Why are We So Good At Football, and They So Bad? Institutions and National Footballing Performance

Meshael Batarfi[1] · James Reade[1]

Abstract

The basic production technology in football is identical for each team that competes. All around the world, a field, goalposts and a ball is all that is required, in addition to players. It's hard to imagine the quality of informal football in public parks, streets and alleys the world over differs much. Yet at each country's highest level, there exists vast quality differences in the national football teams across countries. This paper sketches out broad patterns in this variation in performance, and seeks to understand why some countries are very good, whilst others perform poorly. We investigate a range of macroeconomic, demographic and political explanations, alongside more conventional sporting metrics. We also consider the extent to which they explain the observed variation in footballing performance historically. We find that higher level of GDP per capita helps nations to win more often, but that population hinders this. A more developed domestic footballing structure appears to be helpful too.

Keywords Development · Contests · Sport

JEL Classification O1 · C20 · L83

1 Introduction

Outcomes matter. Economists are interested in the distribution of outcomes, especially if many salient features that matter for outcomes remain fixed. Around the world, people play football on streets, often in any kind of space that's available.

✉ James Reade
   j.j.reade@reading.ac.uk

1   Department of Economics, University of Reading, Reading, UK

Even without grass and goalposts, or even a football, substitutes can be used—"jumpers for goalposts."

Almost every country in the world has a national football team that represents it. National teams play against each other in international competition and the pinnacle of the game is the showpiece event: the World Cup. This takes place every four years.

Every country's youngsters kick balls around, and there is no reason to believe kick-abouts in England or Germany differ much in quality from kick-abouts in Canada or China. Yet, in the 90-year history of the World Cup, only eight countries have won the event.

Indeed, the variation can be stark; in the calendar year of 2011, Bolivia played 16 football matches and won none, while in 1984 France played 12 football matches and won all of them.

As with general economic activity then, performance in football varies across countries, and it is interesting and important to ask why some countries perform better than others. In the spirit of Landes (1990) we ask why are certain countries so good at football and others so bad?

In this paper, we answer this question by using a range of explanatory variables that range from the sporting through to the macro-economic. Although production technology is identical at grassroots informal level, between there and a country's national team is tremendous scope for variation. Footballing infrastructure can vary greatly, from a single amateur league through to numerous professional leagues, youth structures to select and develop talent can vary dramatically, while opportunity costs will differ depending on the level of economic opportunity in a country. Countries may have preferences for a much wider range of sports than just football, and political structures may exert an influence through, say, compulsory military service.

The remainder of this paper is structured as follows: Sect. 2 will expand on our motivation and discuss the relevant literature, Sect. 3 outlines the methodology, Sect. 4 presents the data, Sect. 5 sets out the results, Sect. 6 provides a discussion of our results, and Sect. 7 concludes.

2 Motivation and Literature Review

Neale (1964) described the output of a football match as the "product joint": two teams using identical production technology to produce a single output. The more productive is a team, ceteris paribus, the more likely it wins. The more recent economics literature casts outcomes in terms of Tullock (1980) contest functions. That is, the outcome of a match between team 1 and team 2 is some function of the effort expended by each team, e_1 and e_2, such as:

$$p_1(e_1, e_2) = \frac{e_1^r}{e_1^r + e_2^r},$$

(1)

where p_1 is the probability team 1 wins, and $0 \leq r \leq 1$. The cost of effort, however, for teams will differ, and for national teams will be a function of various national characteristics.

While the production technology for a basic football match is identical across countries, much of the associated infrastructure varies a great deal. Football clubs, at all levels, organise football, and select players. Football clubs then play within leagues, in competition with each other. Neale notes that the economic unit of interest is the sports league, rather than the sports club—at this level externalities relating to the quality of teams are internalised. As such, sports leagues matter, as do the bodies that govern them.

Sports leagues and governing bodies attempt to ensure sufficient competition throughout their structures such that interest remains; the uncertainty of outcome hypothesis proposed by Rottenberg (1956) surely applies to participants as well as spectators. More interesting league structures attract more attention; more attention would be expected to correlate with higher willingness to pay to participate or spectate. Such greater willingness to pay implies greater resources for sports leagues, in principle reducing the cost of effort in (1) for the national team of that country.

Equally, though, sport is, ostensibly, a leisure pursuit. As such, it may be proposed that as any economy develops, and productivity increases, the opportunity cost of leisure rises. Hence conversely, it may be the cost of effort increases with greater income levels in a country.

Moreover, productivity growth is rather limited; teams can only win matches. Teams like France in 1984 may win all of their matches, and from this point further productivity growth cannot occur.[1] Because of productivity gains in other sectors in an economy, the cost of producing a football match has increased over the years, a manifestation of the Baumol effect (Baumol and Bowen 1966). This will be particularly strong in countries where incomes are higher for other types of production where productivity gains occur.

It is thus interesting to explore the relationship between sporting performance and measurable factors both from within sport, and from the broader economies in which football exists. Koning and McHale (2012) incorporate population and income per capita to a simulation model forecasting the winners of the Fifa World Cup. They find both to be positive and significant.

