



The following is an analysis from Goldman Sachs Research economists Jan Hatzius, Jari Stehn and Donnie Millar, originally published June 3, 2016.

NEW: SEE THE PREDICTION MODEL UPDATE PUBLISHED ON 3 JULY 2016 [[/our-thinking/pages/euro-2016-update-03-jul-2016.html](#)]

The Econometrician's Take on Euro 2016

- We present a model for predicting the outcome of the 2016 European Football championship in France from June 10 to July 10. It is similar to our model for the 2014 World Cup. Using historical performance data for each team—most importantly the Elo rating system originally devised to rank chess players—we estimate a set of probabilities that a particular team will reach a particular round, up to and including the championship. We also provide a modal “most likely” case for how the tournament will unfold (although “most likely” does not mean “likely”).
- The model says that France has a 23% probability of winning the trophy, followed by Germany at 20%, Spain at 14%, and England at 11%. Although Germany has the highest Elo rating, France is slightly favored because of its home advantage. After each day of play, we will re-run the model using updated historical performance data in order to generate new probabilities and a new modal forecast.
- How much faith should we have in these predictions? On the plus side, our approach carefully considers the stochastic nature of the tournament using statistical methods; also, the predictions are not far from bookmakers' odds. On the minus side, the environment is “stochastic” indeed, i.e., football is quite an unpredictable game!
- That charming unpredictability was on full display two years ago, when our model failed to anticipate the elimination of heavyweights Spain and Italy in the group stage and gave Brazil a 48% probability of winning the trophy. More encouragingly, it identified three of the four semifinalists before the start of the tournament, and the fully updated version predicted the winner of every match in the knockout stage except for the 7-1 semifinal between Germany and Brazil.

Below we introduce our statistical model for predicting the outcome of the 2016 European Football Championship in France from June 10 to July 10.

How the Model Works

At a high level, our approach is as follows. First, we estimate a regression model to predict the number of goals scored by a particular team (“team i”) against a particular opponent

on predicting football matches, we assume that the number of goals scored by team i is described by a so-called Poisson distribution and explained by the following statistical factors:^[2]

1. The difference in team performance as reflected in Elo ratings prior to the match. The Elo system was originally devised to rank chess players. It is a composite measure of national football team success that evolves depending on a team's results and the strength of its opponents.
2. The number of goals scored by team i in the last 10 competitive matches.
3. The number of goals conceded by team j in the last 2 competitive matches.
4. A home dummy.
5. A European Championship dummy to capture whether a team does systematically better at European Championships than in other competitive matches

Second, we use these regression estimates and our assumed Poisson distribution in a Monte Carlo simulation with 100,000 draws to generate a distribution of outcomes for each of the 52 matches, from the opener between France and Romania on June 10 to the final on July 10. We use the rounded prediction of the goals scored to determine the outcome of each match during the group stage and the unrounded prediction to pick the winner in the knockout stage.

Third, we use the estimation results to generate both a set of probabilities that a particular team reaches a particular stage of the tournament, up to and including the championship, and a modal—that is, single most likely—forecast for the outcome of each match, which we then run forward through the tournament until the final.

A Summary of Predictions

Our probabilities are shown in Exhibit 1. The model says that France has a 23% probability of winning the trophy, followed by Germany at 20%, Spain at 14%, and England at 11%. Although Germany has the highest Elo rating, France is favored because of its home advantage.

OVERVIEW HOW THE MODEL WORKS A SUMMARY OF THE PREDICTIONS HOW CONFIDENT CAN WE BE?

						Rating
France	97.5	73.0	54.1	34.8	23.1	1947
Germany	95.7	72.5	50.0	31.9	19.9	2037
Spain	88.3	61.6	39.4	24.0	13.6	1977
England	92.0	60.7	35.7	20.4	10.5	1934
Portugal	91.9	56.0	30.3	16.5	8.0	1885
Belgium	84.7	47.8	25.3	12.5	5.7	1896
Russia	77.7	39.9	17.1	6.9	2.6	1747
Croatia	63.6	30.5	14.0	5.9	2.2	1794
Ukraine	68.8	32.3	14.4	5.7	1.9	1797
Czech Republic	61.8	28.8	12.7	5.1	1.7	1728
Poland	68.9	31.2	13.3	4.8	1.5	1762
Italy	65.2	27.2	11.1	4.5	1.5	1844
Turkey	57.0	25.0	10.5	3.9	1.3	1797
Switzerland	65.0	26.5	10.2	3.5	1.2	1753
Austria	70.6	29.1	10.1	3.6	1.1	1766
Ireland	62.5	25.0	9.9	3.5	1.1	1751
Romania	65.2	26.2	10.3	3.7	1.0	1738
Sweden	58.0	22.3	8.1	2.8	0.9	1735
Slovakia	58.8	23.6	7.9	2.3	0.7	1715
Hungary	57.8	19.6	5.7	1.7	0.3	1671
Iceland	47.5	14.3	3.5	1.0	0.2	1647
Wales	39.5	11.9	2.7	0.5	0.1	1638
Northern Ireland	30.3	7.6	1.9	0.5	0.1	1578
Albania	31.6	7.6	1.8	0.4	0.0	1578

