universität innsbruck

Research Platform
Empirical and Experimental Economics

working paper

eeecon
[triple:e:con]

Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model

Achim Zeileis, Christoph Leitner, Kurt Hornik

Working Papers in Economics and Statistics

2018-09

University of Innsbruck
https://www.uibk.ac.at/eeecon/

University of Innsbruck
Working Papers in Economics and Statistics

The series is jointly edited and published by

- Department of Banking and Finance
- Department of Economics
- Department of Public Finance
- Department of Statistics

Contact address of the editor:
research platform “Empirical and Experimental Economics”
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
Tel: + 43 512 507 71022
Fax: + 43 512 507 2970
E-mail: eeecon@uibk.ac.at

The most recent version of all working papers can be downloaded at
https://www.uibk.ac.at/eeecon/wopec/

For a list of recent papers see the backpages of this paper.

Probabilistic Forecasts for the 2018 FIFA World Cup Based on the Bookmaker Consensus Model

Achim Zeileis
Universität Innsbruck

Christoph Leitner
WU Wirtschafts-
universität Wien

Kurt Hornik
WU Wirtschafts-
universität Wien

Abstract

Football fans worldwide anticipate the 2018 FIFA World Cup that will take place in Russia from 14 June to 15 July 2018. 32 of the best teams from 5 confederations compete to determine the new World Champion. Using a consensus model based on quoted odds from 26 bookmakers and betting exchanges a probabilistic forecast for the outcome of the World Cup is obtained. The favorite is Brazil with a forecasted winning probability of 16.6%, closely followed by the defending World Champion and 2017 FIFA Confederations Cup winner Germany with a winning probability of 15.8%. Two other teams also have winning probabilities above 10%: Spain and France with 12.5% and 12.1%, respectively.

The results from this bookmaker consensus model are coupled with simulations of the entire tournament to obtain implied abilities for each team. These allow to obtain pairwise probabilities for each possible game along with probabilities for each team to proceed to the various stages of the tournament. This shows that indeed the most likely final is a match of the top favorites Brazil and Germany (with a probability of 5.5%) where Brazil has the chance to compensate the dramatic semifinal in Belo Horizonte, four years ago. However, given that it comes to this final, the chances are almost even (50.6% for Brazil vs. 49.4% for Germany). The most likely semifinals are between the four top teams, i.e., with a probability of 9.4% Brazil and France meet in the first semifinal (with chances slightly in favor of Brazil in such a match, 53.5%) and with 9.2% Germany and Spain play the second semifinal (with chances slightly in favor of Germany with 53.1%).

These probabilistic forecasts have been obtained by suitably averaging the quoted winning odds for all teams across bookmakers. More precisely, the odds are first adjusted for the bookmakers' profit margins ("overrounds"), averaged on the log-odds scale, and then transformed back to winning probabilities. Moreover, an "inverse" approach to simulating the tournament yields estimated team abilities (or strengths) from which probabilities for all possible pairwise matches can be derived. This technique (Leitner, Zeileis, and Hornik 2010a) correctly predicted the winner of 2010 FIFA World Cup (Leitner, Zeileis, and Hornik 2010b) and three out of four semifinalists at the 2014 FIFA World Cup (Zeileis, Leitner, and Hornik 2014). Interactive web graphics for this report are available at: https://eeecon.uibk.ac.at/~zeileis/news/fifa2018/.

Keywords: consensus, agreement, bookmakers odds, tournament, 2018 FIFA World Cup.

1. Bookmaker consensus

In order to forecast the winner of the 2018 FIFA World Cup, we obtained long-term winning odds from 26 online bookmakers (also including two betting exchanges, see Tables 3 and 4 at the end). However, before these odds can be transformed to winning probabilities, the

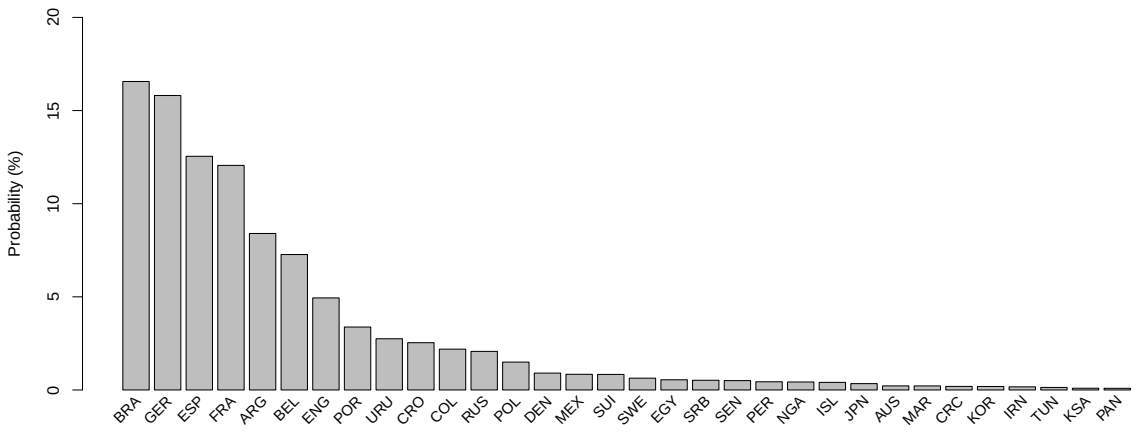

Figure 1: 2018 FIFA World Cup winning probabilities from the bookmaker consensus model.

stake has to be accounted for and the profit margin of the bookmaker (better known as the "overround") has to be removed (for further details see Henery 1999; Forrest, Goddard, and Simmons 2005). Here, it is assumed that the quoted odds are derived from the underlying "true" odds as: $\textit{quoted odds} = \textit{odds} \cdot \delta + 1$, where $+1$ is the stake (which is to be paid back to the bookmakers' customers in case they win) and $\delta < 1$ is the proportion of the bets that is actually paid out by the bookmakers. The overround is the remaining proportion $1 - \delta$ and the main basis of the bookmakers' profits (see also Wikipedia 2018 and the links therein). Assuming that each bookmaker's δ is constant across the various teams in the tournament (see Leitner *et al.* 2010a, for all details), we obtain overrounds for all 26 bookmakers with a median value of 15.2%.

