

*The PwC World Cup
Index: what can the
dismal science tell
us about the
beautiful game?*

June 2014

Contents

<i>Executive summary</i>	1
Key features of our analysis	2
Money can't buy World Cup success	2
Quantity of players and footballing interest is important to success	4
Tradition matters... but so does recent form	5
Home advantage bodes well for Brazil... and for the rest of South America	5
Are England perennial underachievers?	6
Who is in the 'Group of Death'?	7
Annex: Technical details of regression models	9
Contacts and services	11

Executive summary

As the 2014 World Cup in Brazil draws closer, there is increasing interest in what makes a successful World Cup nation. This paper (which builds on an earlier one in 2010) contributes to this debate by using econometric analysis to test the correlation between several variables and World Cup performance. We aim to answer the following questions:

- Which characteristics explain historic World Cup performance?
- Which countries have underachieved and overachieved in past World Cups?
- Who are the strongest countries in the 2014 World Cup, and which group is the ‘Group of Death’?

Home advantage bodes well for Brazil...and for the rest of South America

Our modelling shows that, on average, host countries can expect to progress a further two rounds than if they were not hosting. There is also a clear ‘home region’ effect, which may reflect stronger home support and familiar climatic conditions.

Money can’t buy World Cup success

In our previous analysis of the Olympics, we found a strong correlation between medal totals and the size of the economy. However, no such relationship was found for the World Cup, with GDP per capita and population failing to explain how well a country performs after controlling for other factors.

Instead, we found a number of football-specific variables which explain World Cup performance better, such as the number of registered football players in each country and average football attendance levels. The other factors that we found to be important were a long-term footballing tradition and recent form as reflected in the FIFA world rankings.

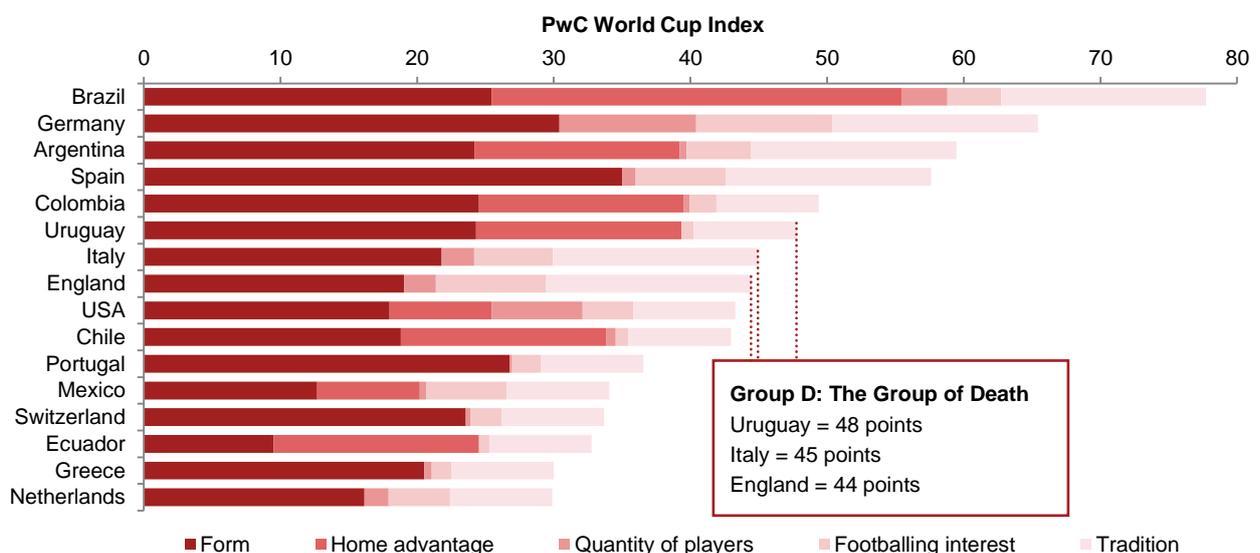
Brazil are the favourites, England face a struggle in the Group of Death

After identifying the extent to which different variables explain World Cup performance, we created a composite indicator which measures the relative strength of each competing country across a weighted average of these characteristics. This produces a ‘PwC World Cup Index’ that provides a novel way of assessing each country’s relative prospects in the 2014 World Cup.

Our indicator suggests that Brazil may be the favourite this summer, due to their combination of footballing tradition and home advantage. But Germany, Argentina and Spain also have good chances.

Meanwhile, England rank eighth on our index, which on the face of it might suggest that the quarter-finals should be attainable. But they are in for a tough battle to progress from the group stages given the slightly superior strength of their rivals Italy and Uruguay. Our analysis suggests this is the ‘Group of Death’, with the highest combined index score of any of the eight preliminary groups.

