union won the first tournament in which 4 teams of 17 had made it to the final stage.[34] Here again, the number of teams selected for the

[33] For more details on the FIFA World Cup, see https://www.fifa.com/tournaments/mens/worldcup

[34] For more details on the UEFA EUROs, see https://www.uefa.com/uefaeuro/history/.

final stages increased over time (8 teams in 1980, 16 in 1996, and 24 in 2016).

Figure 8. Number of teams in the sample, by year and tournament



Notes: In Figure 8, we plot the evolution of the number of teams participating to the EUROs (on the left) and to the World Cup (on the right). The equivalent number of teams in the qualification and final stages are a result of our unilateral panel construction, where we avoid teams dropping out to maintain balance and prevent out-selection issues.

In terms of team squad members, there is an upperbound for the final stage, whereas virtually no limitations exist for the qualification matches, leaving eligibility criteria aside [35]. The limit of 22 players per squad was increased by 1 in the 2002 World Cup and EUROs 2004.

[35] The detailed regulations for the 2018 World Cup can be found at this site: https://www.uefa.com/MultimediaFiles/Download/Regulations/ uefaorg/Regulations/ 01/87/54/21/1875421_DOWNLOAD.pdf

Appendix D

Additional tables

D.1 First-stage regressions

Table 14. Unilateral framework, benchmark, first-stage regressions

	First stage: Dependent variable: team's diversity			
	(1)	(2)	(3)	(4)
	IV	IV	IV	IV
IV				
IV, 18y lag	0.011***	0.011***	0.012***	0.008***
	(0.002)	(0.002)	(0.003)	(0.002)
Stand. dev. appearances		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Control variables				
Log of GDP/capita			-0.299**	-0.129
			(0.117)	(0.128)
Population (mln)			-0.015	
			(0.011)	
Log immig. stocks, 18y lag				-0.060***
				(0.017)
Observations	1900	1900	1900	1676
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	20.34	20.31	17.47	11.83
Team FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes

Notes: First stage for the baseline estimates of the unilateral framework, Table 3. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) /years 1978–2018 (in Column 4). Dependent variable in the second stage: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 15. Bilateral framework, benchmark, first-stage regressions

	First stage: Dependent variable: bilateral diversity				
	(1)	(2)	(3)	(4)	(5)
	IV	IV	IV	IV	IV
Instrumental variable					
IV, home vs. away	0.008***	0.007***	0.008***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Control variables					
Log of GDP/capita, home		-0.260**	-0.247**	-0.056	-0.056
		(0.086)	(0.085)	(0.103)	(0.103)
Log of GDP/capita, away		0.312***	0.292**	0.150	0.150
		(0.091)	(0.091)	(0.110)	(0.110)
Stand. dev. appearances, away		-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Stand. dev. appearances, home		0.001***	0.001***	0.002***	0.002***
		(0.000)	(0.000)	(0.000)	(0.000)
Population (mln), home			-0.005***		
			(0.001)		
Population (mln), away			0.005***		
			(0.001)		
Log immig. stocks, 18y lag, home				-0.057**	-0.057**
				(0.019)	(0.019)
Log immig. stocks, 18y lag, away				0.039**	0.039**
				(0.020)	(0.020)
Observations	3877	3877	3877	3568	3568
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	31.46	24.52	32.29	14.48	14.49
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances		Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: First stage for the baseline estimates of the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in columns 4–5). Dependent variable in the second stage: goal difference. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–5 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 16. Unilateral framework, alternative measure of rating, first-stage regressions

	First stage: Dependent variable: team's diversity			
	(1)	(2)	(3)	(4)
	IV	IV	IV	IV
IV				
IV, 18y lag	0.011***	0.012***	0.012***	0.008***
	(0.002)	(0.003)	(0.003)	(0.002)
Control variables				
Log of GDP/capita		-0.299**	-0.299**	-0.129
		(0.117)	(0.117)	(0.128)
Population (mln)		-0.015	-0.015	
		(0.011)	(0.011)	
Stand. dev. appearances		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Log immig. stocks, 18y lag				-0.060***
				(0.017)
Observations	1900	1900	1900	1676
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	20.34	17.47	17.47	11.83
Team FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes

Notes: First stage for the estimates of the unilateral framework, Table 4. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in Column 4). Dependent variable in the second stage: Elo score levels of the national team (end of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 17. Unilateral framework, opponent's strength and diversity, first-stage regressions