Several studies have examined the economic determinants of national sporting performance and country wealth and development level. Most of the existing literature has found that countries with better life conditions and more resources have a higher ability to win at sport when compared to developing countries with poor life conditions and less available resources (Andreff 2006).

Peeters et al. (2019) consider the role that immigration plays in the development of knowledge capital in countries, via coaches of national teams. They find a positive effect of GDP and population.

---

[1] This is, of course, somewhat debatable. In 1984 France won all matches and won the European Championships, a continental tournament. In 1998 they didn't win all matches, yet won the World Cup, arguably a greater achievement marking greater productivity.

Eber (2003) shows that sport practice and economic wealth go hand in hand. Improvement in sport performance in developing countries are related to both technology transfer and local information spillover. Some developing countries have shown the ability to catch up with developed countries in many different disciplines including sport (Yamamura 2009).

A country's FIFA ranking may be considered as an indicator of the country's development (Gásquez and Royuela 2014). Szymanski (2016) noted that the FIFA ranking of national teams depends negatively on population size and income. Economic prosperity is positively associated with football resource acquisition, which helps to boost football performance (Omondi-Ochieng 2015). Country wealth and population size may contribute to international success for both male and female soccer teams (Bredtmann et al. 2014). In this context, other researchers have found similar results showing income and population having a positive diminishing effect on football success (Leeds and Leeds 2009). Moreover, population size has been found to be strongly correlated with football success, especially in Latin American countries (Macmillan and Smith 2007).

Similar results have been found for women's football. Higher income leads to better availability of resources, such as strong infrastructure and more leisure time, which enables potential athletes to practice more (Hoffmann et al. 2006). African countries with a higher level of GDP and population have also shown better sport performance(Luiz and Fadal 2011). More broadly, Krause and Szymanski (2019) investigate the classic economic development hypothesis of convergence by using international football match outcomes, noting that while production technology differs considerably across industries, and thus countries, the production technology for a football match is the same regardless of the country the match is played in. Most studies have confirmed the significant impact of economic development on sport performance (Andreff 2006). Moreover, GDP was significantly related to the success of football clubs in some European countries (Klobučník et al. 2019).

Other research has suggested that the more elderly members in a community, the more likely that policy makers will pay attention to non-sport sectors instead, such as health (Barros 2006). Wealthy countries with better nutritional levels, more sufficient sport facilities and more qualified coaches have a higher chance of winning more medals in the Olympic games (Barros 2006).

Countries with a large population have more opportunities to win more medals in the Olympics Games, as it is more likely that skilled and talented players will be drawn from a larger population (Andreff 2006). Bernard and Busse (2004) consider the impact of population and income per capita on Olympic games success, finding both have a positive impact.

Most governments profess to be favourably inclined towards sport, and for many reasons this ought to have a positive impact on health, social control and cohesion, diplomacy and economic development (Houlihan and Green 2006). Countries like the UK have created their own charter aiming to obtain excellent sport results at the world class level, whereas other countries, such as France, have modified their regime according to their level of performance in the Olympic games (Nys 2006).

3 Methodology

We carry out an empirical analysis, informed by the above literature and reasoning, to investigate the institutional quality of footballing infrastructure in different countries.

The value of football match outcomes is the detail with which they have been recorded over a very long time period—the first international football match took place in 1872, and over 30,000 have occurred since then. This can be used to evaluate productivity, and sporting output of nations.

We explain variation in win proportions by national football teams using explanatory variables that relate to footballing and sporting aspects of a country, as well as more broader economic and political variables.

We consider how often a country's national team wins in international competition to be our measure of footballing development; the more developed a country, the more often it will win. We scale the number of wins a country has in a season (where we take the conventional northern hemisphere season, which spans two calendar years running from the Autumn of one year through to the Spring of the next year) by the total number of matches that country's team plays.

Teams will win more often, however, the weaker is their opposition, and hence an initial step is to control for opposition strength. We do this by calculating Elo ratings for each team, and we include the average Elo strength of a country's opponents in a season.

We propose a number of hypotheses:

Hypothesis 1 A national team will perform better with more economic resources, since more resource can be devoted to sporting development. That is, the income effect will dominate the Baumol effect.

Hypothesis 2 A national team will perform better with a larger population, since if we assume talented individuals exist with some small probability, then a larger population increases the expected number of talented footballers.

Hypothesis 3 A national team will perform better the higher is its unemployment rate. Football is a leisure pursuit fundamentally, and hence more wealthy nations may have less leisure time as the opportunity cost of leisure is higher. Hence if unemployment is higher, there are more individuals free to play sport and develop, and the lack of job opportunities may encourage young people to pursue a career in sport.

Hypothesis 4 A national team will perform better if that country has a greater military spend. A country whose population must conduct military service will be better trained in team work and discipline, arguably important attributes of successful teams. We measure this by the proportion of GDP spent on military activity.

Hypothesis 5 A country whose national sport is (not) football will perform better (worse), ceteris paribus. We measure this by using information on country performance at the Summer Olympic games, in terms of the number of sports participated in. A country with more medals and participants arguably is a country with more diverse sporting interests.

Hypothesis 6 A country with a more developed domestic football scene will perform better. Countries with more leagues, more clubs and more players, hence a more developed local footballing infrastructure, will perform better. We measure this using information on the number of clubs, and competitions in a country in a season. The more of either, the more developed is the domestic football scene.