Figure 1: Estimated relative strength of leading teams in the 2014 World Cup



What can the dismal science tell us about the beautiful game?

The 19th century historian and essayist, Thomas Carlyle, dubbed economics the 'dismal science' in response to the gloomy writings of Malthus, but it can have its lighter side. In this paper, we apply standard econometric techniques to the 'beautiful game' of football to illustrate how it can help to assess prospects for the 2014 World Cup, which starts in Brazil later this month. This builds on our earlier analyses of the 2010 World Cup¹, as well as several past Summer and Winter Olympic Games.

We aim to answer the following questions:

- Which factors explain the historic performance of countries in the World Cup?
- Which countries have historically under- and over-performed in the World Cup?
- Who are the strongest countries based on a new PwC World Cup Index for 2014, and which group is the 'Group of Death'?

Key features of our analysis

We use historic data on the 56 countries who have played at least 6 World Cup finals matches. We measure historical World Cup performance using FIFA's all-time rankings table, which awards three points for a win, one point for a draw and zero points for a loss for every World Cup finals game played².

We find that the following factors are important in explaining past World Cup performance:

- The number of football players available
- The average attendance of top division football matches
- How many times the country has bid to host the World Cup
- Whether the country is from Europe or South America

- Whether the country is the home nation or from the same continent as the home nation
- Recent form, as measured in our historical analysis by how the country performed in the last two World Cups, and in our 2014 assessment by the latest FIFA world rankings.

As described further below, these factors were then combined in a new PwC World Cup Index as an indicator of relative prospects in the 2014 World Cup.

Meanwhile, as in our 2010 analysis, we found that economic variables such as GDP per capita and population were statistically insignificant after controlling for football-specific factors. Further details of our technical modelling results can be found in the Annex.

Money can't buy World Cup success

In the case of the Olympics, we found a statistically significant positive relationship between the number of medals won and average incomes, as measured by GDP per capita. But no such relationship is evident in international football, as shown in Figure 2 below.

Although there is a slight upward slope to the trend line it is statistically insignificant, and GDP per capita explains less than 5% of the variation in World Cup performance across countries.

It seems, therefore, that richer countries do not generally outperform poorer countries at football. This may reflect the fact that football is relatively inexpensive to practice compared to some Olympic sports – footballing skills can as easily be honed in the back streets as in an expensive sports centre.

¹ J. Hawksworth, 'What can econometrics tell us about World Cup performance?', PwC, May 2010.

² FIFA World Cup all-time rankings, <http://www.fifa.com/worldfootball/statisticsandrecords/tournaments/worldcup/alltimerankings.html>

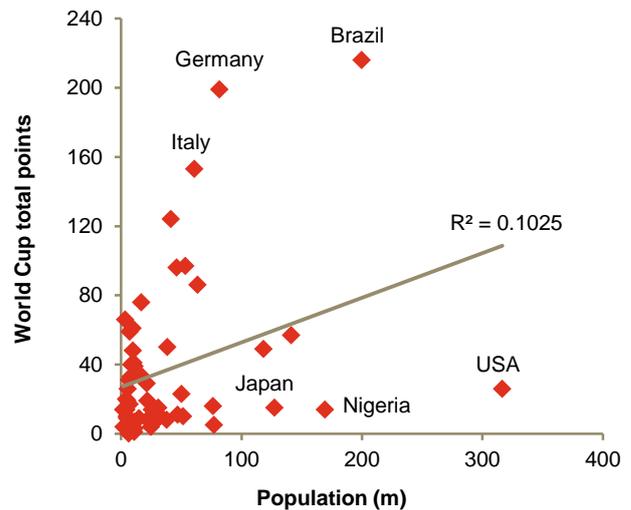
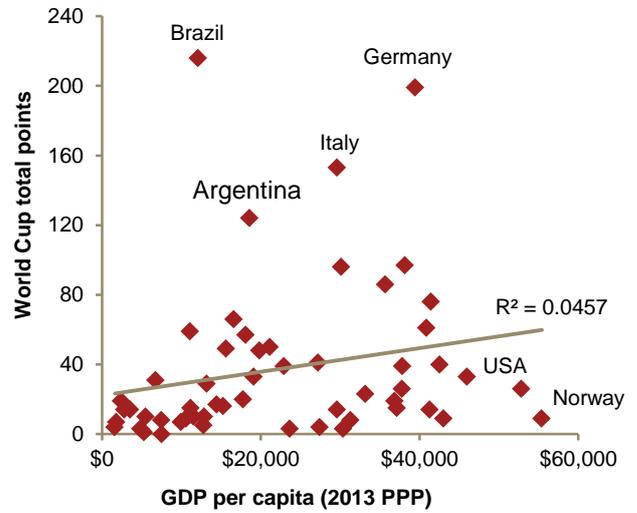
In fact, when we add GDP per capita to our final specification its coefficient is negative – so it appears that poorer countries perform better, holding all other variables constant. Previous studies have claimed that this effect arises because as countries become wealthier they provide a greater range of sporting and other leisure activities to choose from, which means that individuals are less likely to choose football as their preferred activity³.