To aggregate the overround-adjusted odds across the 26 bookmakers, we transform them to the log-odds (or logit) scale for averaging (as in Leitner *et al.* 2010a). The bookmaker consensus is computed as the mean winning log-odds for each team across bookmakers (see column 4 in Table 1) and then transformed back to the winning probability scale (see column 3 in Table 1). Figure 1 shows the barchart of winning probabilities for all 32 competing teams.

According to the bookmaker consensus model, Brazil is most likely to win the tournament (with probability 16.6%) followed by the current FIFA World Champion Germany (with probability 15.8%). The only other teams with double-digit winning probabilities are France (with 12.5%) and Spain (with 12.1%).

Although forecasting the winning probabilities for the 2018 FIFA World Cup is probably of most interest, we continue to employ the bookmakers' odds to infer the contenders' relative abilities (or strengths) and the expected course of the tournament. To do so, an "inverse" tournament simulation based on team-specific abilities is used. The idea is the following:

1. If team abilities are available, pairwise winning probabilities can be derived for each possible match (see Section 2).

2. Given pairwise winning probabilities, the whole tournament can be easily simulated to see which team proceeds to which stage in the tournament and which team finally wins.

3. Such a tournament simulation can then be run sufficiently often (here 1,000,000 times) to obtain relative frequencies for each team winning the tournament.

Team	FIFA code	Probability	Log-odds	Log-ability	Group
Brazil	BRA	16.6	−1.617	−1.778	E
Germany	GER	15.8	−1.673	−1.801	F
Spain	ESP	12.5	−1.942	−1.925	B
France	FRA	12.1	−1.987	−1.917	C
Argentina	ARG	8.4	−2.389	−2.088	D
Belgium	BEL	7.3	−2.546	−2.203	G
England	ENG	4.9	−2.957	−2.381	G
Portugal	POR	3.4	−3.353	−2.486	B
Uruguay	URU	2.7	−3.566	−2.566	A
Croatia	CRO	2.5	−3.648	−2.546	D
Colombia	COL	2.2	−3.799	−2.626	H
Russia	RUS	2.1	−3.856	−2.650	A
Poland	POL	1.5	−4.186	−2.759	H
Denmark	DEN	0.9	−4.695	−2.897	C
Mexico	MEX	0.8	−4.769	−2.908	F
Switzerland	SUI	0.8	−4.777	−2.929	E
Sweden	SWE	0.6	−5.055	−3.009	F
Egypt	EGY	0.5	−5.202	−3.010	A
Serbia	SRB	0.5	−5.252	−3.033	E
Senegal	SEN	0.5	−5.288	−3.061	H
Peru	PER	0.4	−5.421	−3.043	C
Nigeria	NGA	0.4	−5.448	−3.067	D
Iceland	ISL	0.4	−5.498	−3.063	D
Japan	JPN	0.3	−5.680	−3.163	H
Australia	AUS	0.2	−6.121	−3.250	C
Morocco	MAR	0.2	−6.131	−3.278	B
Costa Rica	CRC	0.2	−6.261	−3.321	E
South Korea	KOR	0.2	−6.277	−3.298	F
Iran	IRN	0.2	−6.388	−3.281	B
Tunisia	TUN	0.1	−6.599	−3.389	G
Saudi Arabia	KSA	0.1	−6.975	−3.514	A
Panama	PAN	0.1	−7.024	−3.473	G

Table 1: Bookmaker consensus model for the 2018 FIFA World Cup, obtained from 26 online bookmakers. For each team, the consensus winning probability (in %), corresponding log-odds, simulated log-abilities, and group in tournament is provided.

Here, we use the iterative approach of Leitner *et al.* (2010a) to find team abilities so that the resulting simulated winning probabilities (from 1,000,000 runs) closely match the bookmaker consensus probabilities. This allows to strip the effects of the tournament draw (with weaker/easier and stronger/more difficult groups), yielding the log-ability measure (on the log-odds scale) in Table 1.

2. Pairwise comparisons

A classical approach to modeling winning probabilities in pairwise comparisons (i.e., matches between teams/players) is that of Bradley and Terry (1952) similar to the Elo rating (Elo 2008), popular in sports. The Bradley-Terry approach models the probability that a Team A beats a Team B by their associated abilities (or strengths):