We run a linear regression model on an unbalanced panel data of country win proportions over calendar years:

$$w_{it} = X_{it}\beta + e_{it}, \qquad e_{it} \sim N(0, \sigma^2). \tag{2}$$

Here, w_{it} is the win proportion of country i in year t, and X_{it} is a set of explanatory variables including those representing each of the above hypotheses. The error term is assumed to have mean zero, constant variance.

The win proportion is the ratio of the number of wins a country team has in a calendar year to the number of matches played. We focus on wins and ignore draws, even though a draw can often represent an achievement for a team.[2] It remains, however, that the win is the most salient aspect of a football match's outcome. We consider all matches a national team plays in a year, hence friendlies as well as competitive matches. This is because while friendlies may be argued to be less informative since they do not count towards any tournament outcome, they still provide information about a national team's performance. Furthermore, a greater proportion of a national team's matches are friendlies than for domestic clubs. We control for the strength of opposition a country faces, to guard against the possibility that national teams only pick friendly matches against easier teams.

Because of the range of variables we include, we are unable to exploit the full richness of the history of international football; our effective sample starts in 1950 when our macroeconomic data becomes available, and ends in 2017.

We investigate the impact of a range of variables on the win proportion for a country's national team. In Fig. 1 we plot the win proportion against our measure of average wealth in a country's population: the log of real GDP per capita. The plot reveals that the simple relationship between these variables, whilst positive (the red least squares regression line), is very noisy. Hence we add in a range of additional explanatory variables in X_{it}.

We add in variables to measure the six hypotheses listed above, and we include the lagged win proportion; in addition to country fixed effects, this enables us to

---

[2] Indeed, Reade et al. (2020) argue that the draw is what makes forecasting football match outcomes so *strange*, and difficult.

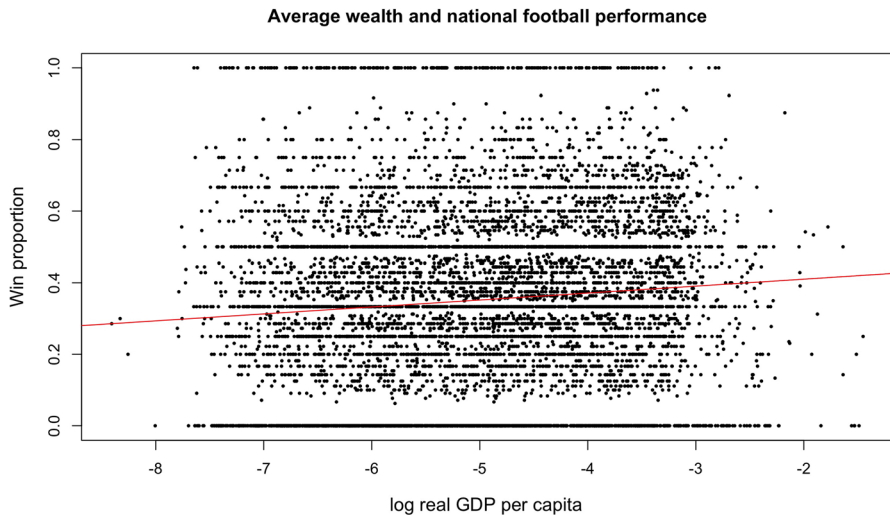Average wealth and national football performance

Fig. 1 Annual income per capita and annual win proportion

capture the persistence in the performance of national football teams; as playing careers at the pinnacle of the game are relatively short, and cohorts overlap, good teams will last for a number of years before gradually being broken up. Equally, bad teams will remain bad for some time even as better players begin to enter the national team.

It is plausible that some of the variables to measure the six hypotheses are related to each other. For example, Olympic success is known to be correlated with the income levels in a country. However, the level of correlation is not sufficiently high between any of our explanatory variables as to warrant concern; by including these distinct variables in the same multi-variate regression model we are able to evaluate the contribution each makes, once the other is conditioned upon.

4 Data

Data on sport is plentiful, as over 30,000 international football matches have taken place since 1872, the year when England and Scotland played the first recognized football match (Wikipedia 2020a). On May 21 1904, after 154 international matches had taken place between nine different countries, The Fédération Internationale de Football Association, or FIFA, was formed as the world's football governing body.[3] FIFA now has more than 200 members (Wikipedia 2020b)

We collect data on football match outcomes back to 1872 from Kaggle.com, and domestic league structures of 113 countries from worldfootball.net. Supplementary