The correlation between population and World Cup success is stronger, with a statistically significant and positive relationship. However, Figure 3 shows there are a number of exceptions with the US in particular failing to capitalise on its larger population, while Uruguay ranks ninth in the all-time World Cup table despite having a population of only 3 million. In addition, the vast populations of China and India have not helped their World Cup success, and they were excluded from our analysis having only played a total of three tournament matches between them.

Although there is a significant correlation between World Cup performance and population, this relationship breaks down when we control for the number of football players available to each country, which provides a much better fit in our model. This means that having a large population does little to enhance footballing performance if it does not produce a larger pool of players from which to choose.

The conclusion that economic variables do not significantly explain World Cup performance was also reached in our 2010 analysis, but in this new paper we extend our analysis to include a number of football-specific variables that greatly enhance the explanatory power of the model.

Figures 2 and 3: The relationship between historic World Cup performance, average income level and population



³ Berlinschi et al, When Drains and Gains Coincide, 2010

Quantity of players and footballing interest is important to success

Our analysis confirms the intuitive finding that countries with a greater pool of football players to choose from perform better at the World Cup. This was a strong statistical relationship, with the number of players explaining around 50% of the variation in countries' World Cup performance, as compared to only 10% for total population.

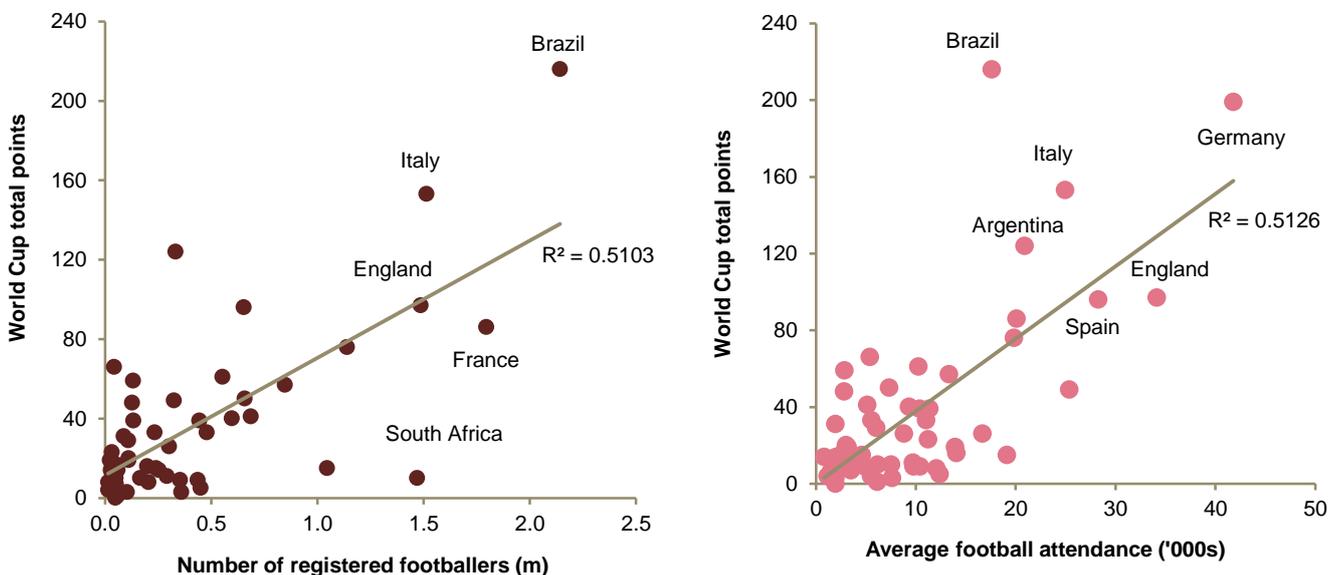
Using the logarithm of registered players provides a similar level of fit, which indicates that there aren't any significant diminishing returns to scale from the number of football players. This means that Germany reaps the full benefits of its 6 million-strong pool of registered football players, the largest by some distance. The next largest is the US with over 4 million, although this figure includes a much larger proportion of female players than other countries, so its relationship with male success in the World Cup is likely to be less strong. Both of these data points are excluded from Figure 4 to demonstrate the correlation for the other countries. England lies on the trend line, with its historic World Cup performance in line with expectations given its number of registered players.

It is sometimes claimed that a joint UK or GB team (as in the Olympics) might have had more success in the World Cup than the home nations competing individually. If Wales, Scotland and Northern Ireland were added to the England team, this would increase the quantity of players available by around 200,000, which according to our model would have earned England an additional 2 points, equivalent to less than one extra win over the last 19 World Cups.

