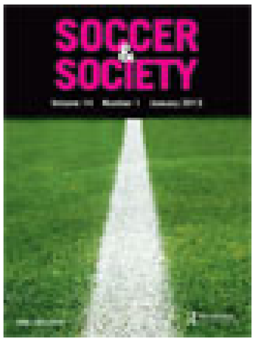

Soccer & Society

ISSN: 1466-0970 (Print) 1743-9590 (Online) Journal homepage: http://www.tandfonline.com/loi/fsas20

Economic, population and political determinants of the 2014 World Cup match results

Gregory T. Papanikos

To cite this article: Gregory T. Papanikos (2015): Economic, population and political determinants of the 2014 World Cup match results, Soccer & Society, DOI: 10.1080/14660970.2015.1067799

To link to this article: http://dx.doi.org/10.1080/14660970.2015.1067799

Published online: 24 Jul 2015.

Soccer & Society, 2015
http://dx.doi.org/10.1080/14660970.2015.1067799

Economic, population and political determinants of the 2014 World Cup match results

Gregory T. Papanikos*

ATINER, Athens, Greece

In this study, it is found that economic, population and political variables are significant determinants of the goal differences in the 2014 football World Cup matches after taking into consideration the football strength of the two teams on the field. It is shown that the strength difference of the two teams had a strong positive non-linear impact on goal differences and that the goal difference increased with strength but at a decreasing rate. Non-linear impact was also evident with regard to the political conditions of one country relative to the rival country. Furthermore, per capita income and population have a positive linear impact. A number of dummy variables were also found to have a non-significant impact on the game results. It was noted that only the variable that relates to the continent of the football teams had a positive impact.

Introduction

In a recent study published in this journal,[1] it was argued that the 2014 World Cup presents many opportunities for research. This study is a contribution to this research. Football is the most popular sport in the world. Every four years, 32 national football teams compete in a final World Cup series. One team wins the trophy and becomes the world champion for the next four years. These 32 teams qualify in preliminary continental rounds with the participation of over 200 national teams. These rounds start three years before the World Cup finals. One country hosts the finals selected in advance by the Fédération Internationale de Football Association (FIFA), the governing body of world football. In 2014, the World Cup finals were held in Brazil, and Germany won the trophy. The 32 national teams were divided into eight groups of four countries. In each group, every team played three games and the first two qualified for the next knockout round. In total, 64 games were played. Every team plays at least three games, and the teams which qualify for the semi-finals play a maximum of seven matches.

This study examines the economic, demographic and political determinants of the goal differences in the 2014 World Cup matches taking into account the relative strength of each national team. Two approaches have been used to analyse match results and team strengths. The first approach uses Poisson models to predict match outcomes, including which team will win the World Cup.[2] Probabilities are based on the a priori information retrieved from bookmakers' odds.[3] This approach looks very much like the Delphi method, an approach that assigns probabilities based on expert opinions which are revised through a Bayesian process. In football, experts may include coaches, players, journalists and football industry analysts.

*Email: gtp@atiner.gr

The second approach uses regression models. The dependent variable is usually the number of goals scored by each team in a match. Explanatory variables are the strength of the opponent and other variables including country-specific characteristics such as economic, demographic and political indicators of the opponent's country. They use different explanatory variables including economic and population indicators of countries of the opposing teams.[4]

Very similar to the second approach are models which aim at explaining the strength differences of national football teams as these are measured almost every month by FIFA itself.[5] These models developed a regression model to explain the FIFA ranking differences using as regressors economic, population, political and a number of dummy variables.[6] On the other hand, models used FIFA rankings as an explanatory variable in a regression model with a dummy variable as the regressand (one for a win and zero otherwise).[7]

The approach adopted here is very similar to regression models. The dependent variable is goal differences. This variable captures better the strength differences of two football teams. In addition, the distribution of goal differences is proportional to relative differences of economic, population and political conditions that exist in the countries of the two teams. For example, if the per capita income of country **A** is $\boldsymbol{x}$ times higher than that of country **B**, then not only is the richer country expected, *ceteris paribus*, to win, but to win with a goal difference which is proportional to $\boldsymbol{x}$. Similar arguments can be made for population and political determinants. The purpose of the regression model here is not to predict outcomes but to explain the determinants of the goal differences of the 64 matches played in the 2014 World Cup.

Unlike most studies that use goals scored by each team, this study uses goal differences because this is the essence of a football match, i.e. to put at least one more goal in your rival team's net than the number of goals they put in your own net. However, there is a good statistical reason why the goal difference should be preferred to a model that explains the number of goals scored by every team in the tournament.[8] First, the correlation between the scores of the two teams is removed and second the model does not assume Poisson marginal distributions. However, as explained later, statistical problems do arise, if the distribution of goal differences is not normally distributed.

The basic conclusions of the study are as follows. FIFA rankings have a positive and non-linear impact on goal differences. The ratios of per capita GDP, as well as population and political variables of the two opponents have a statistical significant positive impact on the goal differences. Higher per capita income relative to the rival team's income results in a higher goal difference. Proportionally larger populations have a positive impact on goal differences. Similarly, national teams that come from more open and democratic countries relative to the rival team's country had larger goal differences in the 2014 World Cup matches.

The study is organized in five sections, including this introduction. The second section discusses the model of goal differences. The third section presents descriptive and summary statistics of the 2014 World Cup match results. The fourth section analyses the empirical results, and the final section concludes.