	First stage: Dependent variable: team's diversity				
	(1)	(2)	(3)	(4)	(5)
	IV	IV	IV	IV	IV
IV					
IV, 18y lag	0.011***	0.012***	0.012***	0.008**	0.008**
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
Control variables					
Log of GDP/capita		-0.300**	-0.292**	-0.123	-0.120
		(0.116)	(0.115)	(0.127)	(0.126)
Population (mln)		-0.015	-0.015		
		(0.011)	(0.011)		
Adversary's diversity		0.046**			0.022
		(0.021)			(0.022)
Adversary's strength			0.001***	0.001***	0.001***
			(0.000)	(0.000)	(0.000)
Stand. dev. appearances			0.000	0.001	0.001
			(0.000)	(0.000)	(0.000)
Log immig. stocks, 18y lag				-0.063***	-0.064***
				(0.017)	(0.017)
Observations	1900	1900	1900	1676	1676
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	20.34	16.93	16.72	10.27	10.05
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes

Notes: First stage for the estimates of the unilateral framework, Table 6. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in Column 4). Dependent variable in the second stage: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 18. Unilateral framework, further results, first-stage regressions

	First stage: Dependent variable: team's diversity						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV	IV	IV	IV	IV	IV	IV
IV							
IV, 18y lag	0.011***	0.012***	0.012***	0.008**	0.027**		
	(0.002)	(0.003)	(0.003)	(0.002)	(0.009)		
Log of GDP/capita		-0.300**	-0.292**	-0.123	-1.116***	-0.214*	-0.147
		(0.116)	(0.115)	(0.127)	(0.329)	(0.120)	(0.128)
Population (mln)		-0.015	-0.015		-0.025	-0.025**	-0.033**
		(0.011)	(0.011)		(0.030)	(0.012)	(0.013)
Adversary's diversity		0.046**					
		(0.021)					
Adversary's strength			0.001***	0.001***			
			(0.000)	(0.000)			
Control variables							
Stand. dev. appearances			0.000	0.001	0.000	-0.000	-0.000
			(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Log immig. stocks, 18y lag				-0.063***			
				(0.017)			
L.IV, 18y lag						0.016***	
						(0.004)	
L2.IV, 18y lag							0.022***
							(0.004)
Observations	1900	1900	1900	1676	1900	1784	1670
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	20.34	16.93	16.72	10.27	9.88	20.54	24.29
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: First stage for the estimates of the unilateral framework, Table 7. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3 and 5) / years 1978–2018 (in Column 4). Dependent variable in the second stage: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 19. Unilateral framework, coach quality, first-stage regressions

	First stage: Dependent variable: team's diversity			
	(1)	(2)	(3)	(4)
	IV	IV	IV	IV
IV, 18y lag	0.011***	0.011***	0.013***	0.008***
	(0.002)	(0.002)	(0.003)	(0.002)
Control variables				
Coach age	-0.003	-0.003	-0.003	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)
Foreign_coach==1	-0.030	-0.031	-0.017	-0.040
	(0.054)	(0.053)	(0.053)	(0.057)
Stand. dev. appearances		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Log of GDP/capita			-0.293**	-0.114
			(0.117)	(0.129)
Population (mln)			-0.015	
			(0.011)	
Log immig. stocks, 18y lag				-0.060***
				(0.017)
Observations	1900	1900	1900	1676
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	20.53	20.51	17.57	12.12
Team FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes

Notes: First stage for the estimates of the Unilateral framework, Table 8. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in Column 4). Dependent variable in the second stage: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 20. Bilateral framework, controlling for initial strength, first-stage regressions

	First stage: Dependent variable: bilateral diversity				
	(1)	(2)	(3)	(4)	(5)
	IV	IV	IV	IV	IV
Instrumental variable					
IV, home vs. away	0.008***	0.007***	0.008***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Control variables					
Initial Elo score, home	-0.000	-0.000**	-0.000**	-0.000**	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Initial Elo score, away	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log of GDP/capita, home		-0.268**	-0.256**	-0.076	-0.076
		(0.086)	(0.086)	(0.101)	(0.101)
Log of GDP/capita, away		0.314***	0.294**	0.157	0.156
		(0.091)	(0.091)	(0.109)	(0.109)
Stand. dev. appearances, away		-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Stand. dev. appearances, home		0.002***	0.002***	0.002***	0.002***
		(0.000)	(0.000)	(0.000)	(0.000)
Population (mln), home			-0.005***		
			(0.001)		
Population (mln), away			0.005***		
			(0.001)		
Log immig. stocks, 18y lag, home				-0.049**	-0.049**
				(0.018)	(0.018)
Log immig. stocks, 18y lag, away				0.037*	0.037*
				(0.019)	(0.019)
Observations	3877	3877	3877	3568	3568
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	31.87	25.31	33.49	15.26	15.27
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances		Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: First stage for the estimates of the bilateral framework, Table 10. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970—2018 (columns 1–3 and 5) / years 1978–2018 (in Column 4) first year available for the instrument to 2018 (columns 6 and 7). Dependent variable in the second stage: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 21. Bilateral framework, further results, first-stage regressions