We found that a country's interest in football was as important as the quantity of players available, as shown in Figure 5. We used the average attendance at club football matches in the top-tier division of each country as a proxy for footballing interest.

Germany, England and Spain have the highest club football attendances respectively, which is not surprising given their reputation as Europe's most prestigious leagues. Brazil and Argentina have a better World Cup record than their club football attendances suggest, which is driven by the high level of exports of their footballers to European leagues, decreasing the quality of their domestic leagues.

Figures 4 and 5: The relationship between historical World Cup performance, number of registered football players and average football attendance



Tradition matters... but so does recent form

We found that the footballing 'tradition' of a country is important in explaining historic World Cup success, in the same way that a winter sports tradition allowed certain countries to consistently over-perform in the Winter Olympics.

Having played competitive football for the longest time, the two continents with the greatest footballing tradition are Europe and South America. Including dummy variables for each of these two continents (a variable with a value of 1 if the country is from the continent and a value of 0 otherwise) has a positive and statistically significant effect on World Cup performance. The coefficients on these variables suggest that, after accounting for our other factors such as the number of registered players and average attendance, countries from South America and Europe on average have collected an additional 40 and 29 points respectively compared to countries from other continents.

Also, countries with a tradition and enthusiasm for football are more likely to bid to host the World Cup, so we also included an indicator for countries which have bid to host the World Cup at least three times. Instead of these variables, we could have proxied for tradition by using the number of years each national FA has been established, or the

number of years since their first World Cup appearance, but these alternatives fit slightly less well in our model and so were not used.

While long-term tradition matters, recent form is also very important. We used a time series analysis to analyse performance in individual World Cup tournaments, and found that performance in the last two World Cups was significant in explaining current performance.

For a wider measure of current form FIFA ranking points could be used instead, which take into account performance in all international matches, and also take into account the difficulty of the opposition. We could not include FIFA ranking points in our time series model as they were only introduced in 1992 and have undergone numerous methodological changes which means that they are not easily comparable over time.

However, we modelled the relationship between FIFA rankings points and World Cup performances for the year 2010 only, which revealed a strong correlation between the two. We found that an extra 50 FIFA ranking points is associated with a country progressing one further round in the World Cup, which is a strong impact given that the current difference in ranking points between the top two ranked teams (Spain and Germany) is 120.

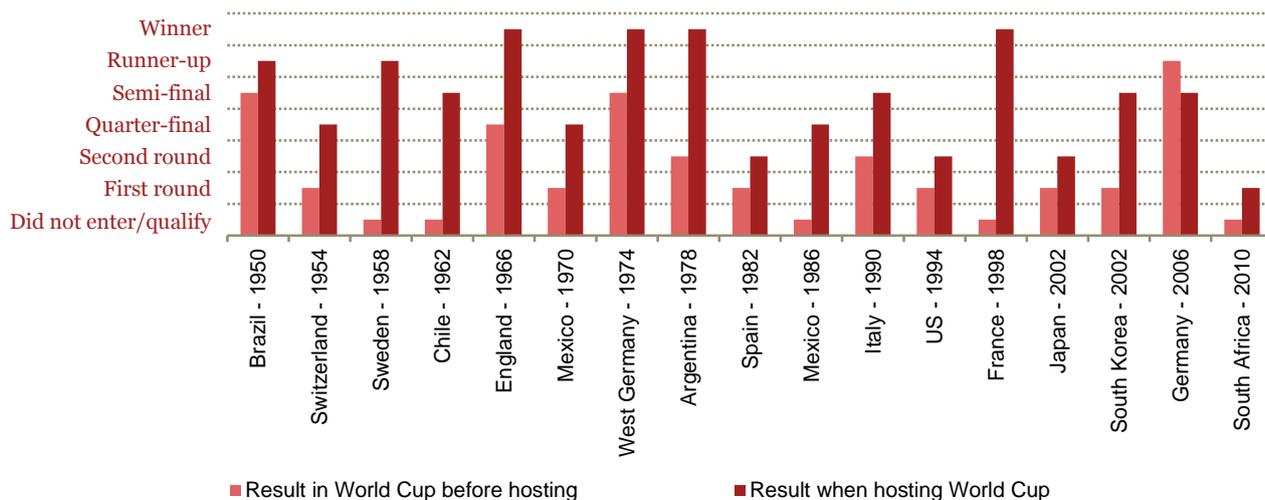
Home advantage bodes well for Brazil... and for the rest of South America

We found in previous modelling work that host countries tended to significantly outperform their prior or subsequent medal totals at the Olympics, and this trend thankfully continued at London 2012. The footballing evidence for a host country advantage appears just as compelling, as shown in Figure 6 below. Since 1950, host countries have produced a better performance when hosting the World Cup compared to their previous tournament in every year apart from 2006. In that year, Germany reached the semi-finals as host, compared to finishing runner-up in 2002.