A model to explain goal differences in a world cup final competition

In the proposed model, goal differences (GD) in the ith match between team $\boldsymbol{A}$ and team $\boldsymbol{B}$ depend on:

1. FIFA ranking differences of country A's and country B's national football teams. This measures strength differences and football abilities of the two teams on the pitch.
2. Economic differences between the two teams of a match as these are measured by country A's per capita income ($YCAP_A$) relative to country B's per capita income ($YCAP_B$).
3. The population of country A (POP_A) relative to the population of country B (POP_B).
4. The level of democracy in country A (DEM_A) relative to country B (DEM_B), and
5. Dummy variables (DUMMIES) that account for the conditions of the 2014 World Cup such as Brazil's home advantage, the differences in team performance due to the climatic conditions of the continent, the confederation of the teams and the round of the match.

The general functional form of the model can be specified as follows:

(1) $GDi = F(FD_{AB}, YCAP_A/YCAP_B, POP_A/POP_B, DEM_A/DEM_B, DUMMIES)$

$i = 1, 2, 3, \ldots. 64.$

The FIFA ranking difference is expected to have a positive impact. However, the beauty of football lies in results which are unexpected. Even weak teams have a chance to beat the strongest team of the world. Good luck is required, but also a good performance of the weak team and a bad performance of the strong team on the day might result in very surprising results. Startling is also the result of two teams that are of equal strength but which in a specific match and for the same reasons of bad–good performance, the goal difference is very large, say three goals or more. For example, in the 2014 World Cup, the first match was between two teams of about the same strength, Spain and the Netherlands, the very two teams that played in the 2010 World Cup final. Actually, Spain (the 2010 World Champions) was ranked higher than the Netherlands. The goal difference of the match was four goals in favour of the Dutch team. But the biggest surprise of all was the result of one of the two games in the semi-finals. Germany hammered Brazil by the shocking goal difference of six goals. These results make football a beautiful and exciting game but as we shall see below, such outcomes are a nightmare for bookmakers and statisticians that use regression models to predict results.[9]

Is the relationship between football strength and goal difference linearly related? Does a *ceteris paribus* doubling of the strength difference double the goal difference of a match? For a number of reasons, the goal difference is expected to increase with strength but at a decreasing rate. First, the motivations which imply lower or higher effort are different when the strength difference is increasing. A much weaker team playing a much stronger team puts more effort into the game for a number of reasons including the weaker team's players' needs for world recognition which eventually may result in better contract opportunities after the World Cup. This recognition will definitely come if they play as good as the players of the stronger team. Even if they are beaten by the stronger team, the margin of goal difference is very important. Second, the players of the stronger team have the opposite strategy. They are expected to win and putting an extra effort into building a goal difference does not make them much better off. Third, a strong team when it plays a weaker team may use different line-ups. The best players are protected for more competitive matches, especially if the strong team has either qualified for the next round or has built a safe goal difference early in the match. The coach of the team protects his star players from exhaustion, injury and the risk of being shown a card, which would exclude them from the next most important match.

Thus, the strength difference of two teams of a match is expected to have a positive impact on the goal differences but at a decreasing rate. Doubling the strength difference, increases the goal difference, but the increase is less than double. This will affect how the strength variable will enter the regression model. A polynomial of strength difference is fitted into a model of goal differences, and the statistical significance of the parameter estimates will determine the degree of the function. As shown below (see Figure 2), the graphical examination of the goal and strength difference of teams played against each other appears to have two turning points. In such a case, a polynomial of strength differences of degree three might fit the data better.

The per capita income of one team relative to the per capita income of the opponent is expected to have a positive impact on goal differences. Rich countries can afford to spend more on the development of football. If the national football team or football in general is considered a public good,[10] then as per capita income increases, more can be spent to improve the performance of the national football team. This means better training facilities a better coach, better management, etc. If football is considered a normal good, then the income elasticity of football demand is positive. Few people would consider the national football team a necessity, and thus, the elasticity might be close to one or greater than one, in which case the national football team becomes a luxury good. In the latter case, an increase in (per capita) income will increase proportionally more the demand for national football. In the long run, countries with higher per capita income will spend more in building strong national teams relative to countries with lower per capita income. This is an additional impact to that of the strength of a national team based on the quality of players at any given point in time. Money by itself cannot build a strong national team. Good players and in particular football stars make the difference and money helps.

If football stars and good players are produced through a stochastic probabilistic process, then the higher the population of a country, the higher the probability that the national football team will have more stars and more good players than a team from a country which is less populated. All countries that won the World Cup trophy have populations that exceeded forty million people.

Strong countries with healthy economies and large populations may not produce a good national football team if the political, social and cultural conditions of the country are not conducive. Does democracy and civil liberty promote the building of a strong national team? Some people have argued that dictators have used national football to strengthen their stay in power. Brazil has been a good example. Football is more than a game and politics have played an important role, as is evident from the many demonstrations that took place prior and during the 2014 World Cup games.[11] On the other hand, more democratic countries with political freedom and an efficient and non-corrupt government may build stronger national teams. This will be empirically tested using an index of democracy.

From a theoretical point of view, the above arguments of the economic, population and political impacts on goal differences are not very persuasive. After all, these variables explain the current strength of a national football team. Thus, a strength measure should be sufficient to explain goal differences. A strong national football team is built if a country is richer, has a larger population and is politically stable. Is there anything more that these variables can add to a model of goal differences when football strength is one of the explanatory variables? The answer is 'yes' because the World Cup final games are different. The measurement of the strength of a national football team is based on previous results which are not matches of a World Cup final series. Unlike the preliminary rounds of World Cup matches, the finals take place every four years over a period of one month in a specific country and are watched by billions of people. The media coverage is unprecedented. Countries and players have a motivation to show their best. Players' rewards are different. Rich countries can afford to reward their national team players with more money. This is enhanced even more if the home population is larger and the political system is more democratic and less corrupt. Higher home population means larger awards for players from private sources such as advertising, sponsoring and wealthier individual contracts. A political free country with democratic institutions allows more efficient exchange transactions. It is a well-known tenet of economic analysis that free trade maximizes labour rewards, i.e. players will be paid the value of their marginal product. A good performance in a World Cup shifts the individual's players marginal product curve to the right. This might explain why these variables might add to the explanatory power of a model which includes the strength of a team as measured by FIFA.