---

[3] These nations were All Ireland, Austria, Belgium, Bohemia, England, France, Hungary, Scotland and Wales.
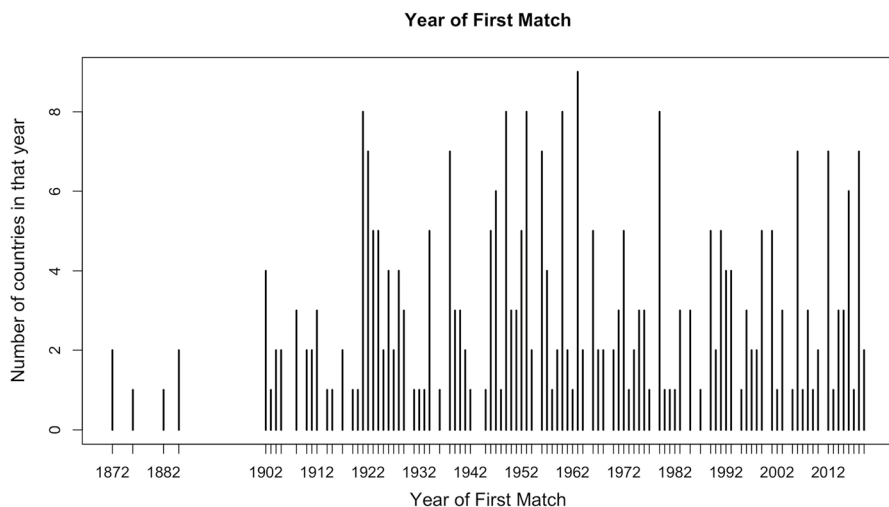
Year of First Match

Fig. 2 Year of first international match by countries

information on domestic league structures was collected from The Rec.Sport.Soccer Statistics Foundation (RSSSF, rsssf.com) and Wikipedia.

Data on GDP per capita, employment and population were collected from Penn World Tables, version 9.1; see rug.nl/ggdc/productivity/pwt/. Data on military spending was collected from the Stockholm International Peace Research Institute (SIPRI, sipri.org).

As many national teams have only relatively recently started playing football, the data constitute an unbalanced panel. Figure 2 provides a graphical representation of the years in which countries played their first international football match. After football's early spread at the start of the twentieth century, the next burst of international starts came after the First World War, with a second burst after the Second World War (Fig. 2).

Data from the Summer Olympics games was collected from the website www. sports-reference.com. The Summer Olympic Games are arguably a better indicator of how 'sporty' a nation is, as the Summer Games have more than 42 sports, whereas the Winter Olympic Games only has 15 sports (Wikipedia 2020c). Furthermore, in general summer sports are more accessible to a wider range of nations due to both the required climactic conditions, and the necessary cost to participate in such sports.

Descriptive statistics are presented in Table 1. From here we see that there is significant variation in the number of observations available from different sources. While the win proportion for countries goes back to their first recorded match, the economic variables are limited to the period between 1950 and 2017 (around 7500 observations). Military spending, as a proportion of GDP, is available for yet fewer observations (6202), and the domestic league arrangements for slightly fewer than that (though these samples do not necessarily overlap).

Table 1 Descriptive statistics

Statistic	N	Mean	SD	Min	Pctl(25)	Pctl(75)	Max
Win proportion	12,065	0.354	0.271	0	0.1	0.5	1
Log real GDP per capita	7847	−4.987	1.195	−8.408	−5.935	−4.046	−1.452
Log population	7847	15.811	1.935	8.384	14.870	17.156	21.067
Employment rate	7422	0.395	0.093	0.117	0.333	0.461	0.898
Military spending (% of GDP)	202	0.028	0.033	0.000	0.013	0.034	1.173
Summer olympics: no. of participants	8012	43.623	83.453	0.000	3.000	40.000	648.000
Summer olympics: total medals	8012	3.659	11.272	0.000	0.000	1.000	174.000
Total number of matches played	12,065	188.982	191.487	1	40	282	1018
First international match	12,065	1,938.250	29.277	1872	1921	1959	2019
Mean opposition elo rating	12,065	1,043.742	106.738	511	981.0	1,113.7	1453
Total clubs in domestic competition	4713	22.643	23.963	0.000	12.000	20.000	292.000
Total number of domestic competitions	4713	2.290	2.943	1.000	1.000	3.000	53.000

It is, nonetheless, the domestic league data that yields novel insights relative to previous studies. We use the total number of clubs in elite competitions in a country per year, and the total number of elite competitions.[4] It points to an internal institutional structure that may provide the basis for international development policy in the area of sport.

In Table 2 the correlation matrix between our explanatory variables, and also the dependent variable, is presented. The largest correlations, of above 0.8, are for the two measures of Olympic participation, and the two measures of the domestic footballing scene. We thus interpret these coefficients with some degree of caution in our results.

5 Results

We present our results in Table 3. We present a number of regression models as we add variables in. Estimation is by ordinary least squares as we have both large *N* and large *T* in our panel dataset. The time period for the regression model in the first column, with the most observations, begins in 1950, the first years of observations for our income and population data. With subsequent additions of explanatory variables, some observations are lost as the coverage of these variables is less complete.

It is important to check the quality of a regression model; we assume that the error term is mean zero, constant variance. In Fig. 3 we present analysis of the quality of the model. From the left plot, which is a histogram of residuals from our

---

[4] We define elite to be those for which data exists, either on worldfootball.net or rsssf.com.

