Rethinking the FIFA World Cup^{TM} Final Draw

Julien Guyon

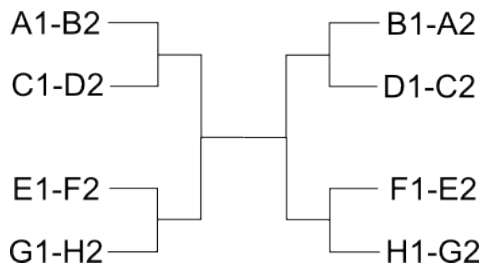
Bloomberg L.P., Quantitative Research
Columbia University, Department of Mathematics
NYU, Courant Institute of Mathematical Sciences

New England Symposium on Statistics in Sports
Harvard University, September 26, 2015

jguyon2@bloomberg.net, jg3601@columbia.edu, julien.guyon@nyu.edu

The FIFA World Cup^{TM}: Basic facts

- The most popular sporting event in the world.
- 32 senior men's national soccer teams.
- In 2014, 5 continents were represented: Europe (UEFA, 13 teams), South America (CONMEBOL, 6 teams), Africa (CAF, 5 teams), North and Central America (CONCACAF, 4 teams), and Asia (AFC, 4 teams).
- Group stage: the 32 finalists are divided into 8 groups of 4, labeled A through H. Each group plays a round-robin tournament, and the winner and runner-up advance to the knockout stage:



This talk is about how the 8 groups of the first stage are built, and how we believe they should be built.

Principles guiding the draw rules

Current building procedure indicates that FIFA is guided by 4 legitimate principles:

- Randomness: Teams placed into groups randomly.
- Tractability: Small number of bowls and balls + TV show of about one hour.
- Balance: Procedure should produce eight balanced groups.
- Geographic separation: Teams from the same continent cannot be drawn into the same group. Exception: European teams, since there are more than 8 of them—a maximum of 2 European teams per group is allowed.

The 4 pots of the final draw of the 2014 FIFA World Cup

Pot 1: seeded teams	Pot 2: S.A. & Africa	Pot 3: N.A. & Asia	Pot 4: Europe
11 Brazil (11)	**12** Chile (12)	**13** USA (13)	**8** Netherlands (8)
1 Spain (1)	**17** Côte d'Iv. (17)	**23** Mexico (24)	**9** Italy (9)
2 Germany (2)	**21** Ecuador (22)	**24** Costa Rica (31)	**10** England (10)
3 Argentina (3)	**22** Ghana (23)	**27** Honduras (34)	**14** Portugal (14)
4 Colombia (4)	**25** Algeria (32)	**28** Japan (44)	**15** Greece (15)
5 Belgium (5)	**26** Nigeria (33)	**29** Iran (49)	**16** Bosnia (16)
6 Uruguay (6)	**32** Cameroon (59)	**30** Korea Rep. (56)	**18** Croatia (18)
7 Switzerland (7)	1 team drawn from Pot 4	**31** Australia (57)	**19** Russia (19)
			20 France (21)

Lack of balance

Gr. A	**11** Brazil (11)	**32** Cameroon (59)	**23** Mexico (24)	**18** Croatia (18)
Gr. B	**1** Spain (1)	**12** Chile (12)	**31** Australia (57)	**8** Netherlands (8)
Gr. C	**4** Colombia (4)	**17** Côte d'Iv. (17)	**28** Japan (44)	**15** Greece (15)
Gr. D	**6** Uruguay (6)	**9** Italy (9)	**24** Costa Rica (31)	**10** England (10)
Gr. E	**7** Switzerland (7)	**21** Ecuador (22)	**27** Honduras (34)	**20** France (21)
Gr. F	**3** Argentina (3)	**26** Nigeria (33)	**29** Iran (49)	**16** Bosnia(16)
Gr. G	**2** Germany (2)	**22** Ghana (23)	**13** USA (13)	**14** Portugal (14)
Gr. H	**5** Belgium (5)	**25** Algeria (32)	**30** Korea Rep. (56)	**19** Russia (19)

Group	A	B	C	D	E	F	G	H	Range	Std dev
Sum of relative ranks **1–32**	**84**	**52**	**64**	**49**	**75**	**74**	**51**	**79**	**35**	**13.0**
Sum of FIFA rankings	112	78	80	56	84	101	52	112	60	21.6
Sum of 3 best relative ranks **1–32**	**52**	**21**	**36**	**25**	**48**	**47**	**29**	**49**	**31**	**11.4**
Sum of 3 best FIFA rankings	53	21	36	25	50	52	29	56	35	13.2

Other flaws of the current draw system

Lack of fairness: Some teams have a greater chance of ending up in a tough group than the rest.