Our time series analysis shows that home country advantage is equivalent to progressing an additional two rounds on average, and this pattern is evident in Figure 6. We also found a smaller, but statistically significant, impact of 'home region' advantage, where countries perform better in their own continent.

This home region effect may reflect a mix of stronger home crowd support and familiar climatic conditions. A European country has never won a World Cup held in the Americas, while only once has a South American country won the World Cup in Europe (Brazil in 1958). The spoils of the two World Cups held outside of Europe and the Americas (South Korea and Japan in 2002 and South Africa in 2010) have been shared equally between Europe and South America.

Figure 6: Host country performance at the World Cup



Are England perennial underachievers?

We have therefore established several variables which explain, to some extent, the differences in historic World Cup performance across countries. These are the number of registered football players, attendance at top division club matches, number of bids to host the World Cup, and whether the country is from Europe or South America.

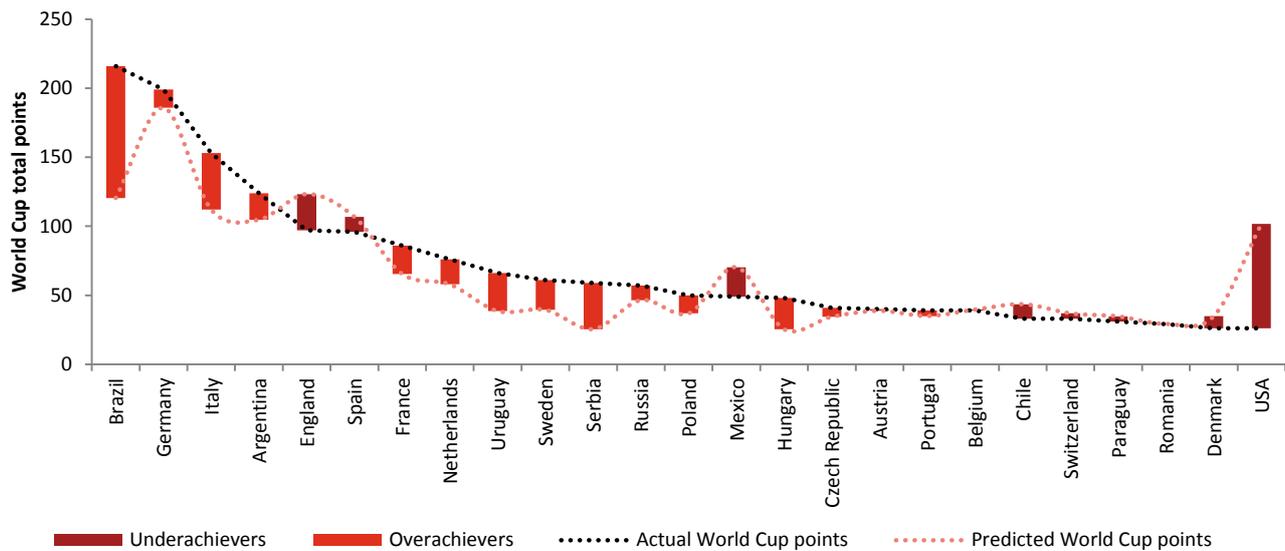
Using a regression model, we can produce estimates of how each country would be expected to perform given their endowments of these different variables. The model is discussed in more detail in the Annex.

Our estimates of each country’s total World Cup points compared to their actual points is shown in Figure 7 below, which reveals whether each country has under or overachieved relative to model estimates. Brazil is clearly the most considerable overachiever, collecting an additional 95 World Cup points compared to our model estimate, which is equivalent to an additional 31 wins. They have

consistently had greater success than other countries with a similar pool of registered players and level of club attendances.

England’s reputation among some as perennial underachievers is somewhat justified in our model. They have collected 26 fewer World Cup points than estimated, which is equivalent to an additional eight wins and two draws over the course of 19 tournaments. The largest underachievers are in fact the USA, who should have collected an additional 76 points according to our model, given their abundance of registered football players and that they often bid to host the tournament. However, it should be noted that this estimate does not take into account the fact that a larger proportion of the USA’s total registered players are female than other countries, which may mean that their underachievement should not be quite as large. The USA do have a much better record in the Women’s World Cup, which is consistent with the greater pool of registered female players there.

Figure 7: Model estimates vs actual historic World Cup performance



Who is in the ‘Group of Death’?

When the World Cup draw is made, one area of fervent debate is which group is the most difficult, often referred to as the ‘Group of Death’. It appears that opinion is divided on which is the Group of Death for the 2014 World Cup, and there have already been attempts to quantify the strength of each World Cup group⁴. Using the information we have gained in our econometric analysis, we add to this debate by identifying our own ‘Group of Death’.