Table 2 Descriptive statistics: correlation matrix

	Win	GDP	Pop	Emp	Mil	Olym prt	Olym tot	Tot match	1st match	Opp elo	Total clubs	Total comps
Win proportion	1	0.100	0.240	0.070	-0.020	0.170	0.120	0.190	-0.170	-0.010	0.190	0.110
Log real GDP per capita		1	-0.130	0.430	0.020	0.280	0.240	0.440	-0.330	0.300	0.320	0.330
Log population			1	0.070	0.020	0.360	0.300	0.390	-0.380	0.280	0.220	0.130
Employment rate				1	-0.150	0.210	0.210	0.320	-0.260	0.150	0.150	0.170
Military spending (% of GDP)					1	0.020	0.030	-0.150	0.060	-0.020	-0.070	-0.110
Summer olympics: no. of participants						1	0.860	0.290	-0.420	0.270	0.270	0.230
Summer olympics: total medals							1	0.220	-0.330	0.200	0.220	0.200
Total number of matches played								1	-0.360	0.470	0.460	0.430
First international match									1	-0.330	-0.290	-0.190
Mean opposition elo rating										1	0.250	0.270
Total clubs in domestic competition											1	0.800
Total number of domestic competitions												1

Histogram of Residuals
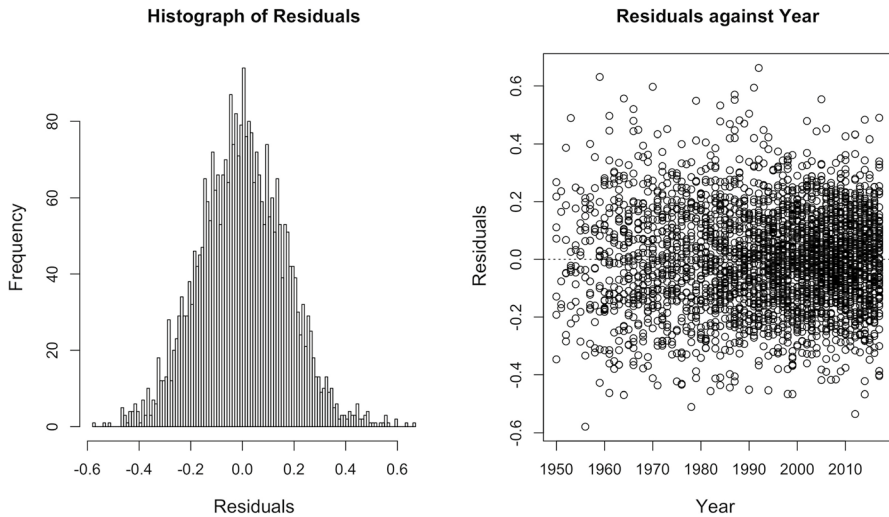
Residuals against Year

Fig. 3 Residual analysis. Left panel is the histogram of residuals from the full model as reported in Table 3, and the right panel is the residuals plotted against the year of observation

preferred, full model, these are symmetric around zero, with no evidence of outliers. From the right plot, residuals against time, no clear patterns stand out, suggesting that most systematic variation in our variable of interest has been included in the model.

In the first column we present a regression model of win proportion on per capita GDP, population and the employment rate. GDP per capita has a positive and significant effect on win proportion, as does the population—richer and more populous countries win more. The employment rate is insignificant. The coefficient of variation, R^2, is 0.071, suggesting that GDP per capita, population and the employment rate are able to explain about 7% of the variation in win ratios recorded by countries.

In the second column, we add fixed effects. With country and year fixed effects, the population variable becomes negative, while the GDP per capita effect is barely affected. The employment rate remains insignificant. In this column, R^2 increases to 0.237, suggesting that the fixed effects alone, included to capture unobserved heterogeneity, account for about 23% of variation in win ratios. While it is commonplace to estimate with fixed effects, it is worth considering their role here explicitly. They enable differences in means over given years and countries, for unobserved effects, to be accounted for. It is likely that adding fixed effects will reduce the explanatory power of other variables, but this should not be the primary concern, provided, as is the case here, the overall explanatory power of the model increases. By including more variables, more discernment is afforded in terms of which variables are causing which variation in the dependent variable.[5]

---

[5] A related concern is that as win proportions must sum to 1, that the interpretation of fixed effects is confusing. In principle, however, all series could be scaled such that they must sum to a particular number. Fixed effects enables unobserved factors that cause differences between countries in their win proportions to be represented in our model.

Table 3 Regression results

| | Dependent variable: | | | | | |
| | Win proportion | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	−0.023	0.673***	0.843***	0.910***	3.065***	30.957***
	(0.029)	(0.182)	(0.216)	(0.222)	(1.001)	(9.363)
Log real GDP per capita	0.029***	0.028***	0.043***	0.051***	0.034***	0.031**
	(0.003)	(0.009)	(0.011)	(0.011)	(0.010)	(0.015)
Log population	0.034***	−0.023*	−0.037**	−0.038**	−0.072***	−0.013
	(0.002)	(0.013)	(0.015)	(0.016)	(0.015)	(0.019)
Employment rate	−0.033	0.077	0.163**	0.157*	0.307***	0.344***
	(0.032)	(0.068)	(0.080)	(0.081)	(0.077)	(0.107)
Military spending (% of GDP)			−0.077	−0.028	0.038	−0.079
			(0.122)	(0.123)	(0.117)	(0.130)
Summer olympics: no. of participants				0.0001*	0.0002***	0.0001
				(0.0001)	(0.0001)	(0.0001)
Summer olympics: total medals				−0.001	−0.001*	−0.001
				(0.001)	(0.001)	(0.001)
Lagged win proportion					0.061***	0.069***
					(0.013)	(0.016)
Total number of matches played					0.0003***	0.0003***
					(0.0001)	(0.0001)
First international match					−0.0004	−0.015***
					(0.001)	(0.005)
Mean opposition elo rating					−0.001***	−0.001***
					(0.00004)	(0.00005)