- The high-ranked teams that are placed in pots together with low-ranked ones are more likely to end up in tough groups than they should.
- 2 teams particularly aggrieved last year: Chile and the United States.

Uneven distribution: All possible outcomes of the draw are not equally likely.

- Much better than in the past, see Jones (1990) and Rathgeber and Rathgeber (2007).
- But still imperfect: $\mathbb{P}(\text{Chile or Ecuador are placed into Group B})$ should have been $2/7 = 14/49$. It was actually $13/49$ (or $24/49$, depending on interpretation of draw rules)

Our suggested procedure

1. **Build pots by level**
2. **Add an S-curve constraint**
3. **Draw the continents first, then the teams**

	Pot 1	Pot 4	Pot 5	Pot 8
Draw I	*1* Brazil (11)	*16* Bosnia (16)	*17* Côte d'Iv. (17)	*32* Cameroon (59)
	2 Spain (1)	*15* Greece (15)	*18* Croatia (18)	*31* Australia (57)
	3 Germany (2)	*14* Portugal (14)	*19* Russia (19)	*30* Korea Rep. (56)
	4 Argentina (3)	*13* USA (13)	*20* France (21)	*29* Iran (49)
	Pot 2	Pot 3	Pot 6	Pot 7
Draw II	*5* Colombia (4)	*12* Chile (12)	*21* Ecuador (22)	*28* Japan (44)
	6 Belgium (5)	*11* England (10)	*22* Ghana (23)	*27* Honduras (34)
	7 Uruguay (6)	*10* Italy (9)	*23* Mexico (24)	*26* Nigeria (33)
	8 Switzerland (7)	*9* Netherlands (8)	*24* Costa Rica (31)	*25* Algeria (32)

The crucial question

How can we ensure that the geographic constraint is satisfied, in a tractable, evenly distributed way?

- The obvious evenly distributed rules are not tractable:
  - list all admissible outcomes and draw one uniformly
  - rejection method
- UEFA faces the same issue when they draw the groups of the Champions League™ but, as explained in Kloessner and Becker (2013) and Guyon (2014), the tractable procedure that they built is unevenly distributed.