Source: Goldman Sachs Global Investment Research

Exhibits 2 and 3 provide a different perspective by showing the modal prediction for the entire tournament. There are some interesting contrasts with the probabilities in Exhibit 1. For example, Exhibit 1 says that Germany is more likely than Spain to win the tournament because it is more likely to succeed across the entire range of possible tournament configurations. But Exhibit 3 says that in the single most likely case, Spain beats England in Semifinal 1, France beats Germany in Semifinal 2, and France then wins the final—i.e., Germany finishes behind Spain.

Which approach is better, the probabilistic one in Exhibit 1 or the modal one in Exhibits 2 and 3? A modal forecast does have the advantage of being more “crisp.” The sentence “Goldman Sachs says France will win” has a better ring to it than “Goldman Sachs says France has a 23% probability of winning, with Germany close behind.” Nevertheless, we think that a probabilistic approach is more useful—for predicting the outcome of football tournaments and, increasingly, for our day-to-day work on economic forecasting.^[3]

OVERVIEW

HOW THE MODEL WORKS

A SUMMARY OF THE PREDICTIONS

HOW CONFIDENT CAN WE BE?

A	France 3	Albania 1
A	Albania 1	Switzerland 1
A	Romania 1	Switzerland 1
A	France 2	Switzerland 1
A	Romania 1	Albania 1
A	France 2	Romania 1
B	Russia 1	Slovakia 1
B	Russia 2	Wales 1
B	England 2	Slovakia 1
B	England 2	Russia 1
B	Wales 1	Slovakia 1
B	England 2	Wales 1
C	Ukraine 1	Poland 1
C	Germany 2	Northern Ireland 1
C	Ukraine 1	Northern Ireland 1
C	Poland 1	Northern Ireland 1
C	Germany 2	Poland 1
C	Germany 2	Ukraine 1
D	Spain 2	Czech Republic 1
D	Croatia 1	Spain 1
D	Turkey 1	Czech Republic 1
D	Croatia 1	Czech Republic 1
D	Turkey 1	Spain 1
D	Turkey 1	Croatia 1
E	Ireland 1	Belgium 1
E	Belgium 1	Italy 1
E	Sweden 1	Italy 1
E	Ireland 1	Italy 1
E	Sweden 1	Belgium 1
E	Ireland 1	Sweden 1
F	Austria 1	Hungary 1
F	Austria 1	Iceland 1
F	Hungary 1	Portugal 2
F	Portugal 2	Iceland 1
F	Hungary 1	Iceland 1
F	Austria 1	Portugal 1

Source: Goldman Sachs Global Investment Research

Exhibit 4 provides more insight into the results by breaking down the probabilities of winning for the top four teams in a “waterfall chart” format. It shows that the most important factor is the Elo score, followed by home advantage and the European Championship dummy. The chart illustrates that the front-runner position for France derives largely from its home advantage, as its Elo rating is well below Germany’s and also a bit below Spain’s. Meanwhile, Germany benefits from the European Championship dummy, which picks up its historically strong tournament performance.