We have already established a number of variables which correlate with historic World Cup success, so we have combined these with a measure of current form (using FIFA rankings) to gauge the strength of each team. To create a single indicator for team strength, which we refer to as the ‘PwC World Cup Index’, we normalised each variable and applied the following weightings:

- Current form – proxied by FIFA ranking, weighting = 35%
- Home advantage – 2 for the host, 1 for South America, 0.5 for Central America and 0 for all other countries, weighting = 30%
- Quantity of players – proxied by number of registered football players, weighting = 10%
- Footballing interest – proxied by club football attendances, weighting = 10%

- Tradition – 1 if the country is from South America/Europe and an extra 1 if it has bid to host the World Cup at least three times, weighting = 15%.

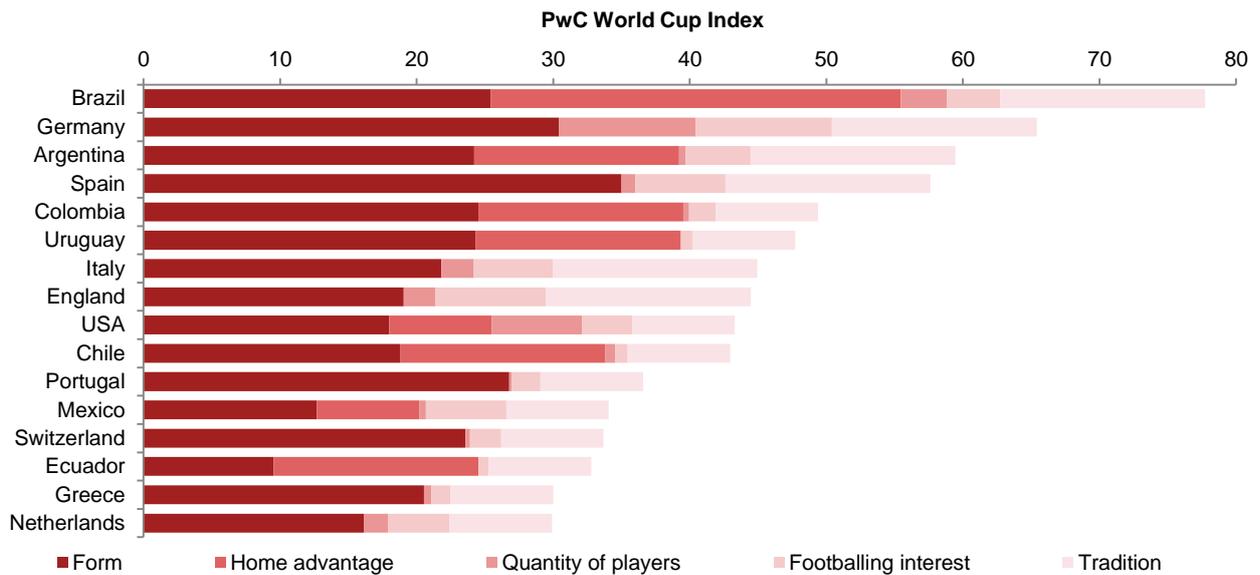
The weightings were based on our analysis of the strength of the historical correlation between each variable and World Cup performance, but with an overlay of judgement to reflect the latest available data on current form. The latter was weighted most heavily because our analysis showed that a small increase in FIFA rankings is associated with a large increase in the likelihood of a country progressing further in the tournament.

The estimated PwC World Cup Index can be used as a yardstick to judge the performances of countries at the 2014 World Cup. From Figure 8 below, Brazil’s combination of footballing tradition and home advantage suggests that they are the favourites based on our index. With the potential economic and sporting gains from hosting a mega-event, Brazil will be looking to win both off the pitch and on the pitch this summer, and our index suggests that they are in a good position to do the latter.

But Brazil will not have it easy. The following three teams of Germany, Argentina and Spain would all be expecting to challenge hard and reach at least the semi-finals, with estimated World Cup strength noticeably greater than the countries below them. There is then a larger group of countries, including England in eighth place, who would be expected based on our index to be competing for a spot in the quarter-finals, given their relative strength.