$$\Pr(A \text{ beats } B) = \frac{ability_A}{ability_A + ability_B}.$$

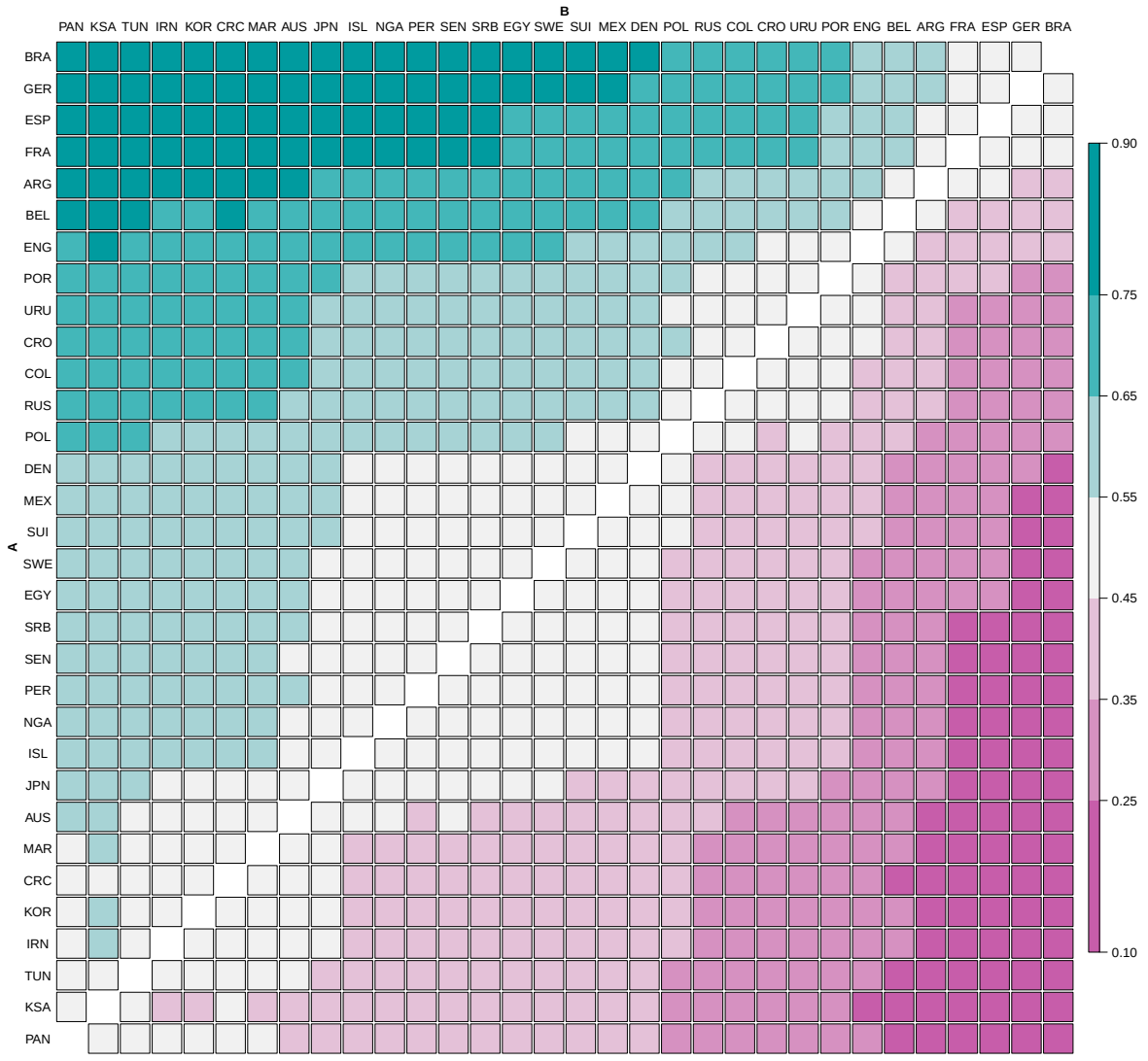

Figure 2: Winning probabilities in pairwise comparisons of all 2018 FIFA World Cup teams. Light gray signals that either team is almost equally likely to win a match between Teams A and B (probability between 45% and 55%). Light, medium, and dark green/pink corresponds to small, moderate, and high probabilities of winning/losing a match between Team A and Team B.

As explained in Section 1, the abilities for the teams in the 2018 FIFA World Cup can be chosen such that when simulating the whole tournament with these pairwise winning probabilities $\Pr(A \text{ beats } B)$, the resulting winning probabilities for the whole tournament are close to the bookmaker consensus winning probabilities. Table 1 reports the log-abilities for all teams and the corresponding pairwise winning probabilities are visualized in Figure 2.

Clearly, the bookmakers perceive Brazil and Germany to be the strongest teams in the tournament that are almost on par, followed by France and Spain which are again virtually on par. The pairwise winning probabilities between these four top teams are close to even with the following winning probabilities for the (slightly) stronger team: Brazil vs. Germany 50.6%, Brazil vs. France 53.5%, Germany vs. France 52.9%, Brazil vs. Spain 53.7%, Germany vs. Spain 53.1%, France vs. Spain 50.2%. Behind this group with the four strongest teams four further teams constitute the "best of the rest": Argentina, Belgium, England, and Portugal. Then, there are several larger clusters of teams that have approximately the same strength (i.e., yielding approximately even chances in a pairwise comparison).

3. Performance throughout the tournament

Based on the teams' inferred abilities and the corresponding probabilities for all matches from Section 2 the whole tournament is simulated 1,000,000 times. As expounded above, the abilities have been calibrated such that the simulated winning proportions for each team closely match the bookmakers' consensus winning probabilities. So with respect to the probabilities of winning the tournament, there are no new insights. However, the simulations also yield simulated probabilities for each team to "survive" over the tournament, i.e., proceed from the group-phase to the round of 16, quarter- and semifinals, and the final.

Figure 3 depicts these "survival" curves for all 32 teams within the groups they were drawn in. France, Brazil, and Germany are the clear favorites within their respective groups C, E and F with almost 90% probability to make it to the round of 16 and also relatively small drops in probability to proceed through the subsequent rounds. Groups B, D, and G also have group favorites with good chances to proceed throughout the tournament but these also have a strong second contender: Spain and Portugal in group B, Argentina and Croatia in group D, Belgium and England in group G. In the remaining two groups, A and H, the strongest teams (Uruguay and Russia in group A, Columbia and Poland in group H) have good chances to proceed to the round of 16 but then the probability to proceed further drops sharply. This is due to probably meeting much stronger teams in the round of 16. See also Table 2 for the underlying numeric values.

4. Conclusions

Our forecasts for the 2018 FIFA World Cup follow closely our previous studies in Leitner *et al.* (2008, 2010b) and Zeileis *et al.* (2012, 2014, 2016) which correctly predicted the winner of the FIFA 2010 and Euro 2012 tournaments. While missing the winner for the other three tournaments, forecasts were still reasonably close: for Euro 2008 the correct final was predicted, three out of four semifinalists for the FIFA 2014 World Cup, and for the Euro 2016 it was correctly predicted that France would beat Germany in the semifinal. However, the latter tournament showed quite clearly that all forecasts are probabilistic and by no means