Admissible continental distributions

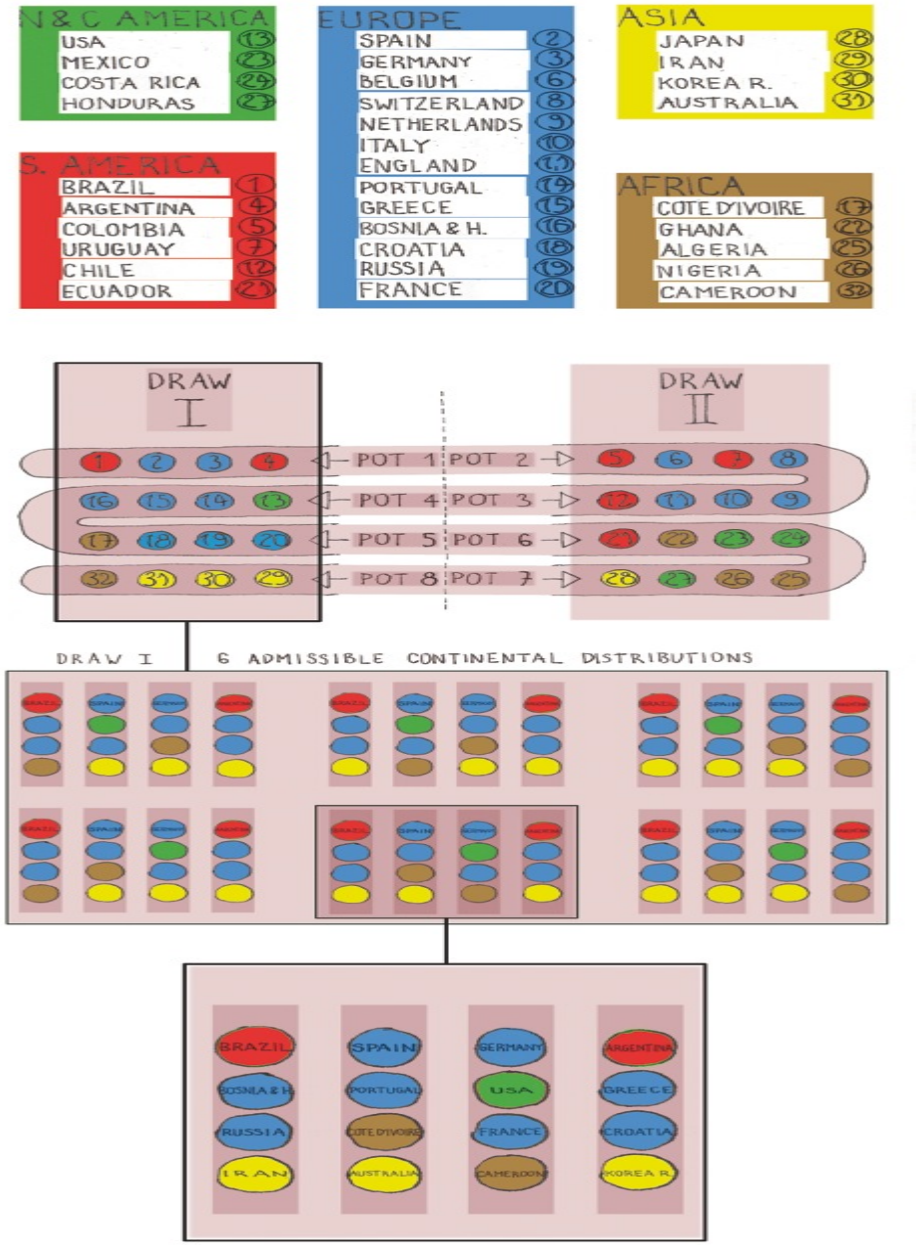
	Pot 1	**Pot 4**	**Pot 5**	**Pot 8**
Draw I	Brazil (11)	Europe	Europe	Asia
	Spain (1)	Europe	Africa	Asia
	Germany (2)	North America	Europe	Africa
	Argentina (3)	Europe	Europe	Asia
	Pot 2	**Pot 3**	**Pot 6**	**Pot 7**
Draw II	Colombia (4)	Europe	North America	Africa
	Belgium (5)	South America	Africa	Asia
	Uruguay (6)	Europe	North America	Africa
	Switzerland (7)	Europe	South America	North America