Table 3 (continued)

	Dependent variable:					
	Win proportion					
	(1)	(2)	(3)	(4)	(5)	(6)
Total clubs in domestic competition						0.001**
						(0.0003)
Total number of domestic competitions						−0.0003
						(0.002)
Observations	7422	7422	5876	5735	5732	3436
Country fixed effects	N	Y	Y	Y	Y	Y
Year fixed effects	N	Y	Y	Y	Y	Y
R^2	0.071	0.237	0.271	0.276	0.347	0.368
Adjusted R^2	0.071	0.211	0.243	0.248	0.322	0.334
Residual SE	0.233 (df = 7418)	0.215 (df = 7180)	0.202 (df = 5662)	0.200 (df = 5520)	0.190 (df = 5513)	0.173 (df = 3263)
F statistic	189.773*** (df = 3; 7418)	9.244*** (df = 241; 7180)	9.875*** (df = 213; 5662)	9.850*** (df = 214; 5520)	13.462*** (df = 218; 5513)	11.032*** (df = 172; 3263)

*p < 0.1; **p < 0.05; ***p < 0.01

In the third column we add in military spending. This has a negative effect, but is insignificant, but it reduces the available sample by almost 2,000 observations. The effect of this is that the employment rate becomes positive and significant, and the population rate becomes significant (and negative). The effect of per capita GDP also increases in magnitude.

In the fourth column we add in Olympic performance. The number of participants a country sends to a Summer Olympics is significant and positive, but the total number of medals is not, although it is negative. Hence the sportiness measured by the number of participants, rather than the success measure by medals, matters.

In the fifth column, football specific variables are added: lagged win percentage, the number of matches played by a country, the year of its first match, and the mean strength of the opposition (as measured by mean Elo (1978) ratings). The lagged win proportion is significant, which establishes that persistence exists in footballing performance. This is inherently sensible since a team can be thought of as a stock concept with flows of younger players into the team, and flows of older players out of it. A country may have a 'golden generation' of players who lead to better performance, but once they begin to retire, performance levels fall. Country experience is positively significant, suggesting that with each extra match of experience, a team's win proportion will increase by 0.0003. The average strength of opposition in a given year also matters for win proportions; there will be natural variation in the quality of opposition a team can face; World Cups and regional tournaments like UEFA's European Championships, only occur every four years, and so outside of these years a team will likely face a lower quality of opposition.

Finally, in column six we add in domestic league structures. The total number of elite clubs in existence has a positive effect, while the total number of competitions is negative but insignificant. Also of note is that while almost all coefficients remain very similar between columns (5) and (6), there are some notable changes. This is likely more the effect of losing around 2,000 observations in adding the domestic league structure variables into the equation, than of the adding of the variables. For example, the population effect becomes insignificant, though still negative. The impact of military spending also turns negative, albeit still insignificant. The impact of summer Olympic participation becomes insignificant, although still of the same sign, while finally, the impact of a country's year of first match (see Fig. 2) jumps in size and becomes significant. It is more likely that this reflects the characteristics of the countries that remain in the sample, but nonetheless it is plausible that a country that began playing football earlier would have more accumulated knowledge capital, and hence win more matches.

The adjusted R^2 for this final equation is 33.7%, suggesting that our model accounts for about a third of the observed variation in win proportions. The same model without fixed effects has an adjusted R^2 of 0.2, implying that our additional variables on top of those in column (1), do account for some of the unobserved heterogeneity, as the fixed effects only account for 13% of unobserved variation in our final model.

Throughout the GDP per capita effect is positive and significant, while the population effect becomes insignificant, and the employment rate effect becomes positive and significant.

6 Discussion

In this Section we provide some discussion of our results, and return to our original hypotheses from Sect. 3 to frame them.

The results presented in Table 3 suggest that for a one log-point increase in GDP per capita, win proportions increase by three to four percentage points. The finding that average wealth does increase the success of national football teams need not necessarily be surprising, and indeed confirms our Hypothesis 1.

The finding of a negative or insignificant population effect is intriguing, particularly given that (Peeters et al. 2019) find a significant and positive impact. Their dependent variable is Elo ranking, however, rather than a country's win proportion in a year. As such, our Hypothesis 2 is not upheld.

Hypothesis 3 is only indirectly measured, since unemployment data is not available across the same wide sample as the employment rate. Our finding of a positive and significant effect of the employment rate is contrary to Hypothesis 3, since it suggests that a country does better at football the more of its population is employed.

Military spending was insignificant in all of our regression models, suggesting that Hypothesis 4 is not upheld, namely that a higher military spend fosters more discipline and fitness and hence better sporting performance.