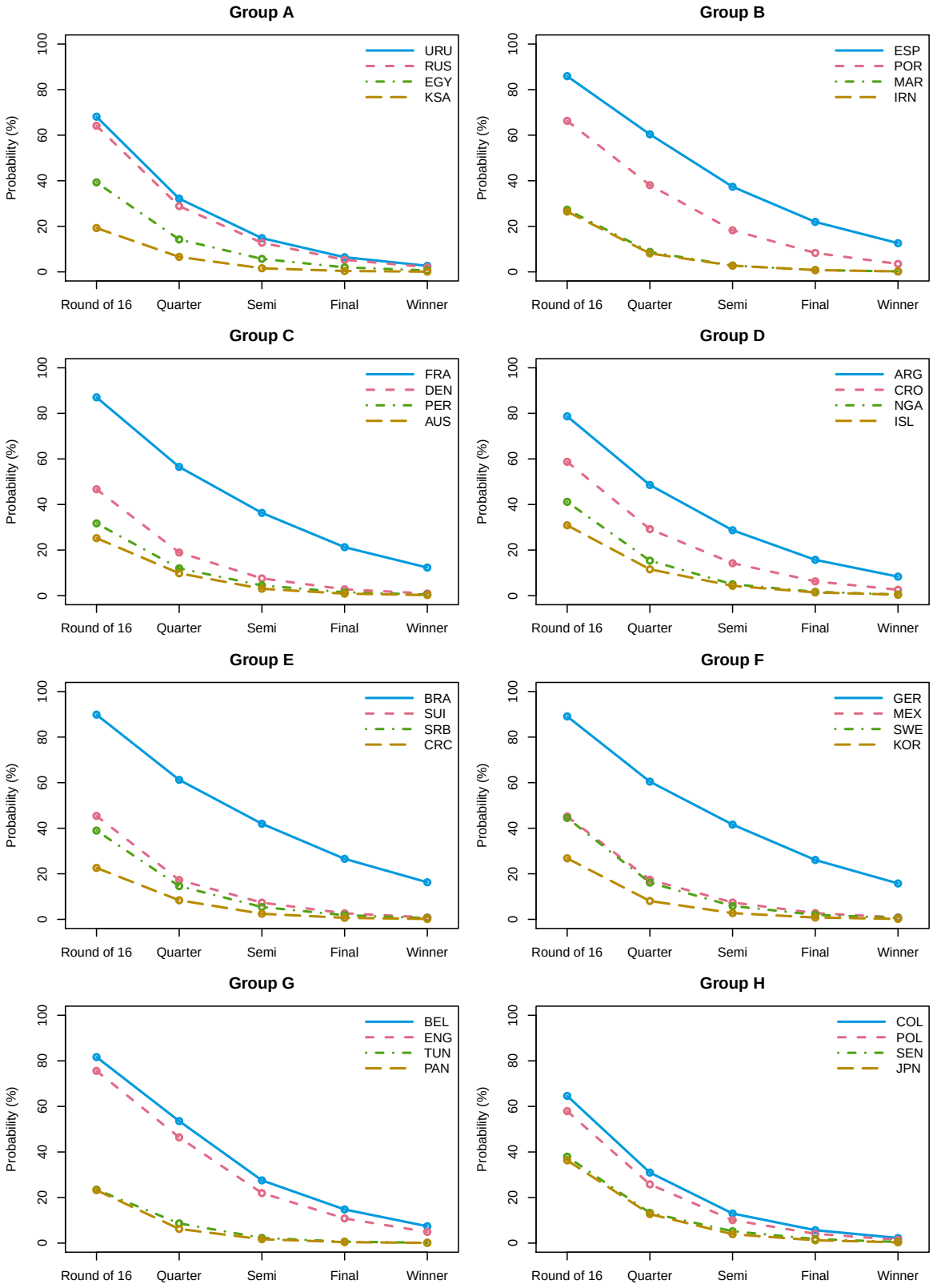

Figure 3: Probability for each team to "survive" in the 2018 FIFA World Cup, i.e., proceed from the group phase to the round of 16, quarter- and semifinals, the final and to win the tournament.

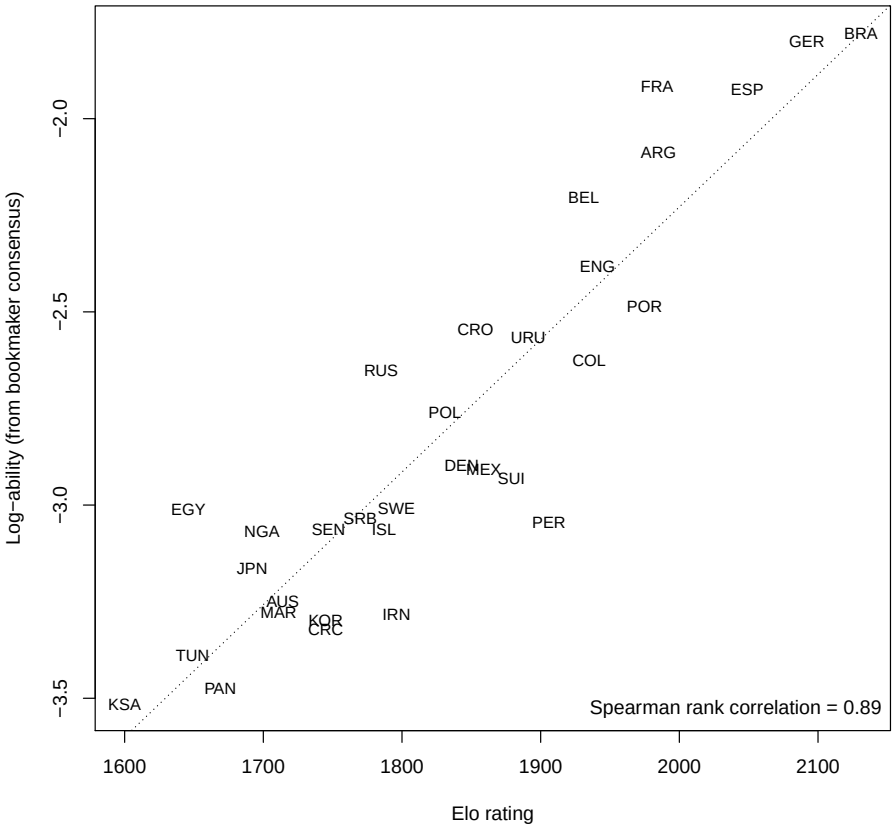

Figure 4: Bookmaker consensus log-ability vs. Elo rating for all 32 teams in the 2018 FIFA World Cup (along with least-squares regression line).

certain: France had a predicted 68.8% probability to beat Portugal, i.e., being expected to win about 2 out of every 3 matches between these two teams. However, in the actual final Gignac failed to seal the deal in added time and Portugal was able to take the victory in overtime.