The empirical specification of the model depends on how these explanatory variables enter into the equation to be estimated. It has already been mentioned that the FIFA ranking difference has a non-linear impact on goal differences. A diagrammatic analysis of the data will give an indication of the type of linear and non-linear relation which exists between the goal difference variable and the explanatory variables. This is taken up in the next section which also presents the raw data and summary statistics of the 32 teams which competed in the 2014 World Cup finals.

Data presentations and analysis

Table 1 reports the basic data of the analysis for each one of the 32 countries whose football teams qualified to play in the 2014 World Cup finals. The strength of a team is measured using the FIFA ranking of June 5 2014. This is the latest measurement before the start of the World Cup on June 12 2014. FIFA ranking is the most known and used indicator of national football team strength.[12] GDP per capita is in US Dollars and is actual 2013 figures or the most recent available provided by the World Bank. Population data are based on the United Nations estimates of 2014. The democracy index is retrieved from the Economist Intelligence Unit of Democracy Index of 2012. The latter changes very little over the years.

The strongest team before the 2014 World Cup was Spain, the 2010 World Champions in the South Africa games. Spain in 2014 failed to qualify from the preliminary round to the round of 16 and they returned back home very early in the games. The weakest team of the 2014 games was Australia, number 62 of the FIFA list with 526 points. The FIFA list includes over 200 countries. This means that the 32 teams which participated in the finals belonged to the 30% strongest national football teams as ranked by UEFA. The 32 teams represent 15% of the total FIFA members. FIFA's system of preliminary rounds is organized around six continent

Table 1. The 32 finalists of the 2014 Football World Cup.

	Country	FIFA ranking	FIFA points	GDP per capita	Population	Democracy index
1	Algeria	22	858	13,304	38,700	383
2	Argentina	5	1175	18,749	42,670	684
3	Australia	62	526	43,550	23,551	922
4	Belgium	11	1074	40,338	11,199	805
5	Bosnia and Herzegovina	21	873	9632	3792	511
6	Brazil	3	1242	15,034	202,818	712
7	Cameroon	56	558	2711	20,387	344
8	Chile	14	1026	21,911	17,773	754
9	Colombia	8	1137	12,371	47,670	663
10	Costa Rica	28	762	13,872	4667	810
11	Côte d'Ivoire	23	809	3012	23,202	100
12	Croatia	18	903	20,904	20,387	693
13	Ecuador	26	791	10,469	15,778	578
14	England	10	1090	50,566	53,010	812
15	France	17	913	36,907	65,931	788
16	Germany	2	1300	43,332	80,716	834
17	Ghana	37	704	3974	27,043	602
18	Greece	12	1064	25,651	11,123	765
19	Honduras	33	731	4591	8555	584
20	Iran	43	641	15,586	77,564	198
21	Italy	9	1104	34,303	60,783	774
22	Japan	46	626	36,315	127,090	808
23	Korea South	57	547	33,140	50,424	813
24	Mexico	20	882	16,463	119,713	690
25	Netherlands	15	981	43,404	16,858	899
26	Nigeria	44	640	5863	178,517	377
27	Portugal	4	1189	25,892	10,478	792
28	Russia	19	893	24,120	146,049	374
29	Spain	1	1485	32,103	46,610	802
30	Switzerland	6	1149	53,705	8161	909
31	Uruguay	7	1147	19,590	3286	817
32	USA	13	1035	53,143	318,349	811

football federations, and some continents have weaker teams than others. In this way, weaker teams in the FIFA ranking (but not too weak) have a possibility to make it to the finals.

Summary statistics are given in Table 2. The average FIFA points of the 32 teams were 933 with a standard deviation of 240 points. From the 32 countries, Cameroon had the lowest per capita income, 2711 US dollars, and Switzerland had the highest per capita income of 53,705 US dollars. The average per capita income was 24,516 dollars with a standard deviation of 15,470 dollars. The average country had a population of 58.8 million people and a standard deviation of 70.3 million. The smallest country to compete was Uruguay with 3.3 million people, and the largest was USA with 318.3 million people. Variations do exist in the democracy index as well, with a standard deviation of 210 points and a mean value of 669 points.

The next summary statistics reflect the variables used in the regression analysis in the following section. The actual raw numbers for each of the 64 games are given in Appendix 1. FIFA ranking difference between the two teams playing a match was on average 32 points with a standard deviation of 337 points. A match between

Table 2. Summary statistics.

Variable	Mean	Standard deviation	Minimum value	Maximum value
FIFA points	933	240	526	1485
GDP per capita (2013 US dollars)	24,516	15,470	2711	53,705
Population in 000s	58,839	70,280	3286	318,349
Democracy index	669	210	100	922
FIFA points difference (FD)	32	337	−959	596
$YCAP_A/YCAP_B$	1.91	2.2	0.07	10.9
POP_A/POP_B	4.9	9.3	0.04	47.2
DEM_A/DEM_B	1.3	1.2	0.12	7.7
Goals difference	−0.14	1.88	−6	4

Australia and Spain had a point difference of −959 which was the maximum difference of any match played in the 2014 finals. This happened because the weakest team played the strongest team in FIFA's list. In other words, the strongest team of the 32 national teams played the weakest but both did not qualify to the next round. And if for Australia this was expected, for Spain it came as a surprise because not only was it the stronger team, but they were the 2010 World Cup winners. On average, per capita income was 1.9 times greater between any two teams of a game with a standard deviation of 2.2. In one match, team A had a per capita income 11 times higher than team B. This was the highest proportion of all matches. Population differences are even larger. On average, the opposing teams' population proportion was 4.9 with a standard deviation of 9.3 and a maximum value of 47. Smaller are the differences of the democratic index with a proportion of 1.3 and standard deviation of 1.2 and a maximum proportion of 7.7.