Our Hypothesis 5 suggests that the larger the number of participants at the Summer Olympics, the more diversified is sporting preferences in a country, and the less likely is that country to be successful at football. However, the effect of the number of participants is positive on the footballing win proportion, albeit insignificant, counter to this.[6] This may be a feature of the relatively high correlation between these two variables; however, while they are of opposite signs, they are not of offsetting magnitude. However, the level of success, measured via medals, is negative if insignificant, suggesting that greater success at a range of sports means footballing success is lower.

A strongly significant, and positive, impact of the total number of clubs suggests that the richness of the domestic football scene matters greatly, supporting our Hypothesis 6. It is total clubs, rather than competitions that matter—indeed, the number of competitions is negative, if insignificant. Again, this opposite sign between these variables may reflect the relatively high correlation between them. However, a reasonable explanation can be provided; it suggests that an increasing number of competitions likely distracts teams by adding more games to the calendar, tiring players who then play for the national team. A common thread of discussion, particularly in England, revolves around the number of matches being played due to the number of competitions.

As a final part of our analysis, we present two sets of residuals from the model. We present the residuals for two countries: England and France.[7] Residuals display the variation left unaccounted for in national team performance after all of our explanatory variables have been taking into account. As such, the residuals should

---

[6] It is significant in column (5), before the domestic football scene variables are added.

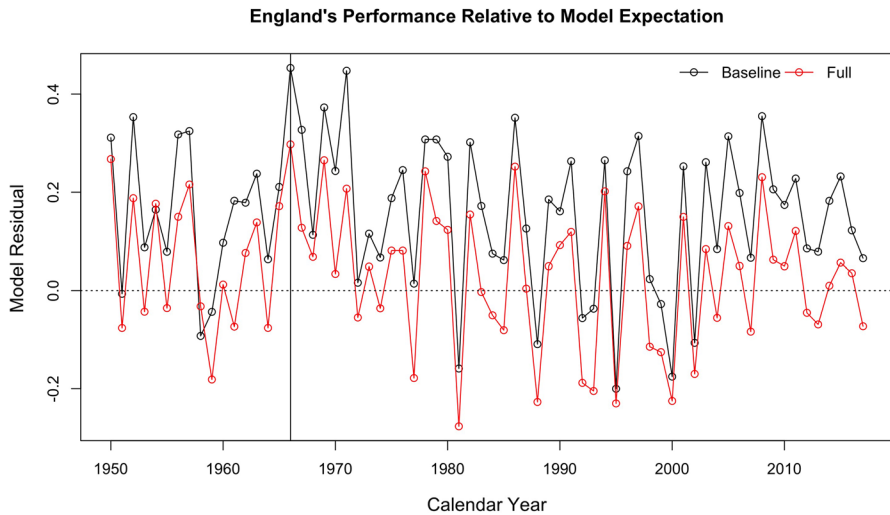[7] Further residual plots for countries are available on request.

England's Performance Relative to Model Expectation

Fig. 4 Residuals from England's performance

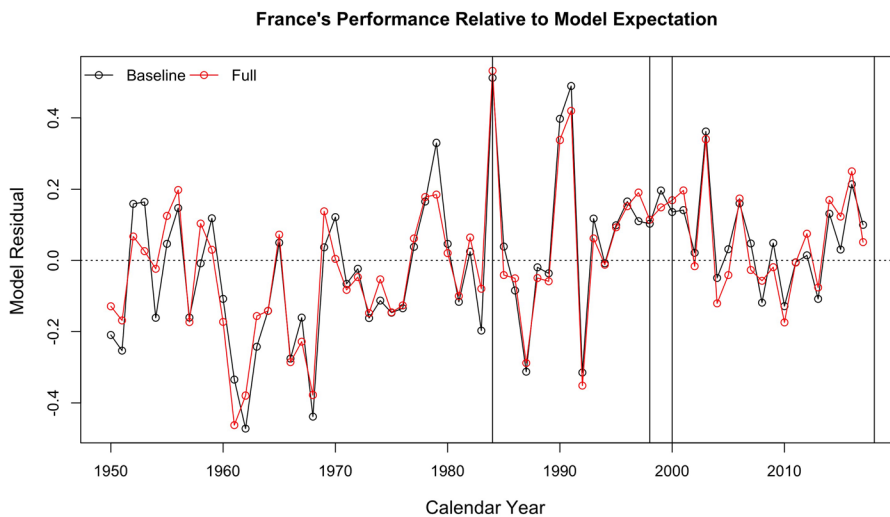**France's Performance Relative to Model Expectation**

Fig. 5 Residuals from France's performance

display little or no systematic pattern. In the Figs. 4 and 5 we plot in black the resid-
uals from the baseline model (column 1), as well as the full model (column 6). The
full model residuals in both cases are mean zero, and constant variance, from a cas-
ual glance.

Another aspect of residual series is that while none should be *too* large, the larger residuals still do reveal the more unexpectedly strong performances in a given year. In England's case in Fig. 4, the largest residual is IIT the year 1966, which corresponds with the single year in which England won the World Cup. In France's performance in Fig. 5, its largest positive residual is 1984, the year it won every single match, including the UEFA European Championships, France's first international tournament.