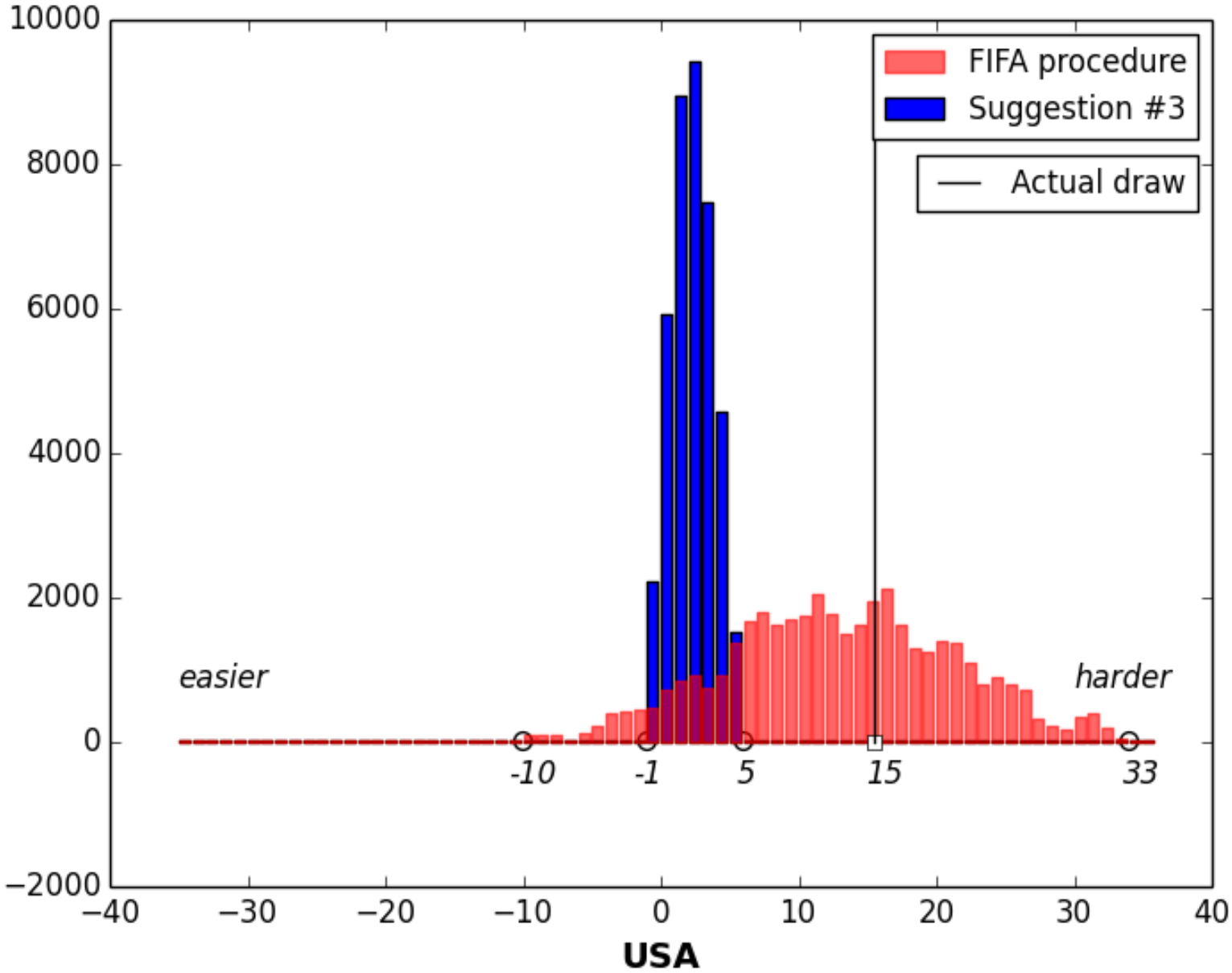
In 2014: Only $N_I = 6$ admissible continental distributions for Draw I, and $N_{II} = 24$ for Draw II.