The goal difference is of great interest to this study. In Table 2, this variable is calculated as the goal difference of the teams as they appear in the FIFA schedule. For example, the match Brazil vs. Germany had a goal difference of −6. The match Germany vs. Portugal had a goal difference of 4, and the match Spain vs. the Netherlands had a −4 goal difference. Which teams go first determines the sign and not the absolute goal difference. FIFA's way of reporting match results is followed in this study. The probability distribution does not change if we change the order of team appearance in one match. The mean value of the goal difference was close to zero and a *t*-test rejects the hypothesis that is different from zero.

Table 3 presents the distribution of the absolute goal differences. Most of the matches (45.31%) had one goal difference and 20.31% ended in a draw. One-third of the games had a goal difference higher or equal to two goals. Four games had a goal difference of 3 goals and one 6 goals.

Which were the matches in the 2014 World Cup whose outcomes could be considered a (big) surprise? In statistical jargon, these realizations are called outliers? Mentioned above was the shocking goal difference in the Brazil–Germany semi-final game. The 6 goal difference (a final score of 7–1 for Germany) was a big upset for Brazil and becomes even bigger if it is taken into consideration that Brazil played at home. A surprise in football occurs when (a) a weak team beats a much stronger team and (b) when one team beats another team of equal strength with a very high score difference, i.e. more than three goals. Outliers can be graphically shown.

Table 3. Goal difference distribution

Value	Count	Per cent	Cumulative count	Cumulative per cent
0	13	20.31	13	20.31
1	29	45.31	42	65.63
2	10	15.63	52	81.25
3	8	12.50	60	93.75
4	3	4.69	63	98.44
6	1	1.56	64	100.00
Total	64	100.00	64	100.00

Figure 1 is a scatter diagram of goal differences and FIFA ranking differences. Included is an ellipsis with a 95% confidence interval. Points that lie outside the ellipse are defined as outliers. Three games lie outside, and they can be considered as outliers. First is the match between Brazil and Germany with -6 goal difference. As can be seen from the vertical grid lines, this point lies almost on the zero score of strength difference. It lies on the negative part of the FIFA strength difference indicating that Germany was ranked higher than Brazil, but if one takes into consideration the alleged home advantage that exists in football games, this strength difference is almost non-existent.

The second outlier is the match between Australia and Spain. Here, it is not the result that makes it an outlier because any score would have been an outlier, but the FIFA difference. The third outlier is the win of the Netherlands over Spain by 5 goals to 1. However, the two matches are very close to the ellipse and only the Brazil vs. Germany match is far away. In this study, only the latter is considered as an outlier for the regression analysis.

Outliers make football matches a spectacular game and it might explain why it is so popular, because great upsets, or as journalists would say, 'miracles' do happen. However, outliers create estimation problems. They violate the conditions necessary to apply the classical linear regression estimation. One of them is the normality assumption of the dependent variable and of the residuals. One ad hoc solution is to remove the outliers and refit the data to see whether the new line provides a better

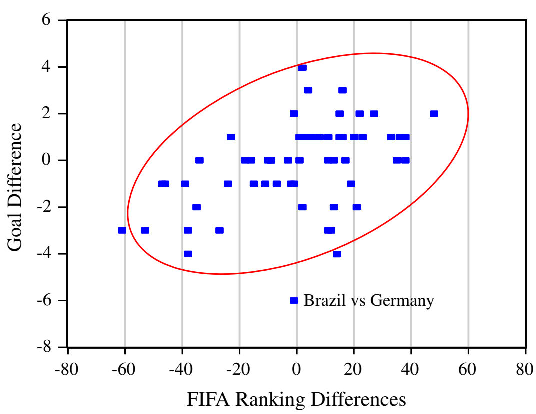

Figure 1. Goal and FIFA ranking points difference.

fit. This approach is followed here and the empirical results are reported with and without the outlier. A better fit will produce residuals which are normally distributed with zero mean and a minimum variance relative to any other possible estimator and model specification. However, the new fit will not necessarily provide a larger R^2 or a lower probability for the *F*-test.

Figure 2 is the same as the previous graph excluding the outlier and with a kernel fit. The graph shows that the relationship between goal differences and FIFA points difference is non-linear. It appears there are two turning points, so a polynomial of degree three might fit the data better.

Empirical findings

Based on the previous analysis of the determinants of the goal differences in the 64 matches played in the 2014 World Cup, the following empirical model is estimated:

$$\mathrm{GD_{AB}} = \mathrm{c_1} + \mathrm{c_2FD_{AB}} + \mathrm{c_3FD_{AB}^2} + \mathrm{c_4FD_{AB}^3} + \mathrm{c_5YCAP_A/YCAP_B} + \mathrm{c_6POP_A/POP_B} + \mathrm{c_7DEM_A/DEMP_B} + \mathrm{u}$$

The hypotheses made for the six parameters of the above equation are the following:

$$\mathrm{c_1} + \mathrm{c_5} + \mathrm{c_6} + \mathrm{c_7} = 0,\ \mathrm{c_2} > 0,\ \mathrm{c_3} < 0,\ \mathrm{c_4} = ?,\ \mathrm{c_5} > 0,\ \mathrm{c_6} > 0 \text{ and } \mathrm{c_7} > 0$$

The first parameter restriction states that when teams are of equal strength according to FIFA (FD = 0) and they come from countries with the same per capita incomes ($\mathrm{YCAP_A/YCAP_B} = 1$), equal populations ($\mathrm{POP_A/POP_B} = 1$), and the same level of democracy ($\mathrm{DEM_A/DEMP_B} = 1$), then the expected goal difference in their match should be zero. The $\mathrm{c_3}$ and $\mathrm{c_4}$ specifies a non-linear function which fits the data of the scatter diagram of goal and FIFA ranking differences. This implies that goal differences increase with an increase in FIFA ranking differences ($\mathrm{c_2} > 0$) but at a decreasing rate ($\mathrm{c_3} < 0$).