	Dependent variable in second stage: goal difference				hyperbolic sine	Home score	Goal difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	IV: Diversity, appearance	**IV:** Diversity, SW	**IV:** 20 years lag	**IV:** 22 years lag	**IV:**Goal difference, hyperbolic sine	**IV:**Outcome: home score	**IV:**Diversity, home vs. away
Instrumental variable							
IV, home vs. away	0.007^{***}	0.007^{***}			0.008^{***}	0.008^{***}	
	(0.001)	(0.002)			(0.001)	(0.001)	
Stand. dev. appearances, away	-0.001^{***}	-0.002^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Stand. dev. appearances, home	0.001^{***}	0.002^{***}	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}	0.001^{**}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Population (mln), home	-0.002^{**}	-0.005^{***}	-0.005^{***}	-0.005^{***}	-0.005^{***}	-0.005^{***}	-0.006^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Population (mln), away	0.003^{**}	0.005^{***}	0.006^{***}	0.006^{***}	0.005^{***}	0.005^{***}	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Control variables							
Log of GDP/capita, home	-0.312^{***}	-0.348^{***}	-0.223^{**}	-0.188^{**}	-0.247^{**}	-0.247^{**}	-0.391^{***}
	(0.088)	(0.096)	(0.084)	(0.085)	(0.085)	(0.085)	(0.078)
Log of GDP/capita, away	0.269^{**}	0.363^{***}	0.278^{**}	0.246^{**}	0.292^{**}	0.292^{**}	0.005
	(0.092)	(0.103)	(0.089)	(0.090)	(0.091)	(0.091)	(0.104)
Observations	3877	3877	3762	3643	3877	3877	3877
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	28.13	22.76	40.42	47.09	32.29	32.29	15.47
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Minute appearances	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: First stage for the estimates of the bilateral framework (match level), Table 11. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–2, 5–7) / the first year available for the instrument to 2018 (in columns 3–4). Dependent variable in the second stage: goal difference for columns 1–4 and 6, its hyperbolic sine transformation in Column 5 and the goals scored by the home team in Column 7. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–7 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 22. Bilateral framework, controlling for coach information, first-stage regressions

	First stage: Dependent variable: bilateral diversity				
	(1)	(2)	(3)	(4)	(5)
	IV	IV	IV	IV	IV
Instrumental variable					
IV, home vs. away	0.008***	0.007***	0.008***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Coach age, home	-0.004*	-0.004**	-0.005**	-0.005**	-0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Coach age, away	0.004*	0.005**	0.005**	0.005**	0.005**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Control variables					
Foreign coach, home	-0.047	-0.053	-0.045	-0.083**	-0.083**
	(0.038)	(0.038)	(0.038)	(0.039)	(0.039)
Foreign coach, away	0.012	0.009	0.000	0.038	0.038
	(0.036)	(0.036)	(0.036)	(0.039)	(0.039)
Log of GDP/capita, home		-0.255**	-0.242**	-0.052	-0.052
		(0.086)	(0.086)	(0.104)	(0.104)
Log of GDP/capita, away		0.326***	0.306***	0.160	0.159
		(0.093)	(0.092)	(0.113)	(0.113)
Stand. dev. appearances, away		-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Stand. dev. appearances, home		0.002***	0.002***	0.002***	0.002***
		(0.000)	(0.000)	(0.000)	(0.000)
Population (mln), home			-0.005***		
			(0.001)		
Population (mln), away			0.006***		
			(0.001)		
Log immig. stocks, 18y lag, home				-0.057**	-0.057**
				(0.019)	(0.019)
Log immig. stocks, 18y lag, away				0.041**	0.041**
				(0.020)	(0.020)
Observations	3873	3873	3873	3568	3568
Kleibergen-Paap LM test	0.00	0.00	0.00	0.00	0.00
Kleibergen-Paap F test	30.97	24.22	32.21	14.64	14.65
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances		Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: First stage for the baseline estimates of the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in columns 4–5). Dependent variable in the second stage: goal difference. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–5 display IV first-stage results, with heteroskedastic robust standard errors in parentheses, clustered at team pair level. For each IV 55 specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Appendix E

Other tables

We report the full table of results for the control function approach reported in the Sensitivity checks section for the bilateral estimations (Table 12 reports the average partial effects).