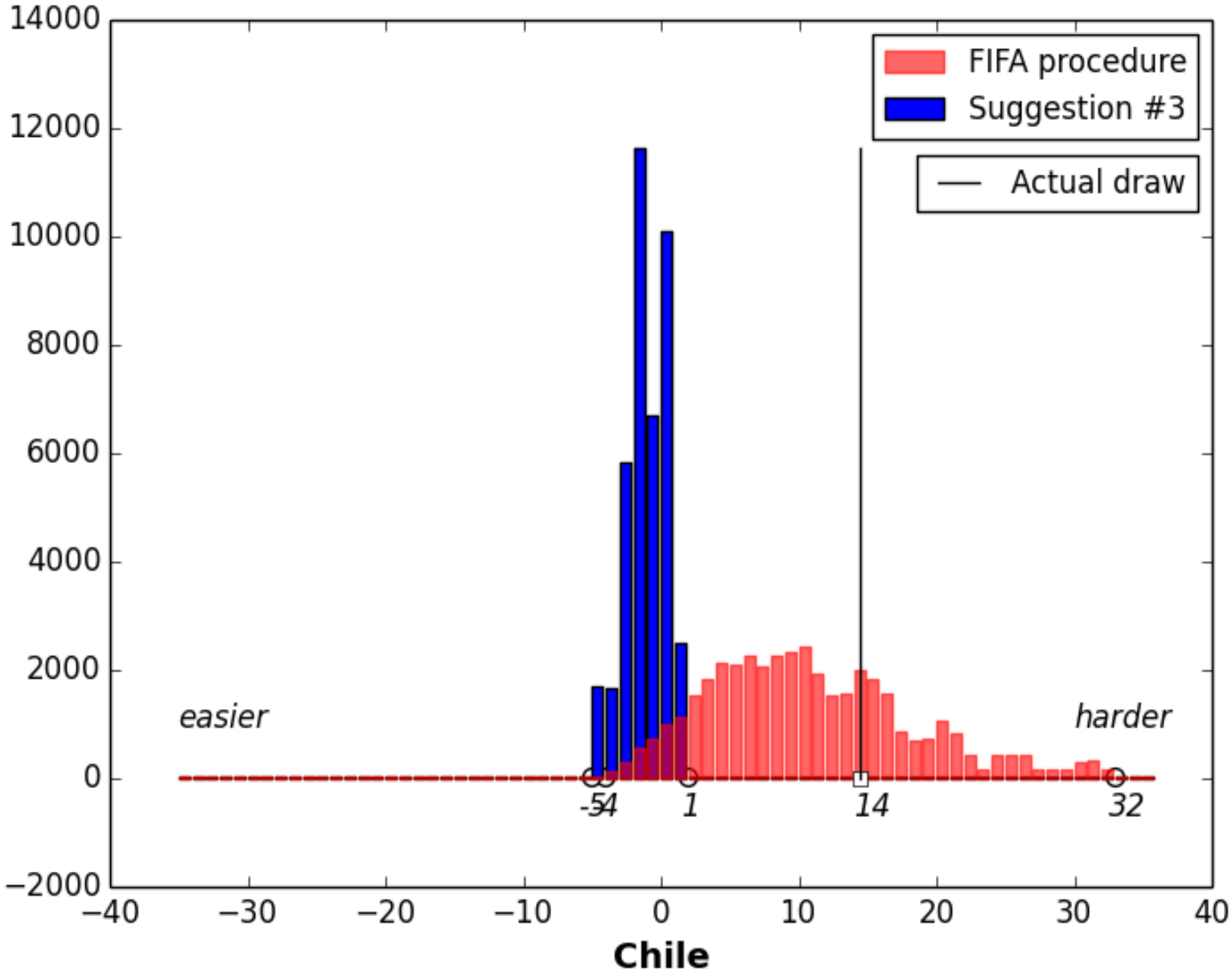
The procedure

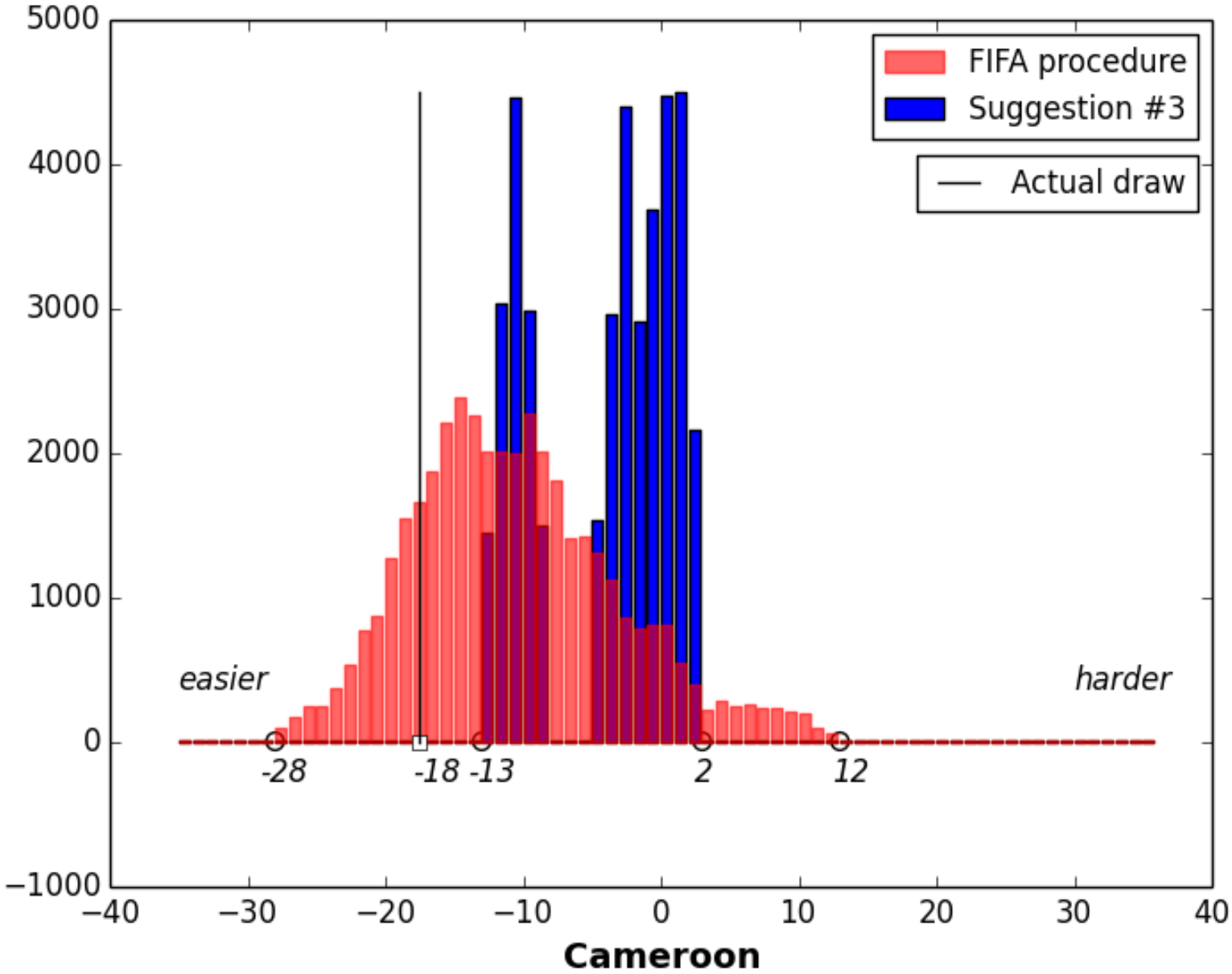
1. Before the draw, the exhaustive list of admissible continental distributions is established for both Draw I and Draw II, and numbered from 0 to $N_I - 1$ and 0 to $N_{II} - 1$ respectively.
2. The day of the draw, two numbers are drawn independently that follow the uniform distribution on the integers from 0 to $N_I - 1$, and 0 to $N_{II} - 1$ respectively, defining the admissible continental distributions
3. Then Pot 8 is emptied sequentially, randomly, and each team drawn goes to the first available position for its continent, from Row 1 (the host, Brazil) to Row 4 (Argentina).
4. The same procedure is repeated for Pots 7, 6, 5, 4 and 3.
5. Eventually, in order to determine the matches for the knockout stage, the host country is allocated to group A—a tradition that we do not question—while the seven remaining seeded teams are allocated randomly to groups B to H in a way that is consistent with the S-curve constraint.