The core idea of our method (Leitner *et al.* 2010a) is to use the expert knowledge of international bookmakers. These have to judge all possible outcomes in a sports tournament and assign odds to them. Doing a poor job (i.e., assigning too high or too low odds) will cost them money. Hence, in our forecasts we rely on the expertise of 26 such bookmakers (which actually also include two betting exchanges). Specifically, we (1) adjust the quoted odds by removing the bookmakers' profit margins (with median value of 15.2%), (2) aggregate and average these to a consensus rating, and (3) infer the corresponding tournament-draw-adjusted team abilities using a classical pairwise-comparison model.

Not surprisingly, our forecasts are closely related to other rankings of the teams in the 2018 FIFA World Cup, notably the FIFA and Elo ratings. The Spearman rank correlation of the consensus log-abilities with the FIFA rating is 0.76 and with the Elo rating even 0.89. However, the bookmaker consensus model allows for various additional insights, such as

the "survival" probabilities over the course of the tournament. Interestingly, when looking at the scatter plot of consensus log-abilities vs. the Elo rating in Figure 4 there are a few teams that are either clearly better (above the dotted least-squares regression line, e.g., Russia and Egypt) or worse (below the dotted line, e.g., Peru and Iran) in the forward-looking bookmakers' odds compared to the retrospective Elo rating. In case of Russia, this is surely the home advantage that bookmakers expect to be higher than the Elo rating – and in case of Egypt, this is likely due to the new superstar Mohamed Salah who is still expected to join his team after his injury in the final of the UEFA Champions League.

In addition to general team ratings like Elo and FIFA, various other methods have been introduced for forecasting FIFA World Cups. For example, some banks use their economic rating techniques and apply them also to forecasting major sports events (e.g., Goldman-Sachs Global Investment Research 2014; Danske Bank Research 2014). However, the probabilistic basis for these is often not so clear which can lead to a rather extreme winning probability for the favorite (e.g., up to almost 50% in the aforementioned reports). Better probabilistic results can be obtained by modeling team strengths and simulating tournament outcome "directly" – as opposed to our "inverse" simulation approach. Flexible modeling techniques have been suggested by Groll, Schauberger, and Tutz (2015) and Schauberger and Groll (2018), see the references therein for related approaches. Finally, Ekstrøm (2018) proposed a platform for comparing direct simulated predictions for the 2018 FIFA World Cup.

In summary, it can only be said with certainty that football fans are eagerly awaiting the tournament while its outcome cannot be known before the end of the final in Moscow on July 15. While Brazil and Germany have the best chances to win the World Cup in the bookmakers' expert opinions, it is still far more likely that one of the other teams wins. This is one of the two reasons why we would recommend to refrain from placing bets based on our analyses. The more important second reason, though, is that the bookmakers have a sizeable profit margin of about 15.2% which assures that the best chances of making money based on sports betting lie with them. Hence, this should be kept in mind when placing bets. We, ourselves, will not place bets but focus on enjoying the exciting football tournament that the 2018 FIFA World Cup will be with 100% predicted probability!

References

Bradley RA, Terry ME (1952). "Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons." *Biometrika*, **39**(3/4), 324–345. doi:10.2307/2334029.

Danske Bank Research (2014). "World Cup 2014 Special." Accessed 2014-06-17, URL http://www.danskeresearch.com/.

Ekstrøm CT (2018). "Predicting the Winner of the 2018 FIFA World Cup Predictions." Presented at "eRum 2018 – European R Users Meeting", Budapest, Hungary. Slides available online at http://biostatistics.dk/talks/eRum2018, video at https://www.youtube.com/watch?v=urJ1obHPsV8, and code at https://github.com/ekstroem/socceR2018.

Elo AE (2008). *The Rating of Chess Players, Past and Present*. Ishi Press, San Rafael.

Forrest D, Goddard J, Simmons R (2005). "Odds-Setters as Forecasters: The Case of En-

glish Football." *International Journal of Forecasting*, **21**(3), 551–564. doi:10.1016/j.ijforecast.2005.03.003.

Goldman-Sachs Global Investment Research (2014). "The World Cup and Economics 2014." Accessed 2018-05-28, URL http://www.goldmansachs.com/our-thinking/outlook/world-cup-and-economics-2014-folder/world-cup-economics-report.pdf.

Groll A, Schauberger G, Tutz G (2015). "Prediction of Major International Soccer Tournaments Based on Team-Specific Regularized Poisson Regression: An Application to the FIFA World Cup 2014." *Journal of Quantitative Analysis in Sports*, **11**(2), 97–115. doi:10.1515/jqas-2014-0051.

Henery RJ (1999). "Measures of Over-Round in Performance Index Betting." *Journal of the Royal Statistical Society D*, **48**(3), 435–439. doi:10.1111/1467-9884.00202.

Leitner C, Zeileis A, Hornik K (2008). "Who Is Going to Win the EURO 2008? (A Statistical Investigation of Bookmakers Odds)." *Report 65*, Department of Statistics and Mathematics, Wirtschaftsuniversität Wien, Research Report Series. URL http://epub.wu.ac.at/1570/.

Leitner C, Zeileis A, Hornik K (2010a). "Forecasting Sports Tournaments by Ratings of (Prob)Abilities: A Comparison for the EURO 2008." *International Journal of Forecasting*, **26**(3), 471–481. doi:10.1016/j.ijforecast.2009.10.001.

Leitner C, Zeileis A, Hornik K (2010b). "Forecasting the Winner of the FIFA World Cup 2010." *Report 100*, Institute for Statistics and Mathematics, WU Wirtschaftsuniversität Wien, Research Report Series. URL http://epub.wu.ac.at/702/.

Schauberger G, Groll A (2018). "Predicting Matches in International Football Tournaments with Random Forests." *Statistical Modelling*. Forthcoming.

Wikipedia (2018). "Odds — Wikipedia, The Free Encyclopedia." Online, accessed 2018-05-27, URL https://en.wikipedia.org/wiki/Odds.

Zeileis A, Leitner C, Hornik K (2012). "History Repeating: Spain Beats Germany in the EURO 2012 Final." *Working Paper 2012-09*, Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics, Universität Innsbruck. URL http://EconPapers.RePEc.org/RePEc:inn:wpaper:2012-09.

Zeileis A, Leitner C, Hornik K (2014). "Home Victory for Brazil in the 2014 FIFA World Cup." *Working Paper 2014-17*, Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics, Universität Innsbruck. URL http://EconPapers.RePEc.org/RePEc:inn:wpaper:2014-17.