Table 23. Bilateral framework, goals for

	Dependent variable: home team's goals scored				
	(1) **IV:**Poisson, control function	(2) **IV:**Poisson, control function	(3) **IV:**Poisson, control function	(4) **IV:**Poisson, control function	(5) **IV:**Poisson, control function
Diversity, away	-0.139	-0.253	-0.242*	-0.411*	-0.411*
	(0.139)	(0.161)	(0.143)	(0.225)	(0.226)
RES_FEd	0.115	0.241	0.231	0.390*	0.389*
	(0.139)	(0.162)	(0.145)	(0.226)	(0.227)
Diversity, home	0.281*	0.333*	0.350**	0.389	0.386
	(0.156)	(0.181)	(0.161)	(0.296)	(0.298)
RES_FEo	-0.255	-0.325*	-0.346**	-0.383	-0.381
	(0.156)	(0.180)	(0.161)	(0.296)	(0.298)
Log of GDP/capita, home		0.245*	0.259*	0.143	0.142
		(0.145)	(0.140)	(0.162)	(0.163)
Log of GDP/capita, away		-0.427***	-0.418***	-0.279**	-0.276*
		(0.129)	(0.121)	(0.140)	(0.141)
Stand. dev. appearances, away		-0.003***	-0.003***	-0.003***	-0.003***
		(0.000)	(0.000)	(0.001)	(0.001)
Stand. dev. appearances, home		0.002***	0.002***	0.002**	0.002**
		(0.000)	(0.000)	(0.001)	(0.001)
Population (mln), home			-0.000		
			(0.001)		
Population (mln), away			-0.000		
			(0.002)		
Log immig. stocks, 18y lag, home				0.041	0.041
				(0.034)	(0.034)
Log immig. stocks, 18y lag, away				-0.068**	-0.068**
				(0.027)	(0.028)
Observations	3877	3877	3877	3568	3568
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances	Yes	Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: Coefficients table, relative to Table 12's home score results. Estimates for the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in columns 4–5). Dependent variable: home team's number of goals scored in the top sub-table, Away team's number of goals scored in the top sub-table. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–5 display results from a Poisson, control-function regression, with bootstrapped (2000 reps) standard errors in parentheses, clustered at team pair level. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 24. Bilateral framework, goals against

	Dependent variable: away team's goals scored				
	(1) **IV:**Poisson, Control function	(2) **IV:**Poisson, control function	(3) **IV:**Poisson, control function	(4) **IV:**Poisson, control function	(5) **IV:**Poisson, control function
Diversity, away	0.370^{**}	0.406^{**}	0.331^{**}	0.456^{*}	0.463^{*}
	(0.161)	(0.181)	(0.159)	(0.261)	(0.262)
RES_FEd	-0.363^{**}	-0.417^{**}	-0.341^{**}	-0.463^{*}	-0.469^{*}
	(0.161)	(0.181)	(0.160)	(0.263)	(0.264)
Diversity, home	-0.169	-0.191	-0.203	-0.452	-0.439
	(0.171)	(0.202)	(0.179)	(0.351)	(0.355)
RES_FEo	0.156	0.199	0.212	0.458	0.447
	(0.172)	(0.203)	(0.181)	(0.351)	(0.355)
Log of GDP/capita, home		-0.092	-0.103	0.020	0.027
		(0.162)	(0.152)	(0.191)	(0.193)
Log of GDP/capita, away		0.343^{**}	0.298^{*}	0.173	0.185
		(0.163)	(0.156)	(0.171)	(0.172)
Stand. dev. appearances, away		0.003^{***}	0.003^{***}	0.003^{***}	0.003^{***}
		(0.000)	(0.000)	(0.001)	(0.001)
Stand. dev. appearances, home		-0.003^{***}	-0.003^{***}	-0.002^{***}	-0.002^{***}
		(0.001)	(0.000)	(0.001)	(0.001)
Log immig. stocks, 18y lag, home				-0.081^{**}	-0.080^{**}
				(0.038)	(0.038)
Log immig. stocks, 18y lag, away				0.061^{*}	0.060^{*}
				(0.034)	(0.034)
Observations	3877	3877	3877	3568	3568
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances	Yes	Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: Coefficients table, relative to Table 12's away score results. Estimates for the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in columns 4–5). Dependent variable: home team's number of goals scored in the top sub-table, Away team's number of goals scored in the top sub-table. Columns 1–5 display results from a Poisson, control-function regression, with bootstrapped (2000 reps) standard errors in parentheses, clustered at team pair level. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 25. Bilateral framework, first stage, for home team diversity