7 Conclusion

In this paper, we investigate variation in observed outcomes for countries playing football, in order to try and explain why such great variation exists at national levels between countries, yet does not at the very grassroots level of the game. We find that GDP per capita has a significant positive impact on wining proportion, though population size might decrease the chances of winning, the more populous the country is, the less chances they have to win. There was no impact of military spending or country experience or summer total participation, but the total matches have shown negative significant effect. The total number of football clubs in a country has a positive impact on a national team's performance, while the number of domestic competitions has a negative, though insignificant, effect.

References

Andreff, W. (2006). Sport in Developing Countries. In W. Andreff, S. Szymanski (Eds.), *Handbook on the Economics of Sport*, chapter 30. Edward Elgar Publishing. https://ideas.repec.org/h/elg/eechap/3274_30.html.

Barros, C. P. (2006). Local government and regional development in sport. In W. Andreff & S. Szymanski (Eds.), *Handbook on the economics of sport*. Cheltenham: Edward Elgar Publishing.

Baumol, W. J., & Bowen, W. G. (1966). *Performing arts-the economic dilemma: a study of problems common to theater, opera, music and dance*. Cambridge: MIT Press.

Bernard, A. B., & Busse, M. R. (2004). Who wins the Olympic games: economic resources and medal totals. *Review of Economics and Statistics*, *86*(1), 413–417.

Bredtmann, J., Crede, C.J., & Otten, S. (2014). The effect of gender equality on international soccer performance. *Ruhr Economic Paper*, (501).

Eber, N. (2003). Sports practice, health, and macroeconomic performances: an endogenous growth model. *Journal of Sports Economics*, *4*(2), 126–144.

Elo, A. E. (1978). *The rating of chessplayers, past and present* (Vol. 3). London: Batsford.

Gásquez, R., & Royuela, V. (2014). Is football an indicator of development at the international level? *Social Indicators Research*, *117*(3), 827–848.

Hoffmann, R., Ging, L Chew, Matheson, V., & Ramasamy, B. (2006). International women's football and gender inequality. *Applied Economics Letters*, *13*(15), 999–1001.

Houlihan, B., & Green, M. (2006). The changing status of school sport and physical education: explaining policy change. *Sport, Education and Society*, *11*(1), 73–92.

Klobučník, M., Plešivčák, M., & Vrábel', M. (2019). Football clubs' sports performance in the context of their market value and GDP in the European Union regions. *Bulletin of Geography, Socio-Economic Series*, *45*(45), 59–74.

Koning, R. H., & McHale, I. G. (2012). Estimating match and World Cup winning probabilities. In *International handbook on the economics of mega sporting events*. Edward Elgar Publishing, Cheltenham

Krause, M., Szymanski, S. (2019). Convergence vs. the middle income trap: the case of global soccer. *Applied Economics*, 51(27):2980–2999. http://www.melanie-krause.de/assets/pdf/Soccer-Convergence-Draft.pdf.

Landes, D. S. (1990). Why are we so rich and they so poor? *The American Economic Review*, *80*(2), 1–13.

Leeds, M. A., & Leeds, E. M. (2009). International soccer success and national institutions. *Journal of Sports Economics*, *10*(4), 369–390.

Luiz, J. M., & Fadal, R. (2011). An economic analysis of sports performance in Africa. *International Journal of Social Economics*.

Macmillan, P., & Smith, I. (2007). Explaining international soccer rankings. *Journal of Sports Economics*, *8*(2), 202–213.

Neale, W. C. (1964). The peculiar economics of professional sports. *Quarterly Journal of Economics*, *78*(1), 1–14.

Nys, J.-F. (2006). Physical activity, sport and health. In Andreff, W., Szymanski, S. (Eds.), *Handbook on the Economics of Sport*. Edward Elgar Publishing.

Omondi-Ochieng, P. (2015). Gross National Income, football workers and national football team performances. *Team Performance Management*.

Peeters, T.L.P.R., Mills, B.M., Pennings, E., & Sung, H. (2019). Manager migration, learning-by-hiring, and cultural distance in international soccer. *Global Strategy Journal*.

Reade, J.J., Singleton, C.A., & Brown, A. (2020). Evaluating strange forecasts: the curious case of football match scorelines. *Scottish Journal of Political Economy*, Forthcoming.

Rottenberg, S. (1956). The baseball players' labor market. *The Journal of Political Economy*, *64*(3), 242–258.

Szymanski, S. (2016). Professional Asian football leagues and the global market. *Asian Economic Policy Review*, *11*(1), 16–38.

Tullock, G. (1980). Efficient rent seeking. In *Toward a theory of the rent-seeking society*. A&M University Press, College Station.

Wikipedia. 1872 Scotland v England football match — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=1872%20Scotland%20v%20England%20football%20match&oldid=950216659, (2020a). [Online; Accessed 12-June-2020].

Wikipedia. FIFA—Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=FIFA&oldid=963895099, (2020b). [Online; Accessed 25-June-2020].

Wikipedia. Olympic Games—Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Olympic%20Games&oldid=963876108, (2020c). [Online; Accessed 25-June-2020].

Yamamura, E. (2009). Technology transfer and convergence of performance: an economic study of FIFA football ranking. *Applied Economics Letters*, *16*(3), 261–266.