Zeileis A, Leitner C, Hornik K (2016). "Predictive Bookmaker Consensus Model for the UEFA Euro 2016." *Working Paper 2016-15*, Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics, Universität Innsbruck. URL http://EconPapers.RePEc.org/RePEc:inn:wpaper:2016-15.

Team	Round of 16	Quarterfinal	Semifinal	Final	Win
Brazil	89.9	61.2	42.0	26.6	16.3
Germany	89.1	60.4	41.6	26.1	15.8
Spain	85.9	60.4	37.4	21.9	12.6
France	87.0	56.5	36.3	21.3	12.3
Argentina	78.7	48.6	28.7	15.7	8.3
Belgium	81.7	53.6	27.5	14.8	7.4
England	75.6	46.4	22.0	10.8	4.9
Portugal	66.3	38.1	18.3	8.3	3.5
Uruguay	68.1	32.1	14.8	6.4	2.6
Croatia	58.7	29.2	14.2	6.3	2.6
Colombia	64.6	30.9	13.0	5.7	2.2
Russia	64.2	28.9	12.8	5.3	2.1
Poland	57.9	25.8	10.1	4.1	1.5
Denmark	46.7	18.9	7.6	2.8	0.9
Mexico	45.2	17.4	7.4	2.7	0.9
Switzerland	45.4	17.3	7.3	2.7	0.9
Sweden	44.5	16.1	5.9	2.0	0.6
Egypt	39.3	14.2	5.7	2.0	0.6
Serbia	39.0	14.6	5.4	1.8	0.6
Senegal	37.9	13.3	5.2	1.7	0.5
Peru	31.7	12.0	4.5	1.5	0.4
Nigeria	41.2	15.3	5.0	1.7	0.5
Iceland	30.9	11.5	4.3	1.4	0.4
Japan	36.3	12.7	3.9	1.2	0.3
Australia	25.2	9.8	3.0	0.9	0.2
Morocco	27.3	8.8	2.8	0.8	0.2
Costa Rica	22.6	8.4	2.5	0.7	0.2
South Korea	26.8	8.1	2.8	0.8	0.2
Iran	26.5	8.1	2.7	0.8	0.2
Tunisia	23.5	8.6	2.3	0.6	0.1
Saudi Arabia	19.2	6.6	1.6	0.4	0.1
Panama	23.2	6.2	1.7	0.4	0.1

Table 2: Simulated probability for each team to "survive" in the 2018 FIFA World Cup, i.e., proceed from the group phase to the round of 16, quarter- and semifinals, the final and to win the tournament.

	BRA	GER	ESP	FRA	ARG	BEL	ENG	POR
bwin	5.0	5.50	7.0	7.5	10.0	13.0	19.0	23
bet365	5.0	5.50	7.0	7.5	10.0	12.0	19.0	26
Sky Bet	5.5	5.50	7.0	7.0	10.0	11.0	17.0	26
Ladbrokes	5.5	5.50	7.0	7.0	10.0	11.0	15.0	26
William Hill	5.5	5.50	7.0	6.5	10.0	12.0	17.0	26
Marathon Bet	5.0	5.50	6.5	6.5	10.0	12.0	17.0	26
Betfair Sportsbook	5.5	5.50	6.5	7.0	10.0	11.0	17.0	23
SunBets	5.5	5.50	6.0	6.5	10.0	11.0	17.0	23
Paddy Power	5.5	5.50	6.5	7.0	10.0	11.0	17.0	23
Unibet	5.0	5.80	7.0	8.0	10.0	12.0	18.0	23
Coral	5.5	5.50	7.0	7.0	10.0	11.0	15.0	26
Betfred	5.5	5.50	7.5	6.5	10.0	11.0	17.0	26
Boylesports	5.5	5.50	7.0	7.0	10.0	12.0	17.0	26
Black Type	5.0	5.50	6.5	7.5	10.0	12.0	17.0	26
Betstars	5.0	5.75	7.5	7.5	10.0	11.0	15.0	23
Betway	5.5	5.50	7.0	7.5	11.0	12.0	17.0	26
BetBright	5.0	5.50	6.5	7.0	10.0	12.0	17.0	26
10Bet	5.0	5.10	7.0	7.0	10.0	12.0	17.0	26
Sportingbet	5.0	5.50	7.0	7.5	10.0	13.0	19.0	23
188Bet	5.6	5.60	7.2	7.5	11.0	13.0	19.0	26
888sport	5.0	5.80	7.0	8.0	10.0	12.0	18.0	23
Bet Victor	5.5	5.50	7.0	7.0	10.0	11.0	17.0	26
Sportpesa	5.0	5.10	7.0	7.0	10.0	12.0	17.0	26
Spreadex	5.5	6.00	7.5	8.0	12.0	13.0	21.0	29
Betdaq	5.5	5.60	7.0	7.6	11.8	12.8	18.6	29
Smarkets	5.5	5.60	7.2	7.6	11.8	12.8	18.0	28
	URU	CRO	COL	RUS	POL	DEN	MEX	SUI
bwin	34	34	41	41	67	101	126	101
bet365	34	34	41	41	51	101	101	101
Sky Bet	29	34	41	41	67	101	101	101
Ladbrokes	26	34	34	41	41	81	81	101
William Hill	26	29	34	51	51	101	101	101
Marathon Bet	34	34	34	41	41	81	101	101
Betfair Sportsbook	29	34	34	41	51	81	101	101
SunBets	29	34	34	41	41	101	81	81
Paddy Power	31	34	34	41	51	81	101	101
Unibet	34	31	41	33	71	101	101	101
Coral	29	34	34	34	34	81	67	101
Betfred	26	34	29	34	41	101	81	81
Boylesports	29	34	41	41	51	81	101	81
Black Type	29	34	34	41	51	81	101	81
Betstars	34	34	41	51	67	81	101	101
Betway	29	34	51	41	67	101	101	101
BetBright	29	34	34	34	51	81	81	101
10Bet	34	34	41	41	67	101	101	101
Sportingbet	34	34	41	41	67	101	126	101
188Bet	31	31	41	41	51	101	101	101
888sport	34	31	41	33	71	101	101	101
Bet Victor	34	34	41	41	67	101	101	101
Sportpesa	34	34	41	41	67	101	101	101
Spreadex	36	36	51	51	81	101	126	151
Betdaq	33	37	47	55	78	108	147	147
Smarkets	31	35	47	49	78	108	138	146

Table 3: Quoted odds from 26 online bookmakers (including two betting exchanges) for the 32 teams in the 2018 FIFA World Cup. Obtained on 2018-05-20 from https://www.oddschecker.com/ and https://www.bwin.com/, respectively.