The estimation of the empirical equation requires that the distribution of goal differences should be sufficiently approximated by the normal distribution, otherwise

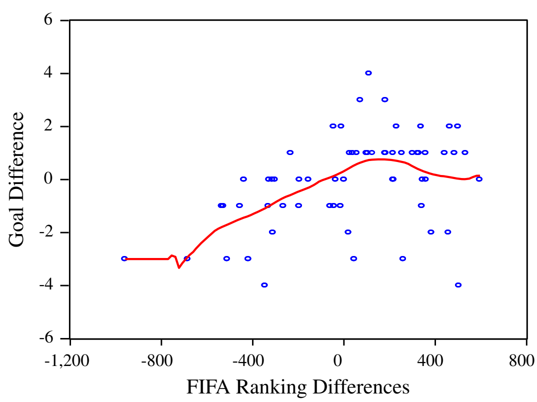

Figure 2. A kernel fit of goal and FIFA ranking points difference.

inferences can be ambiguous.[13] The Jarque–Bera test for normality rejects the hypothesis of normal distribution of the goal differences for the total sample. However, if the game between Brazil and Germany is excluded, then the Jarque–Bera test fails to reject the hypothesis of normal distribution.[14]

Table 4 reports the empirical results. Five specifications are presented. The first four exclude the game between Brazil and Germany. A Harvey test of heteroscedasticity rejects the null hypothesis of homoscedasticity, and the model is estimated with an heteroscedasticity correction. The first specification includes only a polynomial of FIFA measured strength differences between the two opposing teams of each match. All three terms of the polynomial are statistically significant. Goal disparity increases with FIFA measured strength difference, but at a decreasing rate. Strength difference can explain 24% of goal differences in the 63 games of the 2014 World Cup. However, in this model specification, the residuals of the regression are not normally distributed. The Jarque–Bera test rejects the hypothesis of normality, and the sum of squared residuals is not minimized. Specification (2) adds the economic, population and political variables. The latter was found to have a non-linear impact on goal difference. Per capita income and population proportions have a positive linear impact on goal differences.

Specifications (3) and (4) include dummy variables. Only the continental dummy variable is statistically significant. Teams from the Latin America continent had a ‘home’ advantage. This might be due to the climate, but of equal importance was the bigger crowds of supporters which could travel to Brazil because of the short distances. The surprising result is that of Brazil. It appears that there was no home advantage for Brazil. Even though the parameter is not statistically significant, it is important to note that its sign is negative. For some teams which are considered as the favourite to win the Cup because they play at home, the pressure may have the opposite effect, i.e. the home effect becomes a disadvantage. Also, non-significant was the round of the match. The negative sign shows that matches of the knockout rounds have smaller goal differences, something to be expected in these types of games where every team wants to play very good defence.

Specification (5) includes the full sample of the 64 games. The residuals of the regression are not normally distributed. With the exception of the dummy variable, the estimated parameters are not affected.

A Wald test of the parameter restriction shows that countries with similar football strength, per capita income, population and level of democracy will result in a tie.

The impact of the per capita income is statistically significant. An increase in the per capita income ratio increases the goal differences. Countries with 5 times higher per capita income than their opponents will have a goal difference of one goal relative to match results between teams with similar per capita income. Population ratios are important as well. Countries with a population double than the rival team’s national population will result in a goal difference higher by one goal. The impact of the democracy index is positive, but it decreases as the discrepancy between the two countries increases.

Even though the prediction of the goal scores was not the prime motivation for this study, it is of interest to see what the model would have predicted for each of the 63 matches played in the 2014 World Cup finals since one match was an outlier and it has been excluded from the regression analysis. Table 5 presents the actual, the fitted and the residual of the 63 matches using the model specification (3) of Table 4. There are two types of prediction information that can be extracted from

Table 4. Empirical results.

	(1)	(2)	(3)	(4)	(5)
Constant	0.404 (0.1484)	−0.843 (0.1196)	−0.820 (0.1228)	−0.615990 (0.3723)	−1.135468 (0.1273)
FD_{AB}	0.002908 (0.0001)	0.002183 (0.0005)	0.002 (0.0005)	0.002276 (0.0000)	0.002118 (0.0005)
FD^2_{AB}	−5.23E−06 (0.0126)	−6.05E−06 (0.0010)	−6.61E−06 (0.0001)	−6.68E−06 (0.0011)	−5.68E−06 (0.0030)
FD^3_{AB}	−4.65E−09 (0.0810)	−5.02E−09 (0.0162)	−5.46E−09 (0.0055)	−5.94E−09 (0.0077)	−5.00E−09 (0.0268)
$YCAP_A/YCAP_B$		0.141 (0.03)	0.207 (0.01)	0.190 (0.01)	0.188 (0.02)
POP_A/POP_B		0.030 (0.001)	0.027 (0.001)	0.030 (0.002)	0.033 (0.0003)
$DEM_A/DEMP_B$		0.98 (0.01)	0.95 (0.04)	0.89 (0.05)	1.05 (0.05)
$DEM_A/DEMP^2_B$		−0.117235 (0.0115)	−0.123 (0.0143)	−0.116652 (0.0274)	−0.129 (0.0253)
Continent			0.500 (0.0786)	0.491759 (0.0900)	0.220431 (0.6370)
Brazil				−0.539026 (0.2492)	
Round				−0.060085 (0.6829)	
Sample size	63	63	63	63	64
R-squared	0.277013	0.363014	0.391309	0.400735	0.317619
Adjusted *R*-squared	0.240251	0.281943	0.301133	0.285492	0.218364
F-statistic	7.535310 (0.0002)	4.477739 (0.0005)	4.339379 (0.0004)	3.477300 (0.0015)	3.200022 (0.0047)
Sum squared residual	136	120	115	113	153
Jarque–Bera normality test for residuals	5.6 (0.061)	1.55 (0.462)	2.81 (0.246)	2.18 (0.337)	23.93 (0.000)

Notes: *p*-values in brackets. All models have been estimated using HAC standard errors and covariance (Bartlett kernel, Newey–West automatic bandwidth = 19.8022, NW automatic lag length = 3).