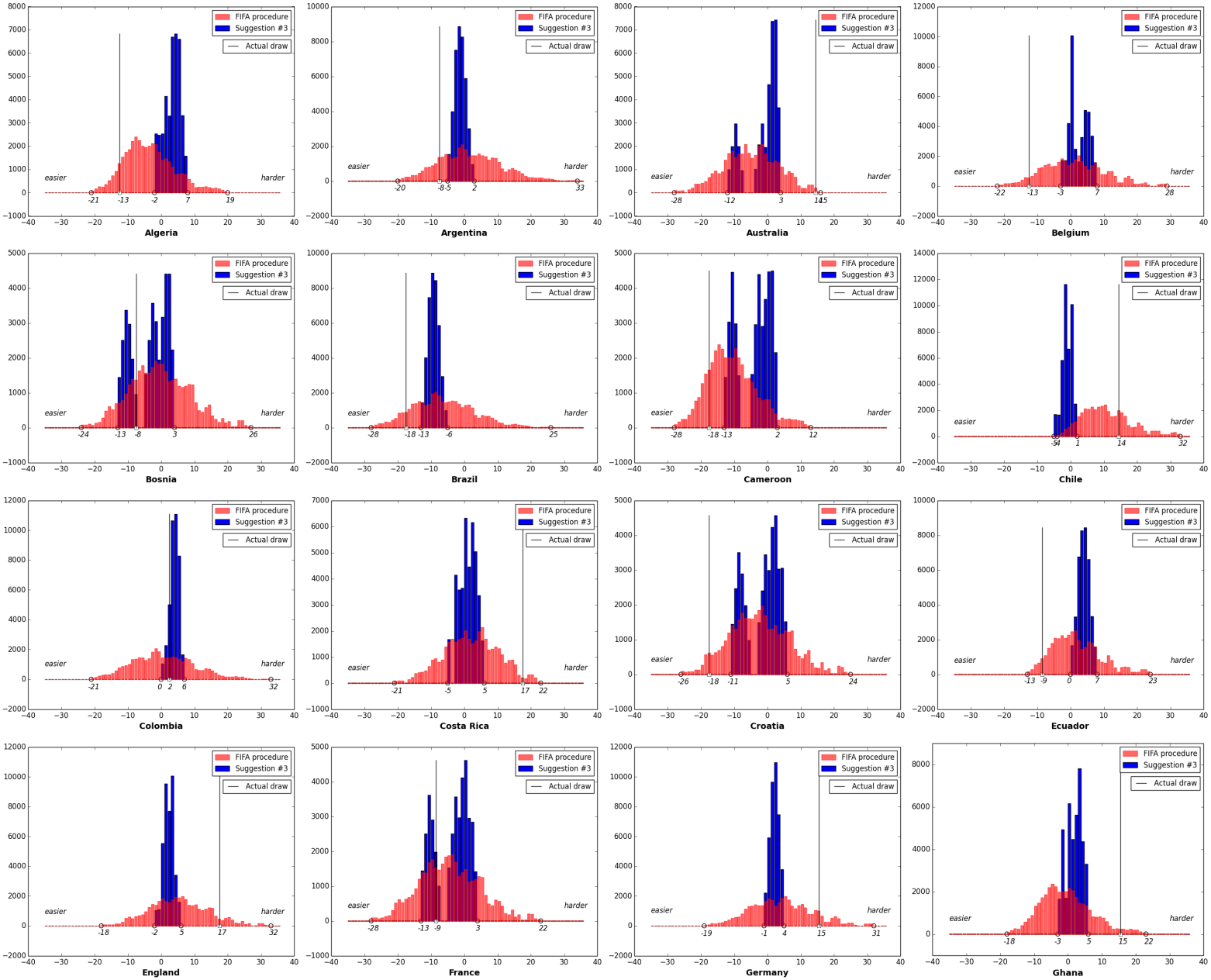
First time that a random procedure is suggested for the final draw of the FIFA World Cup that is tractable, produces balanced groups, and satisfies the geographic constraint. Moreover, it is fair to all teams, and evenly distributed.