	SWE	EGY	SRB	SEN	PER	NGA	ISL	JPN
bwin	151	151	201	151	151	201	201	301
bet365	151	151	201	201	201	201	201	301
Sky Bet	151	151	201	151	251	251	251	201
Ladbrokes	81	151	151	126	151	151	151	151
William Hill	151	151	151	151	201	151	251	251
Marathon Bet	126	201	151	151	201	201	201	201
Betfair Sportsbook	151	151	126	201	201	201	201	251
SunBets	101	151	151	151	151	151	201	251
Paddy Power	151	151	126	201	201	201	201	251
Unibet	101	151	101	151	151	201	201	301
Coral	101	126	126	151	201	201	126	201
Betfred	81	151	101	151	201	151	201	201
Boylesports	151	151	151	151	201	151	201	201
Black Type	101	151	151	151	201	201	201	251
Betstars	151	151	201	201	151	251	251	151
Betway	126	201	151	126	151	201	201	301
BetBright	125	151	151	126	201	151	151	201
10Bet	151	151	201	176	201	201	176	251
Sportingbet	151	151	201	151	151	201	201	301
188Bet	101	201	151	151	251	201	201	301
888sport	101	151	101	151	151	201	201	301
Bet Victor	151	151	201	201	251	201	251	201
Sportpesa	151	151	201	176	201	201	176	251
Spreadex	201	151	251	251	251	251	301	401
Betdaq	245	149	245	284	230	284	368	441
Smarkets	228	166	258	288	258	297	297	297
	AUS	MAR	CRC	KOR	IRN	TUN	KSA	PAN
bwin	301	401	401	501	501	751	501	1001
bet365	301	501	501	751	501	751	1001	1001
Sky Bet	501	251	751	251	751	1001	1001	1001
Ladbrokes	501	251	251	251	501	501	1001	1001
William Hill	501	301	301	401	501	751	1001	1001
Marathon Bet	401	401	301	401	501	501	1001	1001
Betfair Sportsbook	276	326	501	501	501	501	501	501
SunBets	301	501	401	501	501	501	1001	1001
Paddy Power	276	326	501	501	501	501	501	501
Unibet	401	301	451	601	451	451	1001	1001
Coral	501	251	251	251	501	501	1001	1001
Betfred	501	501	401	401	501	501	1001	1001
Boylesports	501	401	501	501	501	501	1001	1001
Black Type	251	501	501	501	501	501	1001	1001
Betstars	501	301	501	251	301	301	501	501
Betway	301	501	501	501	501	751	1001	1001
BetBright	251	251	251	251	501	501	1001	1001
10Bet	301	401	501	401	501	751	751	1001
Sportingbet	301	401	401	501	501	751	1001	1001
188Bet	501	501	501	751	501	751	1501	1501
888sport	401	301	451	601	451	451	1001	1001
Bet Victor	501	501	501	501	501	1001	2001	2001
Sportpesa	301	401	501	401	501	751	751	1001
Spreadex	501	751	751	751	751	1001	1001	1001
Betdaq	613	637	750	779	622	833	833	833
Smarkets	490	490	490	490	490	980	980	980

Table 4: Quoted odds from 26 online bookmakers (including two betting exchanges) for the 32 teams in the 2018 FIFA World Cup. Obtained on 2018-05-20 from https://www.oddschecker.com/ and https://www.bwin.com/, respectively.

Affiliation:

Achim Zeileis
Department of Statistics
Faculty of Economics and Statistics
Universität Innsbruck
Universitätsstr. 15
6020 Innsbruck, Austria
E-mail: Achim.Zeileis@R-project.org
URL: https://eeecon.uibk.ac.at/~zeileis/

Christoph Leitner, Kurt Hornik
Institute for Statistics and Mathematics
Department of Finance, Accounting and Statistics
WU Wirtschaftsuniversität Wien
Welthandelsplatz 1
1020 Wien, Austria
E-mail: Christoph.Leitner@wu.ac.at, Kurt.Hornik@wu.ac.at

University of Innsbruck - Working Papers in Economics and Statistics
Recent Papers can be accessed on the following webpage:

https://www.uibk.ac.at/eeecon/wopec/

2018-09 **Achim Zeileis, Christoph Leitner, Kurt Hornik:** Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model

2018-08 **Lisa Schlosser, Torsten Hothorn, Reto Stauffer, Achim Zeileis:** Distributional regression forests for probabilistic precipitation forecasting in complex terrain

2018-07 **Michael Kirchler, Florian Lindner, Utz Weitzel:** Delegated decision making and social competition in the finance industry

2018-06 **Manuel Gebetsberger, Reto Stauffer, Georg J. Mayr, Achim Zeileis:** Skewed logistic distribution for statistical temperature post-processing in mountainous areas

2018-05 **Reto Stauffer, Georg J. Mayr, Jakob W. Messner, Achim Zeileis:** Hourly probabilistic snow forecasts over complex terrain: A hybrid ensemble postprocessing approach

2018-04 **Utz Weitzel, Christoph Huber, Florian Lindner, Jürgen Huber, Julia Rose, Michael Kirchler:** Bubbles and financial professionals

2018-03 **Carolin Strobl, Julia Kopf, Raphael Hartmann, Achim Zeileis:** Anchor point selection: An approach for anchoring without anchor items

2018-02 **Michael Greinecker, Christopher Kah:** Pairwise stable matching in large economies

2018-01 **Max Breitenlechner, Johann Scharler:** How does monetary policy influence bank lending? Evidence from the market for banks' wholesale funding

2017-27 **Kenneth Harttgen, Stefan Lang, Johannes Seiler:** Selective mortality and undernutrition in low- and middle-income countries