Table 5. Actual and predicted goal score differences.

Match	Actual Goal Difference	Predicted Goal Difference	Under (+) and Over (-) prediction	Predicting the Winner	Predicting Goal Difference
1	2	1.65	0.35	1	1
2	1	2.27	−1.27	1	1
3	−4	−1.20	−2.80	1	0
4	2	−0.83	2.83	0	0
5	3	0.73	2.27	1	0
6	−2	−0.19	−1.81	1	0
7	−1	0.34	−1.34	0	1
8	1	−0.57	1.57	0	0
9	1	0.57	0.43	1	1
10	3	1.72	1.28	1	1
11	1	1.29	−0.29	1	1
12	4	0.73	3.27	1	0
13	0	0.21	−0.21	1	1
14	−1	−1.36	0.36	1	1
15	1	1.33	−0.33	1	1
16	0	−0.12	0.12	1	1
17	0	−0.50	0.50	1	1
18	−1	−1.51	0.51	1	1
19	−2	−1.07	−0.93	1	1
20	−4	−1.49	−2.51	1	0
21	1	1.22	−0.22	1	1
22	1	0.68	0.32	1	1
23	0	−1.06	1.06	0	1
24	−1	0.03	−1.03	1	1
25	−3	0.49	−3.49	0	0
26	−1	−0.03	−0.97	1	1
27	1	0.11	0.89	1	1
28	0	0.29	−0.29	1	1
29	1	0.43	0.57	1	1
30	1	1.11	−0.11	1	1
31	−2	0.08	−2.08	0	0
32	0	0.81	−0.81	1	1
33	2	−0.03	2.03	0	0
34	−3	−2.81	−0.19	1	1
35	−3	−3.08	0.08	1	1
36	−2	−0.19	−1.81	1	0
37	−1	0.22	−1.22	0	1
38	0	−1.12	1.12	0	1
39	−3	−1.71	−1.29	1	1
40	1	1.00	0.00	1	1
41	−1	−2.80	1.80	1	0
42	2	0.98	1.02	1	1
43	−3	−1.32	−1.68	1	0
44	0	−0.74	0.74	0	1
45	1	0.37	0.63	1	1
46	−1	−0.56	−0.44	1	1
47	−1	−1.80	0.80	1	1
48	0	0.06	−0.06	1	1
49	0	−0.01	0.01	1	1
50	2	0.36	1.64	1	0

(*Continued*)

Table 5. (*Continued*).

Match	Actual Goal Difference	Predicted Goal Difference	Under (+) and Over (-) prediction	Predicting the Winner	Predicting Goal Difference
51	2	0.90	1.10	1	1
52	1	0.52	0.48	1	1
53	1	0.38	0.62	1	1
54	0	−0.90	0.90	0	1
55	1	0.58	0.42	1	1
56	1	0.22	0.78	1	1
57	1	0.55	0.45	1	1
58	−1	−0.45	−0.55	1	1
59	0	0.89	−0.89	1	1
60	1	0.72	0.28	1	1
61	−6	NA	NA	NA	NA
62	0	−0.40	0.40	1	1
63	−3	0.69	−3.69	0	0
64	1	0.32	0.68	1	1

Table 5. The first prediction relates to the winner of each of the 63 matches. If the actual and the predicted numbers of Table 5 have the same sign, the model has accurately predicted the winner of the match albeit not always with the correct score difference. Since zero has no sign, a residual of less than 0.5 goals is considered a correct prediction of the winner. In 51 matches out of the 63 of the sample, the result was according to what the model would have predicted. The second prediction relates to goal difference. If the model predicts the score difference with no more than 1.5 goals, then it is assumed to be a good prediction. In this case, 48 scores out of the 63 matches were correctly predicted by the model.

Conclusions

The outcome of football matches between national teams in the World Cup finals depends not only on the relative quality and the strength of the two rivals on the day of the match but on economic, demographic and political considerations of their respective countries. In this study, it was found that these variables have significant impacts on the goal difference of a football match. The 64 games of the 2014 World Cup held in Brazil were used to investigate with an empirical model the quantitative effect of these explanatory variables. FIFA rankings have a strong non-linear positive effect on match outcomes. It is found that both the relative per capita income and the relative population of the two rival teams have a positive impact on the goal difference of the match. Similarly, more democratic countries have a stronger national football team. The impact is positive but decreases as the difference in the democratic index increases. The results of the study confirm the empirical findings of other studies.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes

1. See Conchas, 'Research Possibilities'.
2. See Dyte and Clarke, 'A Ratings Based Poisson Model'; Suzuki et al., 'A Bayesian Approach for Predicting Match Outcomes'; Leitner et al., 'Forecasting Sports Tournaments by Ratings of (Prob)abilities'; and Leitner et al., 'Bookmaker Consensus and Agreement' for an example of such studies.
3. For example, using this approach, Zeileis et al., 'Home Victory for Brazil' foretold that Brazil, Argentina and Germany will win the World Cup with 22.5, 15.8 and 13.4% probabilities, respectively.
4. Goldman-Sachs Global Investment Research, *The World Cup and Economics 2014* have developed a regression model to predict match results and most importantly the winner of the 2014 World Cup. Similarly, Groll et al., *Who will Win the Trophy?*, have developed a regression model to predict the winner of the 2014 World Cup.
5. FIFA provides a ranking of countries using a weighted average of the last four years taking into consideration (a) the result of the match, (b) the importance of the match, (c) the strength of the opponent and (d) the strength of the confederation. FIFA reports rankings since 1993. June 2014 was the last FIFA ranking report before the World Cup in Brazil. For example, Spain with 1485 points was in first position just before the World Cup begun. The most recent ranking report is available on http://www.fifa.com/worldranking/rankingtable/index.html. See also FIFA's method on http://www.fifa.com/worldranking/procedureandschedule/menprocedure/index.html. In general, it is very difficult to make predictions in football and particularly games that are played in World Cup finals. For example, Ramírez and Cardona, 'Which Team will Win the 2014 FIFA World Cup?' developed a Bayesian approach to predict the teams which will qualify for the next rounds. From the 8 matches of the round of 16, their model predicted only one (12.5%) match that of Argentina vs. Switzerland. Furthermore, from their 16 teams, 6 did not make it to the actual round of 16 (37.5%).
6. For example, see Hoffman et al., 'The Socioeconomic Determinants'; Houston and Wilson, 'Income, Leisure and Proficiency'; and Macmillan-Smith, 'Explaining International Soccer Rankings'.
7. See Torgler, 'The Economics of the Football FIFA World Cup'.
8. See Karlis and Ntzoufras, 'Bayesian Modelling of Football Outcomes'.
9. To be fair with the bookmakers and statisticians, the Brazil–Germany match result may not have been the same if two of the best Brazilian players were able to play. For example, in the Goldman-Sachs Global Investment Research, *The World Cup and Economics 2014* is stated that '… if a key player who was responsible for a team's recent successes is injured, this will have no bearing on our predictions'. Neymar, a world class midfielder, and Brazil's captain, Thiago Silva, a world class defence player, could not play in this particular game.
10. See Szymanski and Smith, 'The English Football Industry'.
11. For an example of studies that look at this relationship with explicit reference to Brazil see Couto, 'Football, Control and Resistance' and de Melo and Drumond, 'Globo, the Brazilian Military Dictatorship'.
12. Alternative to FIFA rankings are the Elo rankings and the bookmakers rankings. All these rankings are highly correlated. Interestingly, the Elo ranking was favoured in Brazil for the 2014 World Cup Games as this was admitted by the Goldman-Sachs Global Investment Research, *The World Cup and Economics 2014* study. This is because the Elo ranking gives more emphasis on the most recent game results relative to FIFA rankings.
13. Karlis and Ntzoufras, 'Bayesian Modelling of Football Outcomes', 135.
14. More rigorous tests of empirical distribution tests such as the Anderson–Darling test reject the hypothesis of normal distribution of goal differences.

References

Conchas, M. 'Research Possibilities for the 2014 FIFA World Cup in Brazil'. *Soccer & Society* 15, no. 1 (2014): 167–74.

Couto, E. 'Football, Control and Resistance in the Brazilian Military Dictatorship in the 1970s'. *The International Journal of the History of Sport* 31, no. 10 (2014): 1267–77.

Dyte, D., and S.R. Clarke. 'A Ratings Based Poisson Model for World Cup Soccer Simulation'. *Journal of the Operational Research Society* 51, no. 8 (2000): 993–8.

Goldman-Sachs Global Investment Research. *The World Cup and Economics 2014: A Statistical Model of the 2014 World Cup*, 2014. http://www.goldmansachs.com/our-thinking/outlook/world-cup-sections/world-cup-book-2014-statistical-model.html.

Groll, A., G. Schauberger, and G. Tutz. 'Brazil or Germany – Who Will Win the Trophy? Prediction of the FIFA World Cup 2014 Based on Team-specific Regularized Poisson Regression'. Department of Statistics. University of Munich. Technical Report, no. 166, 2014.

Hoffman, R., V.G. Lee, and B. Ramasamy. 'The Socioeconomic Determinants of International Soccer Performance'. *Journal of Applied Economics* 5, no. 2 (2002): 253–72.

Houston, R.G., and D.P. Wilson. 'Income, Leisure and Proficiency: An Economic Study of Football Performance'. *Applied Economics Letters* 9, no. 14 (2002): 939–43.

Karlis, D., and I. Ntzoufras. 'Bayesian Modelling of Football Outcomes: Using the Skellam's Distribution for the Goal Difference'. *IMA Journal of Management Mathematics* 20 (2009): 133–45.

Leitner, C., A. Zeileis, and K. Hornik. 'Forecasting Sports Tournaments by Ratings of (Prob)abilities: A comparison for the EURO 2008'. *International Journal of Forecasting* 26, no. 3 (2010): 471–81.

Leitner, C., A. Zeileis, and K. Hornik. 'Bookmaker Consensus and Agreement for the UEFA Champions League 2008/2009'. *IMA Journal of Management Mathematics* 22, no. 2 (2011): 183–94.

Macmillan, P., and I. Smith. 'Explaining International Soccer Rankings'. *Journal of Sports Economics* 8, no. 2 (2007): 202–13.

de Melo, V.A., and M. Drumond. 'Globo, the Brazilian Military Dictatorship and the 1970 FIFA Football World Cup: Ambiguous Relations'. *Television & New Media* 15, no. 8 (2014): 703–10.

Ramírez, A., and J. Cardona. 'Which Team Will Win the 2014 FIFA World Cup? A Bayesian Approach for Dummies'. Working Papers of the Universidad EAFIT, No. 14-03; 2014.

Suzuki, A., L. Salasar, J. Leite, and F. Louzada-Neto. 'A Bayesian Approach for Predicting Match Outcomes: The 2006 (Association) Football World Cup'. *Journal of the Operational Research Society* 61, no. 10 (2010): 1530–9.