	Dependent variable: home team's goals scored				
	(1) **IV:**Poisson, Control function	(2) **IV:**Poisson, control function	(3) **IV:**Poisson, control function	(4) **IV:**Poisson, control function	(5) **IV:**Poisson, control function
IV, away	-0.000	-0.000	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
IV, home	0.010***	0.009***	0.010***	0.007***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Log of GDP/capita, home		-0.411***	-0.391***	-0.225**	-0.227**
		(0.078)	(0.077)	(0.090)	(0.093)
Log of GDP/capita, away		0.009	0.005	-0.040	-0.044
		(0.105)	(0.104)	(0.122)	(0.119)
Stand. dev. appearances, away		-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Stand. dev. appearances, home		0.001**	0.001**	0.001**	0.001**
		(0.000)	(0.000)	(0.000)	(0.000)
Population (mln), home			-0.006***		
			(0.001)		
Population (mln), away			0.001		
			(0.002)		
Log immig. stocks, 18y lag, home				-0.064***	-0.064***
				(0.018)	(0.017)
Log immig. stocks, 18y lag, away				0.004	0.003
				(0.021)	(0.020)
Observations	3877	3877	3877	3568	3568
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances	Yes	Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: Coefficients table, relative to Table 12's home score results. First stage on estimates for the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in columns 4–5). Dependent variable: home team's number of goals scored in the top sub-table, Away team's number of goals scored in the top sub-table. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–5 display linear first-stage results from a second-stage Poisson, with a control-function method, and bootstrapped (2000 reps) standard errors in parentheses, clustered at team pair level. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 26. Bilateral framework, first stage, for away team diversity

	Dependent variable: home team's goals scored				
	(1) **IV:**Poisson, control function	(2) **IV:**Poisson, control function	(3) **IV:**Poisson, control function	(4) **IV:**Poisson, control function	(5) **IV:**Poisson, control function
IV, away	0.012***	0.010***	0.012***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
IV, home	-0.000	-0.000	-0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log of GDP/capita, home		-0.055	-0.051	-0.147	-0.150
		(0.097)	(0.094)	(0.108)	(0.105)
Log of GDP/capita, away		-0.419***	-0.394***	-0.244**	-0.248**
		(0.082)	(0.079)	(0.090)	(0.091)
Stand. dev. appearances, away		0.001**	0.000*	0.001**	0.001**
		(0.000)	(0.000)	(0.000)	(0.000)
Stand. dev. appearances, home		-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Population (mln), home			0.000		
			(0.001)		
Population (mln), away			-0.006***		
			(0.001)		
Log immig. stocks, 18y lag, home				0.015	0.015
				(0.019)	(0.019)
Log immig. stocks, 18y lag, away				-0.050**	-0.051**
				(0.017)	(0.018)
Observations	3877	3877	3877	3568	3568
Team FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes
Minute appearances	Yes	Yes	Yes	Yes	Yes
Geo-political controls					Yes

Notes: Coefficients table, relative to Table 12's away score results. First stage on estimates for the bilateral framework (match level). Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3) / years 1978–2018 (in columns 4–5). Dependent variable: home team's number of goals scored in the top sub-table, Away team's number of goals scored in the top sub-table. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–5 display linear first-stage results from a second-stage Poisson, with a control-function method, and bootstrapped (2000 reps) standard errors in parentheses, clustered at team pair level. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 27. Placebo, baseline estimations for the sake of comparison

	Dependent variable: change in rating of national football team (Elo score)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Variable of interest								
Diversity	2.943**	2.626**	2.807**	2.654*	22.417**	19.178*	23.652**	36.505*
	(1.136)	(1.156)	(1.129)	(1.395)	(11.128)	(10.479)	(11.655)	(20.356)
Population (mln)	0.193	0.293	0.322	0.241	-0.010	0.125	0.140	-0.267
	(0.245)	(0.226)	(0.208)	(0.310)	(0.413)	(0.407)	(0.414)	(0.692)
Stand. dev. appearances		0.276***	0.277***	0.310***		0.292***	0.293***	0.316***
		(0.023)	(0.023)	(0.028)		(0.026)	(0.027)	(0.033)
Log of GDP/capita			8.814	9.903			18.034**	17.371*
			(6.494)	(7.609)			(8.061)	(9.464)
Log immig. stocks, 18y lag				-0.476				1.599
				(1.369)				(1.944)
Observations	1900	1900	1900	1676	1900	1900	1900	1676
Kleibergen-Paap LM test					18.61	18.53	16.66	7.34
Kleibergen-Paap F test					20.74	20.60	17.47	7.69
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Baseline estimates for the placebo analysis. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 5–7) / years 1978–2018 (in columns 4 and 8). Dependent variable: changes in the Elo score of the national team (end vs. beginning of the championship stage). In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * p < .10, *** p < .05, *** p < .001.