⁴ The Guardian, *Who has the hardest World Cup draw?*, <http://www.theguardian.com/football/interactive/2013/dec/18/world-cup-2014-draw-strength-of-schedule>

Figure 8: Estimated strength of selected teams in World Cup 2014

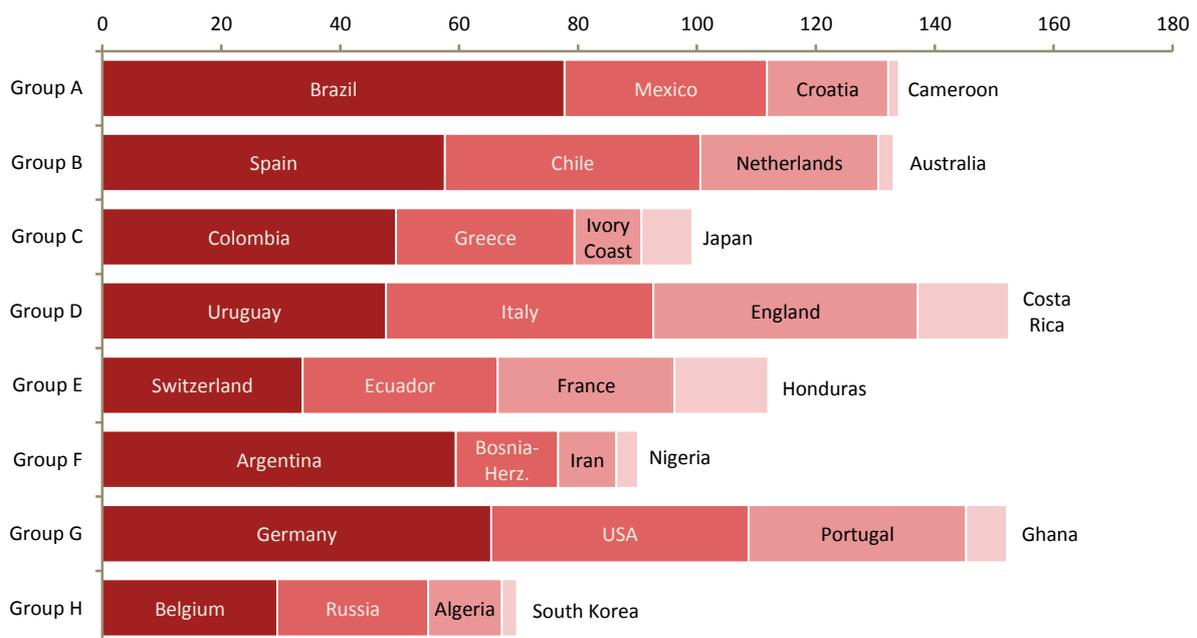


We then added the estimated index scores of the countries in each World Cup group to determine which group contained the strongest combination of teams and was therefore the ‘Group of Death’.

As shown in Figure 9, Group D and Group G are almost identical based on our index. Group D has a slightly higher total combined score, however, and is therefore deemed the Group of Death, given the collective tradition and pedigree of Uruguay, England and Italy. Although our indicator is clearly based on subjective weightings, Group D contains three

countries in the top 10 of the all-time World Cup table, that have won seven out of the 19 previous World Cups, so it does appear that this is a particularly strong group. However, Group G would be the Group of Death if we defined it as the combined strength of the top three teams, using the logic that only the top two countries qualify for the knockout stages. It has the lowest average FIFA ranking of any group, as well as the two countries with the largest pool of footballers, Germany and the USA, with a combined 10.5 million registered players.

Figure 9: The estimated strength of each 2014 World Cup group



Annex: Technical details of regression models

We have estimated two alternative regression models in this study, which are presented in Table 1 and Table 2 below. The first model is a cross-sectional regression of historic World Cup performance, where the dependent variable is total World Cup points – a measure which awards three points for a win, one point for a draw and zero points for a loss for each World Cup finals game played. The 56 countries which have appeared in at least 6 World Cup finals games were included in our model. We regressed total World Cup points against the following variables: GDP per capita, population, number of registered football players, average attendances at top division domestic football matches, a dummy variable for South American countries, a dummy variable for European countries, and a dummy variable for countries that have bid to host the current or previous World Cups at least three times.

Population and GDP per capita were insignificant and were therefore dropped from our final specification. The explanatory variables in our final specification were all statistically significant at the 5% level apart from average attendance, which was significant at the 10% level⁵. The constant was insignificant. The unit for average attendance was thousands and the units for registered players was hundreds of thousands – for example our model estimates that an additional million registered players would be associated with an additional 11 World Cup points.

Table 1: Regression results for Model 1

Explanatory variables	Coefficient	Standard error	t-statistic
Constant	(8.75)	6.80	(1.29)
Registered players	1.10	0.47	2.35
Average attendance	1.26	0.70	1.80
South America	40.02	10.50	3.81
Europe	29.19	8.16	3.58
Hosting bids	43.58	14.32	3.04

R -squared = 71%, F =24.4

⁵ These levels refer to there being less than a 5%, or less than a 10%, chance that, given the modelled correlation, the variables are not in fact related to World Cup performance after controlling for the other variables in the model. We therefore have a high degree of confidence (90-95% or more) that an underlying relationship does in fact exist.