OVERVIEW

HOW THE MODEL WORKS

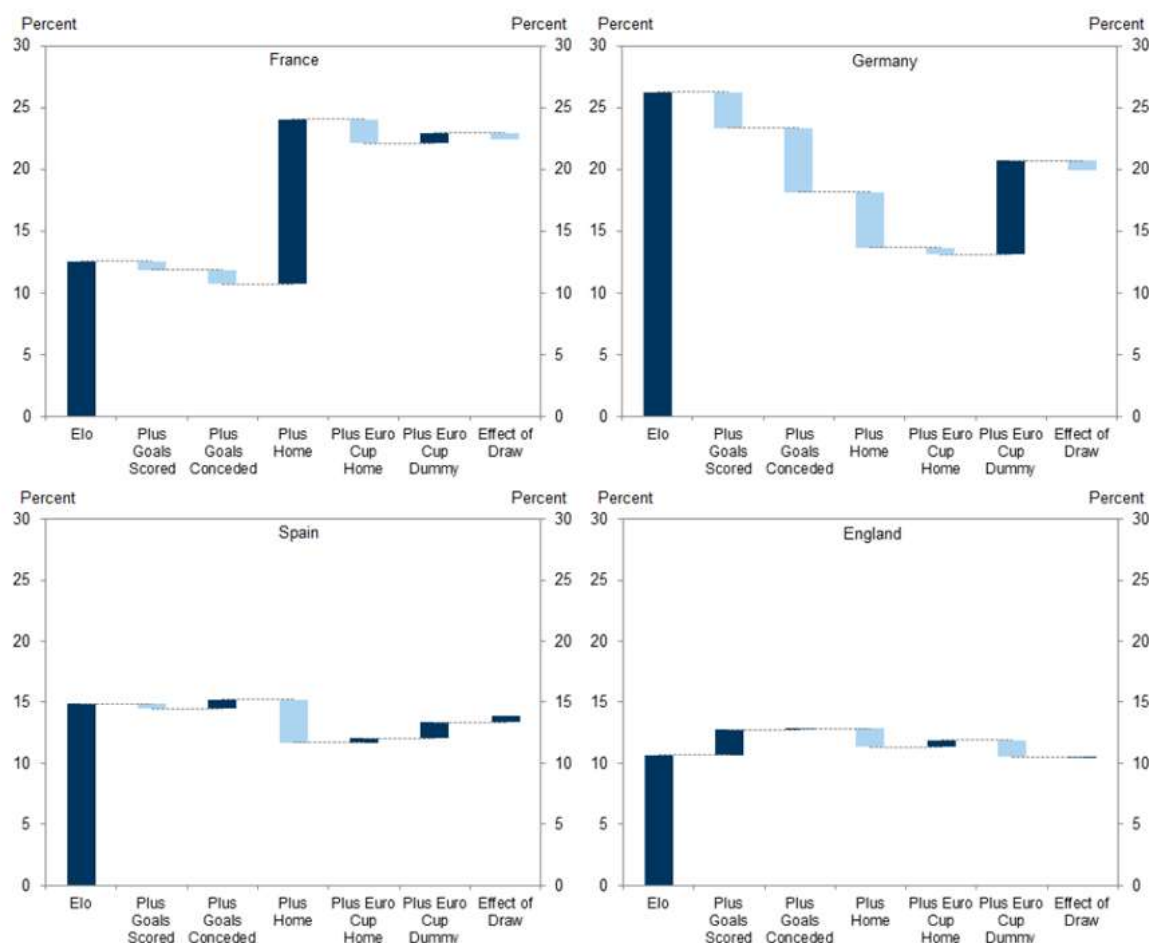
A SUMMARY OF THE PREDICTIONS

HOW CONFIDENT CAN WE BE?



Source: Goldman Sachs Global Investment Research

Exhibit 4: Breaking Down the Championship Probability in Euro 2016



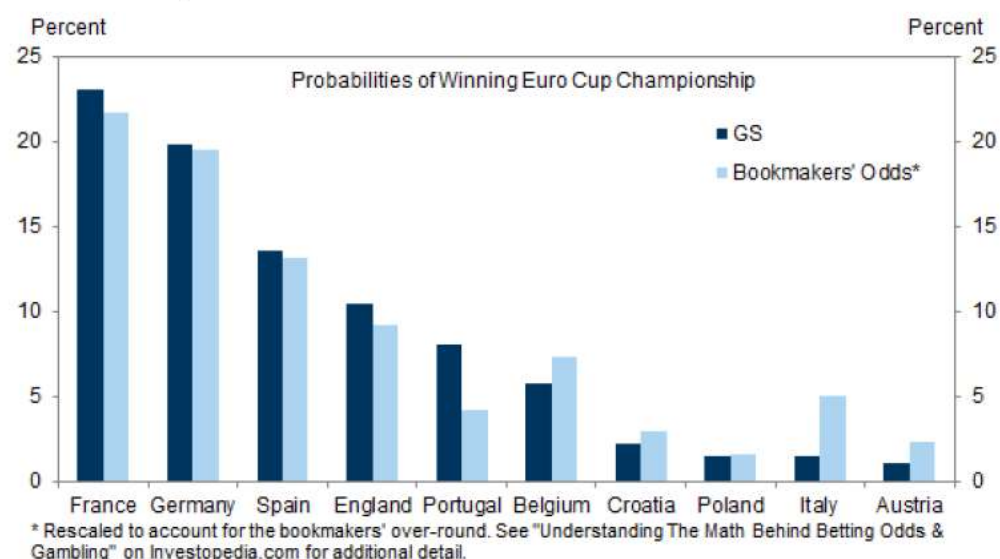
Source: Goldman Sachs Global Investment Research

How Confident Can We Be?