Table 28. Placebo analysis: performances in athletics (all medals) and genetic diversity

	Dependent variable: total medals in athletics championship							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Variable of interest								
Diversity	0.046	0.046	0.044	0.017	-0.151	-0.152	-0.213	-1.112
	(0.155)	(0.155)	(0.154)	(0.157)	(0.438)	(0.436)	(0.500)	(0.925)
Population (mln)	0.170*	0.170*	0.169*	0.226*	0.172***	0.172***	0.172***	0.243***
	(0.099)	(0.099)	(0.099)	(0.114)	(0.031)	(0.031)	(0.031)	(0.048)
Stand. dev. appearances		0.000	0.000	-0.001		0.000	0.000	-0.000
		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
Log of GDP/capita			-0.136	-0.095			-0.244	-0.333
			(0.540)	(0.843)			(0.358)	(0.452)
Log immig. stocks, 18y lag				0.001				-0.068
				(0.140)				(0.095)
Observations	1900	1900	1900	1676	1900	1900	1900	1676
Kleibergen-Paap LM test					18.61	18.53	16.66	7.34
Kleibergen-Paap F test					20.74	20.60	17.47	7.69
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates for the placebo analysis. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 5–7) / years 1978–2018 (in columns 4 and 8). Dependent variable: total medals obtained by the national representative athletes in athletics championships. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Table 29. Placebo analysis: performances in athletics (gold medals) and genetic diversity

	Dependent variable: gold medals in athletics championship							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	OLS	OLS	OLS	IV	IV	IV	IV
Variable of interest								
Diversity	0.035	0.036	0.033	0.028	0.074	0.076	0.050	0.004
	(0.059)	(0.060)	(0.058)	(0.061)	(0.185)	(0.185)	(0.208)	(0.345)
Population (mln)	0.058	0.058	0.057	0.076^{*}	0.057^{***}	0.057^{***}	0.057^{***}	0.076^{***}
	(0.036)	(0.036)	(0.036)	(0.042)	(0.012)	(0.012)	(0.012)	(0.017)
Stand. dev. appearances		-0.000	-0.000	-0.000		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)		(0.001)	(0.001)	(0.001)
Log of GDP/capita			-0.112	-0.046			-0.105	-0.051
			(0.239)	(0.387)			(0.159)	(0.199)
Log immig. stocks, 18y lag				-0.012				-0.014
				(0.063)				(0.043)
Observations	1900	1900	1900	1676	1900	1900	1900	1676
Kleibergen-Paap LM test					18.61	18.53	16.66	7.34
Kleibergen-Paap F test					20.74	20.60	17.47	7.69
Team FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates for the placebo analysis. Estimation sample: football national teams from the UEFA affiliation, performing in World Cup and EUROs in the years 1970–2018 (for columns 1–3, 5–7) / years 1978–2018 (in columns 4 and 8). Dependent variable: gold medals obtained by the national representative athletes in athletics championships. In all regressions, we include team and year fixed effects, as well as a stage dummy. Columns 1–4 display OLS results, with heteroskedastic robust standard errors in parentheses, clustered at team level. Columns 5–8 display IV results, with heteroskedastic robust standard errors in parentheses, corrected for arbitrary autocorrelation of degree 1. For each IV specification, we present the p-value from the Kleibergen-Paap Lagrange Multiplier test for the instrument irrelevance, as well as the F-statistics from Kleibergen-Paap F test for weak instruments. Stars correspond to the following p-values: * $p < .10$, *** $p < .05$, *** $p < .001$.