Szymanski, S., and R. Smith 'The English Football Industry: Profit, Performance and Industrial Structure'. *International Review of Applied Economics* 11, no. 1 (1997): 135–53.

Torgler, B. 'The Economics of the FIFA Football Worldcup'. *Kyklos* 57, no. 2 (2004): 287–300.

Yamamura, E. 'Technology Transfer and Convergence of Performance: An Economic Study of FIFA Football Ranking'. *Applied Economics Letters* 16, no. 3 (2009): 261–6.

Zeileis, A., C. Leitner, and K.Hornik. 'Home Victory for Brazil in the 2014 FIFA World Cup'. Working paper, Faculty of Economics and Statistics, University of Innsbruck, 2014.

Appendix 1.

Table A1. Summary statistics of the 64 matches played in the 2014 World Cup.

M	Team A	Team B	GD	FD_{AB}	$YCAP_A$/ $YCAP_B$	POP_A/ POP_B	DEM_A/ DEM_B
1	Brazil	Croatia	2	339	0.72	47.27	1.03
2	Mexico	Cameroon	1	324	6.07	5.87	2.01
3	Spain	The Netherlands	−4	504	0.74	2.76	0.89
4	Chile	Australia	2	500	0.50	0.75	0.82
5	Colombia	Greece	3	73	0.48	4.29	0.87
6	Uruguay	Costa Rica	−2	385	1.41	0.70	1.01
7	England	Italy	−1	−14	1.47	0.87	1.05
8	Côte d’Ivoire	Japan	1	183	0.08	0.18	0.12
9	Switzerland	Ecuador	1	358	5.13	0.52	1.57
10	France	Honduras	3	182	8.04	7.71	1.35
11	Argentina	Bosnia and Herzegovina	1	302	1.95	11.25	1.34
12	Germany	Portugal	4	111	1.67	7.70	1.05
13	Iran	Nigeria	0	1	2.66	0.43	0.53
14	Ghana	USA	−1	−331	0.07	0.08	0.74
15	Belgium	Algeria	1	216	3.03	0.29	2.10
16	Brazil	Mexico	0	360	0.91	1.69	1.03
17	Russia	Korea South	0	346	0.73	2.90	0.46
18	Australia	The Netherlands	−1	−455	1.00	1.40	1.03
19	Spain	Chile	−2	459	1.47	2.62	1.06
20	Cameroon	Croatia	−4	−345	0.13	4.75	0.50
21	Colombia	Côte d’Ivoire	1	328	4.11	2.05	6.63
22	Uruguay	Αγγλία	1	57	0.39	0.06	1.01
23	Japan	Greece	0	−438	1.42	11.43	1.06
24	Italy	Costa Rica	−1	342	2.47	13.02	0.96
25	Switzerland	France	−3	45	1.46	0.12	1.15
26	Honduras	Ecuador	−1	−60	0.44	0.54	1.01
27	Argentina	Iran	1	534	1.20	0.55	3.45
28	Germany	Ghana	0	596	10.90	2.98	1.39
29	Nigeria	Bosnia and Herzegovina	1	−233	0.61	47.08	0.74
30	Belgium	Russia	1	181	1.67	0.08	2.15
31	Korea South	Algeria	−2	−311	2.49	1.30	2.12
32	USA	Portugal	0	−154	2.05	30.38	1.02
33	The Netherlands	Chile	2	−45	1.98	0.95	1.19
34	Australia	Spain	−3	−959	1.36	0.51	1.15
35	Cameroon	Brazil	−3	−684	0.18	0.10	0.48
36	Croatia	Mexico	−2	21	1.27	0.17	1.00
37	Italy	Uruguay	−1	−43	1.75	18.50	0.95
38	Costa Rica	England	0	−328	0.27	0.09	1.00
39	Japan	Colombia	−3	−511	2.92	2.67	1.22
40	Greece	Côte d’Ivoire	1	255	8.52	0.48	7.65
41	Nigeria	Argentina	−1	−535	0.31	4.18	0.55
42	Bosnia and Herzegovina	Iran	2	232	0.62	0.05	2.58
43	Honduras	Switzerland	−3	−418	0.09	1.05	0.64
44	Ecuador	France	0	−313	0.28	0.24	0.73
45	Portugal	Ghana	1	485	6.52	0.39	1.32

(*Continued*)

Table A1. (*Continued*).

M	Team A	Team B	GD	FD_{AB}	$YCAP_A$/ $YCAP_B$	POP_A/ POP_B	DEM_A/ DEM_B
46	USA	Germany	−1	−265	1.23	3.94	0.97
47	Korea South	Belgium	−1	−527	0.82	4.50	1.01
48	Algeria	Russia	0	−35	0.55	0.26	1.02
49	Brazil	Chile	0	216	0.69	11.41	0.94
50	Colombia	Uruguay	2	−10	0.63	14.51	0.81
51	France	Nigeria	2	464	6.29	0.37	2.09
52	Germany	Algeria	1	442	3.26	2.09	2.18
53	The Netherlands	Mexico	1	99	2.64	0.14	1.30
54	Costa Rica	Greece	0	−302	0.54	0.42	1.06
55	Argentina	Switzerland	1	26	0.35	5.23	0.75
56	Belgium	USA	1	39	0.76	0.04	0.99
57	Brazil	Colombia	1	105	1.22	4.25	1.07
58	France	Germany	−1	−196	0.85	0.82	0.94
59	The Netherlands	Costa Rica	0	219	3.13	3.61	1.11
60	Argentina	Belgium	1	101	0.46	3.81	0.85
61	Brazil	Germany	−6	−58	0.35	2.51	0.85
62	The Netherlands	Argentina	0	−194	2.32	0.40	1.31
63	Brazil	The Netherlands	−3	261	0.35	10.82	0.79
64	Germany	Argentina	1	125	2.31	1.89	1.22