It is difficult to assess how much faith one should have in these predictions. On the plus side, our approach carefully considers the stochastic nature of the tournament using statistical methods, and we do think that the Elo rating—the most important input into our analysis—is a compelling summary of a team's track record. On the minus side, we ignore a number of potentially important factors that are difficult to summarize statistically, including the quality

us are really football experts but some are enthusiastic Germany supporters).^[3]

Exhibit 5: Not Very Far from Bookmakers' Odds



Source: Coral.com, Bet365.com, Betvictor.com, Skybet.com, Ladbrokes.com, Goldman Sachs Global Investment Research

One useful cross-check is to compare our results with bookmakers' odds. Exhibit 5 plots our estimated championship probability against the average probability implied by the odds offered by five different bookmakers. The basic result is clear. Even though our model does not include bookmakers' odds in any way, the probabilities are quite similar. A possible reason is that professional betting firms use many of the same inputs—such as Elo ratings—in their analysis and that they process the information in ways that are ultimately similar to ours.

Another useful check is to evaluate the performance of our model for the 2014 World Cup, which followed an essentially identical approach to the one presented here.^[6] It is safe to say that we had our hits and misses.

First, performance in the group stage was not great. The model only identified 9 of the 16 advancing teams and failed to predict the elimination of heavyweights Spain and Italy, although it correctly anticipated that England would fly home early.

Second, the model gave Brazil a 48% probability of winning the trophy, by far the highest of all the contestants. That failure illustrates a certain lack of imagination that is inherent in our approach. If the greatest football nation on earth—in terms of both past victories and its 2014 Elo rating—plays a World Cup at home, we are bound to project success. At least the probability was below 50%!

Third, the model did correctly identify three of the four semifinalists before the start of the tournament, namely Argentina, Brazil, and Germany, although it incorrectly picked Spain over the Netherlands.

Fourth, the fully updated version of the model—that is, the projection we sent out before each day of play on the basis of updated Elo ratings and other performance measures—was remarkably accurate during the knockout stage. It correctly predicted the winner of every match except the 7-1 semifinal between Germany and Brazil. But that was, by one estimate, the single most surprising result in World Cup history.^[7]

in our comment at the time: “Speaking as forecasters, we regret the miss. But, speaking as Germans, we would note that there are more important things than being right.”

May the best team win and let’s hope that watching Euro 2016 is as much fun as it was to write this article!

Jan Hatzius
Sven Jari Stehn
Donnie Millar

- 1. The database is taken from <http://www.eloratings.net/>.
- 2. See, for example, A. Heuer, C. Müller, and O. Rubner, "Soccer: Is scoring goals a predictable Poissonian process?" Europhysics Letters, 2010.
- 3. As it happens, we recently shifted our approach to forecasting near-term Fed policy in a more probabilistic direction. See “Superforecasting Fed Policy,” US Daily, May 23, 2016.
- 4. This means, for example, that we cannot include factors such as the retirement or injury of a key player.
- 5. One result that looks surprising to our not-so-expert eyes is the low probability of success for Italy in Exhibit 1. The mechanical reason is that Italy’s current Elo rating is not very high and Italy has historically not performed very well at European Championships.
- 6. See The World Cup and Economics, 2014.
- 7. See Nate Silver, “The Most Shocking Result in World Cup History,” <http://fivethirtyeight.com/datalab/the-most-shocking-result-in-world-cup-history/>, July 8, 2014.

CAREERS	WHO WE ARE	WHAT WE DO	OUR THINKING	CITIZENSHIP
INVESTOR RELATIONS	Our Firm At a Glance	Overview	Macroeconomic Insights	Corporate Engagement Ov
MEDIA RELATIONS	Leadership	Investment Banking	The New Energy Landscape	10,000 Women
	People and Culture	Securities	Technology Driving Innovation	10,000 Small Businesses
WORLDWIDE	Business Standards	Investing and Lending	Talks at GS	Goldman Sachs Gives
CONSUMER DEPOSITS	Diversity and Inclusion	Investment Management	Podcast: ‘Exchanges at Goldman Sachs’	Community Teamworks
GS Bank	Locations	Research	Public Policy	Environmental, Social and Governance Report
CONSUMER LENDING	Awards	Engineering	Archive	Environmental Stewardship
Marcus by Goldman Sachs	Alumni		Sources	Sponsorships
	Stories of Progress			Supporting Veterans
	Interactive Guide to Capital Markets			