Appendix F

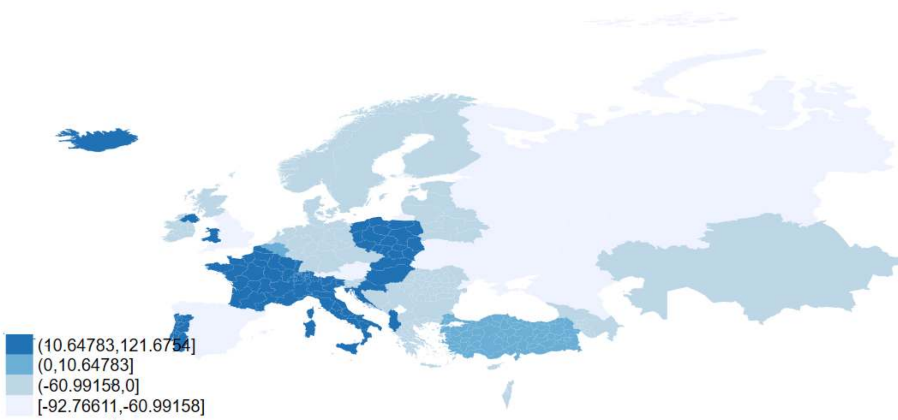
Additional graphs and tables

To complete figures 1 and 3, we present the same cross-sectional maps for the final stages.

Figure 9. Cross-country diversity: descriptive example, EURO 2016, final stage

We present this graph to complete Figure 1. This snapshot represents diversity indices for the final stages of 2016 EURO games. As for Figure 1, the graph broadly presents higher levels of diversity on the Western side of the continent.

Figure 10. Cross-country changes in Elo score: descriptive example, EUROs 2016, final stage



We present this graph to complete Figure 3. This snapshot represents Elo score changes for the final stages of 2016 EURO games. The tournament champion is Portugal, which won a final match against France, the hosting nation.

Table 30. List of countries in the sample, by year in the unilateral sample. Three-letter codes follow the ISO three-letters specification.

ALB	1972 1974 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
ARM	1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
AUT	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
AZE	1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
BEL	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
BGR	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
BIH	2010 2012 2014 2016 2018
BLR	1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
CHE	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
CSK	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994
CYP	1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
CZE	2012 2014 2016 2018
DDR	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990
DEU	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
DNK	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
ENG	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
ESP	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
EST	1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
FIN	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
FRA	1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
GEO	1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
GRC	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
HRV	2010 2012 2014 2016 2018
HUN	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
IRL	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
ISL	1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
ISR	1970 1982 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
ITA	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
KAZ	2006 2008 2010 2012 2014 2016 2018
LTU	1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
LUX	1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
LVA	1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
MDA	1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
MKD	2010 2012 2014 2016 2018
MLT	1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
MNE	2010 2012 2014 2016 2018
NIR	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
NLD	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
NOR	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
POL	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
PRT	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
ROU	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
RUS	1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
SCT	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
SMR	1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
SRB	1998 2000 2002 2010 2012 2014 2016 2018
SUN	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990
SVK	2010 2012 2014 2016 2018
SVN	2010 2012 2014 2016 2018
SWE	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
TUR	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
UKR	1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
WLS	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008 2010 2012 2014 2016 2018
YUG	1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990

List of national teams by year in the Unilateral sample. 3 letter codes follow a ISO 3-letters specification.

List of variables

Table 31. Description of variables, unilateral specifications

Variable name	Variable description	Variable source
Performance measures		
Elo score	Elo score or the team at the end of the stage	Retrieved from eloratings.net
Elo score, computed	Elo score or the team at the end of the stage	Own computation from match level data
Diversity measures		
Diversity	Benchmark Diversity measure, genetic distances are based on dominant populations	Source for surname predictions: forebears.io. Source for genetic distance measures: Spolaore and Wacziarg (2009)
Diversity, apperance	Diversity alternative measure, weighted by minute appearances	Source for surname predictions: forebears.io. Source for genetic distance measures: Spolaore and Wacziarg (2009)
Diversity, SW.	Diversity alternative measure, based on weighted genetic distances	Source for surname predictions: forebears.io; Source for genetic distance measures: Spolaore and Wacziarg (2009)
Team-level variables		
Adversary's diversity	Average diversity level of the teams faced	Own computation from match level data
Adversary's strength	Average Elo score level of the teams faced, measured at the beginning of the stage	Own computation from match level data
Foreign coach	Dummy =1 if the team's manager is foreign	Retrieved from squad-level data on worldfootball.net
Coach age	Age of coach (approximated), computed as year of championship minus year of birth	Retrieved from squad-level data on worldfootball.net
Stand. dev. appearances	Player turnover as computed from the minute appearances	Constructed from squad-level data on worldfootball.net
Stand. dev. squad age	Standard deviation of team members' age	Constructed from squad-level data on worldfootball.net
Squad age	Average of team members' age	Constructed from squad-level data on worldfootball.net
Squad age, squared	Square of squad age.	Constructed from squad-level data on worldfootball.net
Squad size	Number of players in the squad (fixed in the final stages)	Constructed from squad-level data on worldfootball.net
Macroeconomic variables		
Log immig. stocks, 18y lag	Log of the stocks of immigrants, lagged 18 years	World Development Indicators (WDI), International migrant stock, total (SM.POP.TOTL), linear interpolation was conducted on the 5-year-interval data. Complemented by the World Bank's Bilateral Migration Matrix, Özden et al. (2011)
Log of GDP/capita	Log of per capita GDP	National Accounts Section of the United Nations Statistics Division: National Accounts Main Aggregates Database
Population (mln)	Country population size (millions of units)	WDI, SP.POP.TOTL total population; Head et al. (2010)
IV		
IV, 18y lag	Instrumental variable: historical diversity level, 18 years' lag (benchmark lag)	Cline Center for Advanced Social Research. Complemented with the World Bank's bilateral migration matrix (Özden et al. (2011))