2017-26 **Jun Honda, Roman Inderst:** Nonlinear incentives and advisor bias

2017-25 **Thorsten Simon, Peter Fabsic, Georg J. Mayr, Nikolaus Umlauf, Achim Zeileis:** Probabilistic forecasting of thunderstorms in the Eastern Alps

2017-24 **Florian Lindner:** Choking under pressure of top performers: Evidence from biathlon competitions

2017-23 **Manuel Gebetsberger, Jakob W. Messner, Georg J. Mayr, Achim Zeileis:** Estimation methods for non-homogeneous regression models: Minimum continuous ranked probability score vs. maximum likelihood

2017-22 **Sebastian J. Dietz, Philipp Kneringer, Georg J. Mayr, Achim Zeileis:** Forecasting low-visibility procedure states with tree-based statistical methods

2017-21 **Philipp Kneringer, Sebastian J. Dietz, Georg J. Mayr, Achim Zeileis:** Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season

2017-20 **Loukas Balafoutas, Brent J. Davis, Matthias Sutter:** How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior

2017-19 **Martin Geiger, Richard Hule:** The role of correlation in two-asset games: Some experimental evidence

2017-18 **Rudolf Kerschbamer, Daniel Neururer, Alexander Gruber:** Do the altruists lie less?

2017-17 **Meike Köhler, Nikolaus Umlauf, Sonja Greven:** Nonlinear association structures in flexible Bayesian additive joint models

2017-16 **Rudolf Kerschbamer, Daniel Muller:** Social preferences and political attitudes: An online experiment on a large heterogeneous sample

2017-15 **Kenneth Harttgen, Stefan Lang, Judith Santer, Johannes Seiler:** Modeling under-5 mortality through multilevel structured additive regression with varying coefficients for Asia and Sub-Saharan Africa

2017-14 **Christoph Eder, Martin Halla:** Economic origins of cultural norms: The case of animal husbandry and bastardy

2017-13 **Thomas Kneib, Nikolaus Umlauf:** A primer on bayesian distributional regression

2017-12 **Susanne Berger, Nathaniel Graham, Achim Zeileis:** Various versatile variances: An object-oriented implementation of clustered covariances in R

2017-11 **Natalia Danzer, Martin Halla, Nicole Schneeweis, Martina Zweimüller:** Parental leave, (in)formal childcare and long-term child outcomes

2017-10 **Daniel Muller, Sander Renes:** Fairness views and political preferences - Evidence from a large online experiment

2017-09 **Andreas Exenberger:** The logic of inequality extraction: An application to Gini and top incomes data

2017-08 **Sibylle Puntscher, Duc Tran Huy, Janette Walde, Ulrike Tappeiner, Gottfried Tappeiner:** The acceptance of a protected area and the benefits of sustainable tourism: In search of the weak link in their relationship

2017-07 **Helena Fornwagner:** Incentives to lose revisited: The NHL and its tournament incentives

2017-06 **Loukas Balafoutas, Simon Czermak, Marc Eulerich, Helena Fornwagner:** Incentives for dishonesty: An experimental study with internal auditors

2017-05 **Nikolaus Umlauf, Nadja Klein, Achim Zeileis:** BAMLSS: Bayesian additive models for location, scale and shape (and beyond)

2017-04 **Martin Halla, Susanne Pech, Martina Zweimüller:** The effect of statutory sick-pay on workers' labor supply and subsequent health

2017-03 **Franz Buscha, Daniel Müller, Lionel Page:** Can a common currency foster a shared social identity across different nations? The case of the Euro.

2017-02 **Daniel Müller:** The anatomy of distributional preferences with group identity

2017-01 **Wolfgang Frimmel, Martin Halla, Jörg Paetzold:** The intergenerational causal effect of tax evasion: Evidence from the commuter tax allowance in Austria

University of Innsbruck

Working Papers in Economics and Statistics

2018-09

Achim Zeileis, Christoph Leitner, Kurt Hornik

Probabilistic forecasts for the 2018 FIFA World Cup based on the bookmaker consensus model

Abstract
Football fans worldwide anticipate the 2018 FIFA World Cup that will take place in Russia from 14 June to 15 July 2018. 32 of the best teams from 5 confederations compete to determine the new World Champion. Using a consensus model based on quoted odds from 26 bookmakers and betting exchanges a probabilistic forecast for the outcome of the World Cup is obtained. The favorite is Brazil with a forecasted winning probability of 16.6%, closely followed by the defending World Champion and 2017 FIFA Confederations Cup winner Germany with a winning probability of 15.8%. Two other teams also have winning probabilities above 10%: Spain and France with 12.5% and 12.1%, respectively. The results from this bookmaker consensus model are coupled with simulations of the entire tournament to obtain implied abilities for each team. These allow to obtain pairwise probabilities for each possible game along with probabilities for each team to proceed to the various stages of the tournament. This shows that indeed the most likely final is a match of the top favorites Brazil and Germany (with a probability of 5.5%) where Brazil has the chance to compensate the dramatic semifinal in Belo Horizonte, four years ago. However, given that it comes to this final, the chances are almost even (50.6% for Brazil vs. 49.4% for Germany). The most likely semifinals are between the four top teams, i.e., with a probability of 9.4% Brazil and France meet in the first semifinal (with chances slightly in favor of Brazil in such a match, 53.5%) and with 9.2% Germany and Spain play the second semifinal (with chances slightly in favor of Germany with 53.1%). These probabilistic forecasts have been obtained by suitably averaging the quoted winning odds for all teams across bookmakers. More precisely, the odds are first adjusted for the bookmakers' profit margins ("overrounds"), averaged on the log-odds scale, and then transformed back to winning probabilities. Moreover, an "inverse" approach to simulating the tournament yields estimated team abilities (or strengths) from which probabilities for all possible pairwise matches can be derived. This technique (Leitner, Zeileis, and Hornik 2010a) correctly predicted the winner of 2010 FIFA World Cup (Leitner, Zeileis, and Hornik 2010b) and three out of four semifinalists at the 2014 FIFA World Cup (Zeileis, Leitner, and Hornik 2014). Interactive web graphics for this report are available at:
https://eeecon.uibk.ac.at/ zeileis/news/fifa2018/

ISSN 1993-4378 (Print)
ISSN 1993-6885 (Online)