The second model is a panel regression of the performances of each country in each World Cup. The dependent variable captured the round of the tournament that was reached, with the following scores allocated:

- 0 if the country is knocked out in the group stage
- 1 if the country reaches the round of 16
- 2 if the country reaches the quarter finals
- 3 if the country reaches the semi finals
- 4 if the country reaches the finals
- 5 if the country wins the World Cup

Using a panel regression meant that variables which change between each World Cup could be included, so the second model produced a different specification to the first one. We regressed World Cup performance against the following variables: GDP per capita, population, number of registered football players, average attendances at top division domestic football matches, a dummy variable for South American countries, a dummy variable for European countries, a dummy variable for host continent, a dummy variable for host country, and two variables to capture performance in each of the last two World Cups.

GDP per capita, population and average attendance were all found to be insignificant and were dropped from the final specification. The explanatory variables were all significant at the 5% level, apart from host continent which was significant at the 10% level.

Table 2: Regression results for Model 2

Explanatory variables	Coefficient	Standard error	t-statistic
Constant	0.03	0.15	0.22
Host country	1.69	0.31	5.54
Host continent	0.25	0.14	1.71
Europe	0.70	0.18	3.96
South America	0.91	0.21	4.22
Last World Cup	0.13	0.05	2.36
Last but one World Cup	0.16	0.06	2.83
Registered players	0.02	0.01	3.80

R -squared = 30%, F =22.2

Discussion of results: top-down vs bottom-up approaches

In this paper we have identified the relationship between a number of variables and historic World Cup performance, and we have used these relationships to assess the varying prospects of countries competing in the 2014 World Cup. Although this top-down, quantitative approach is useful for explaining the common drivers behind World Cup success, on its own it cannot provide robust predictions of the likely outcomes in Brazil this summer.

This is because it does not capture bottom-up factors that affect teams such as the form of players, injuries such as those to the meniscus or metatarsal, the ability (or otherwise) to strike a stationary ball into goal from 12 yards out, or the opinions of experts regarding football ability, such as the 'golden generation' of Belgian players that have emerged in the last couple of years. Also, because the outcomes of the World Cup rest on the outcome of a small number of matches, where luck can play a major part, there is a small sample size to draw conclusions from.

For these reasons, it is difficult to produce reliable forecasts of the outcomes of the 2014 World Cup based on econometrics alone. However, we do not face these types of issues when forecasting in business, for example when projecting the growth of an industry. We can analyse a large number of repeated transactions such as sales, which provides a much greater sample size for analysis. Also, we can supplement our top-down quantitative approach with bottom-up analysis of microeconomic drivers and qualitative inputs from industry experts, to produce richer and more robust estimates.

Contacts and services

Report authors

For more information about this report please contact one of the authors below:

John Hawskworth

E: john.c.hawskworth@uk.pwc.com

T: +44 (0) 20 7213 1650

William Zimmern

E: william.zimmern@uk.pwc.com

T: +44 (0) 20 7212 2750

Dan Broadfield

E: daniel.t.broadfield@uk.pwc.com

T: +44 (0) 20 7213 3546

PwC's Economics and Policy team

We help you understand how big economic, demographic, social, and environmental changes affect your organisation by setting out scenarios that identify growth opportunities and risks on a global, regional, national and local level. We help make strategic and tactical operational, pricing and investment decisions to support business value creation. We work together with you to achieve sustainable growth.

For more information about these services please visit our website: <http://www.pwc.co.uk/economics-policy>.

PwC's Mega Events Centre of Excellence

We help stakeholders of sports mega events to be more effective in the planning, preparation and staging of mega events. PwC's Mega Events Centre of Excellence comprises a network of professionals experienced in bringing value to all phases of the sports mega events lifecycle.

For more information about these services please visit our website: <http://www.pwc.com/sports-mega-events>.



PwC firms help organisations and individuals create the value they're looking for. We're a network of firms in 158 countries with close to 169,000 people who are committed to delivering quality in assurance, tax and advisory services. Tell us what matters to you and find out more by visiting us at <http://www.pwc.com>.

This publication has been prepared for general guidance on matters of interest only, and does not constitute professional advice. You should not act upon the information contained in this publication without obtaining specific professional advice. No representation or warranty (express or implied) is given as to the accuracy or completeness of the information contained in this publication, and, to the extent permitted by law, PricewaterhouseCoopers LLP, its members, employees and agents do not accept or assume any liability, responsibility or duty of care for any consequences of you or anyone else acting, or refraining to act, in reliance on the information contained in this publication or for any decision based on it.

© 2014 PricewaterhouseCoopers LLP. All rights reserved. In this document, "PwC" refers to PricewaterhouseCoopers LLP (a limited liability partnership in the United Kingdom), which is a member firm of PricewaterhouseCoopers International Limited, each member firm of which is a separate legal entity