Table 32. Variables' description, bilateral specifications

Variable name	Variable description	Variable source
Performance measures		
Goal difference	Goals of team i *home* - Goals of team j. *away*	Mart Jürisoo
Goal difference, hyperbolic sine	Hyperbolic sign transformation of *Goal difference*	*see Goal difference*
Diversity measures		
Bilateral diversity	Diversity score of team i *home* - Diversity score of team j *away*. Benchmark measure, genetic distances are based on dominant populations	Surname predictions: forebears.io. Genetic distance measures: Spolaore and Wacziarg (2009)
Bilateral diversity, appearance	Diversity score of team i *home* - Diversity score of team j *away*. Alternative measure, weighted by minute appearances	Surname predictions: forebears.io. Genetic distance measures: Spolaore and Wacziarg (2009)
Bilateral diversity, SW	Diversity score of team i *home* - Diversity score of team j *away*. Alternative measure, based on weighted genetic distances.	Surname predictions: forebears.io. Genetic distance measures: Spolaore and Wacziarg (2009)
Diversity, home (Diversity, away)	Diversity score of team i *home* (team j *away*)	Surname predictions: forebears.io. Genetic distance measures: Spolaore and Wacziarg (2009)
Team-level variables		
Stand. dev. squad age, home (Stand. dev. squad age, away)	Standard deviation of team i *home* (team j *away*) members' age	Constructed from squad-level data on worldfootball.net
Squad age, home (Squad age, away)	Average of team i *home* (team j *away*) members' age	Constructed from squad-level data on worldfootball.net
Squad age, squared, home (Squad age, squared, away)	Square of squad age, home (*away*)	Constructed from squad-level data on worldfootball.net
Stand. dev. appearances, home (Stand. dev. appearances, away)	Player turnover for team i *home* (team j *away*), as computed from the minute appearances	Constructed from squad-level data on worldfootball.net
Foreign coach, home (Foreign coach, away)	Dummy =1 if the team i *home* (team j *away*)'s manager is foreign	Retrieved from squad-level data on worldfootball.net
Coach age, home (Coach age, away)	Age of team i *home* (team j *away*)'s coach (approximated), computed as year of championship minus year of birth	Retrieved from squad-level data on worldfootball.net
Macroeconomic variables		
Population (mln), home (Population (mln), away)	Team i *home* (team j *away*)'s country population size (millions of units)	WDI, SP.POP.TOTL total population; Head et al. (2010)
Log of GDP/capita, home (log of GDP/capita, away)	Log of per capita GDP for team i *home* (team j *away*)'s country	UN Statistics Division: National Accounts Main Aggregates Database
Log immig. stocks, 18y lag, home (Log immig. stocks, 18y lag, away)	Log of the stocks of immigrants for team i *home* (team j *away*)'s country, lagged 18 years	WDI, International migrant stock (see Unilateral table for details.) Complemented with (Özden et al., 2011)
Adversary's strength, home (Adversary's strength, away)	Average Elo score level of the teams faced, measured at the beginning of the stage	Own computation from match-level data
Contiguity	Dummy =1 if the team i *home* and j *away*) share/ have shared historically a border	Spolaore and Wacziarg (2009)
Same nation	Dummy =1 if the team i *home* and j *away*) are/ have been historically part of the same nation	Spolaore and Wacziarg (2009)
Common language	Dummy =1 if the team i *home* and j *away*) share/ have shared historically an official language	Spolaore and Wacziarg (2009)