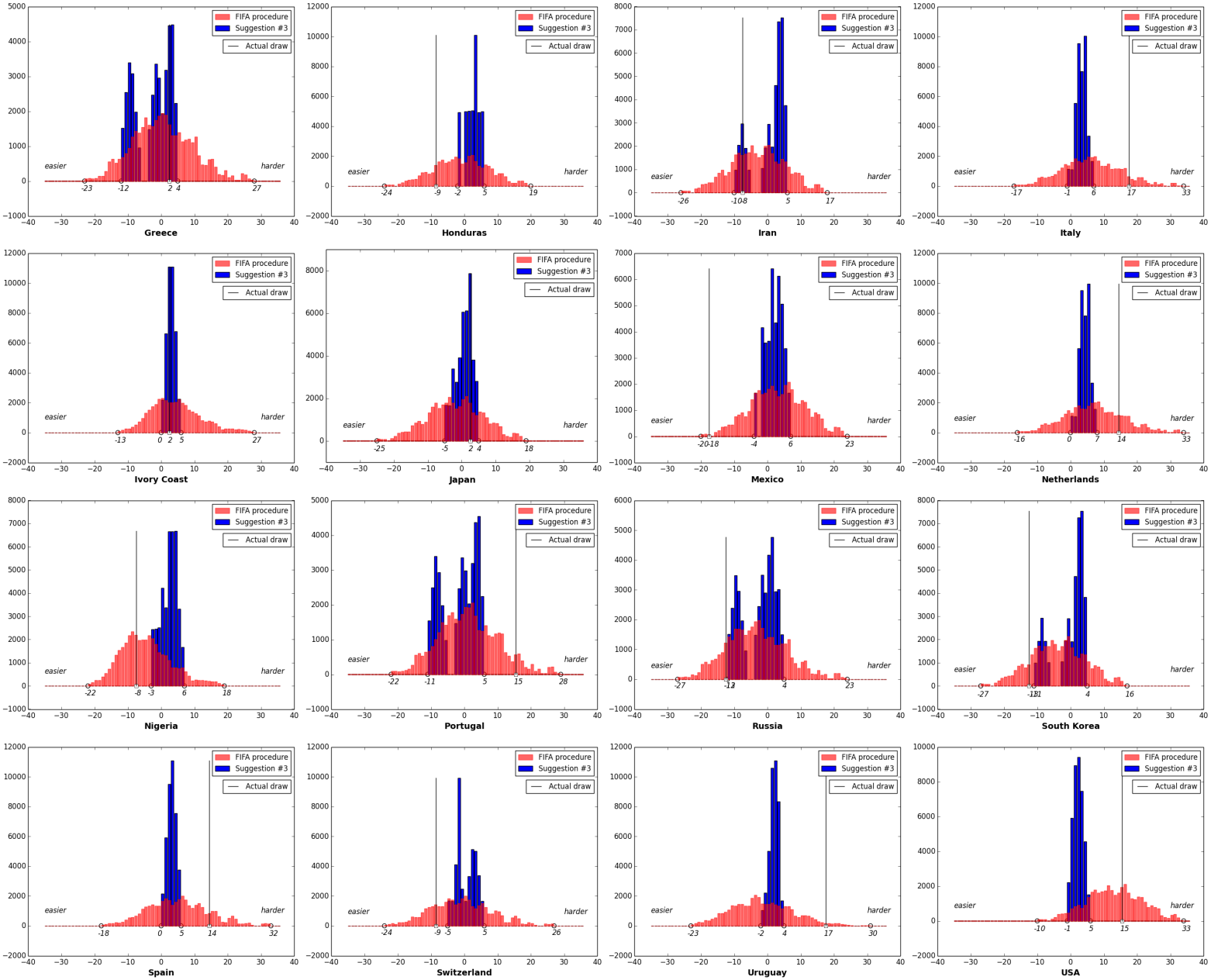

N&C AMERICA
USA 13
MEXICO 23
COSTA RICA 29
HONDURAS 27
S. AMERICA
BRAZIL 1
ARGENTINA 4
COLOMBIA 5
URUGUAY 7
CHILE 12
ECUADOR 21
EUROPE
SPAIN 2
GERMANY 3
BELGIUM 6
SWITZERLAND 8
NETHERLANDS 9
ITALY 10
ENGLAND 11
PORTUGAL 14
GREECE 15
BOSNIA & H. 16
CROATIA 18
RUSSIA 19
FRANCE 20
ASIA
JAPAN 28
IRAN 29
KOREA R. 30
AUSTRALIA 31
AFRICA
COTE D'IVOIRE 17
GHANA 22
ALGERIA 25
NIGERIA 26
CAMEROON 32
DRAW I
DRAW II
POT 1 POT 2
POT 4 POT 3
POT 5 POT 6
POT 8 POT 7
DRAW I 6 ADMISSIBLE CONTINENTAL DISTRIBUTIONS
BRAZIL
SPAIN
GERMANY
ARGENTINA
BOSNIA & H.
PORTUGAL
USA
GREECE
RUSSIA
COTE D'IVOIRE
FRANCE
CROATIA
IRAN
AUSTRALIA
CAMEROON
KOREA R.


FIFA procedure
Suggestion #3
Actual draw
easier
harder
-10
-1
5
15
33
USA
10000
8000
6000
4000
2000
0
−2000
−40
−30
−20
−10
0
10
20
30
40


FIFA procedure
Suggestion #3
Actual draw
easier
harder
-5
-4
1
14
32
14000
12000
10000
8000
6000
4000
2000
0
−2000
−40
−30
−20
−10
0
10
20
30
40
Chile


FIFA procedure
Suggestion #3
Actual draw
easier
harder
-28
-18
-13
2
12
5000
4000
3000
2000
1000
0
−1000
−40
−30
−20
−10
0
10
20
30
40
Cameroon


FIFA procedure
Suggestion #3
Actual draw
easier
harder
Algeria
Argentina
Australia
Belgium
Bosnia
Brazil
Cameroon
Chile
Colombia
Costa Rica
Croatia
Ecuador
England
France
Germany
Ghana


FIFA procedure
Suggestion #3
Actual draw
easier
harder
Greece
Honduras
Iran
Italy
Ivory Coast
Japan
Mexico
Netherlands
Nigeria
Portugal
Russia
South Korea
Spain
Switzerland
Uruguay
USA

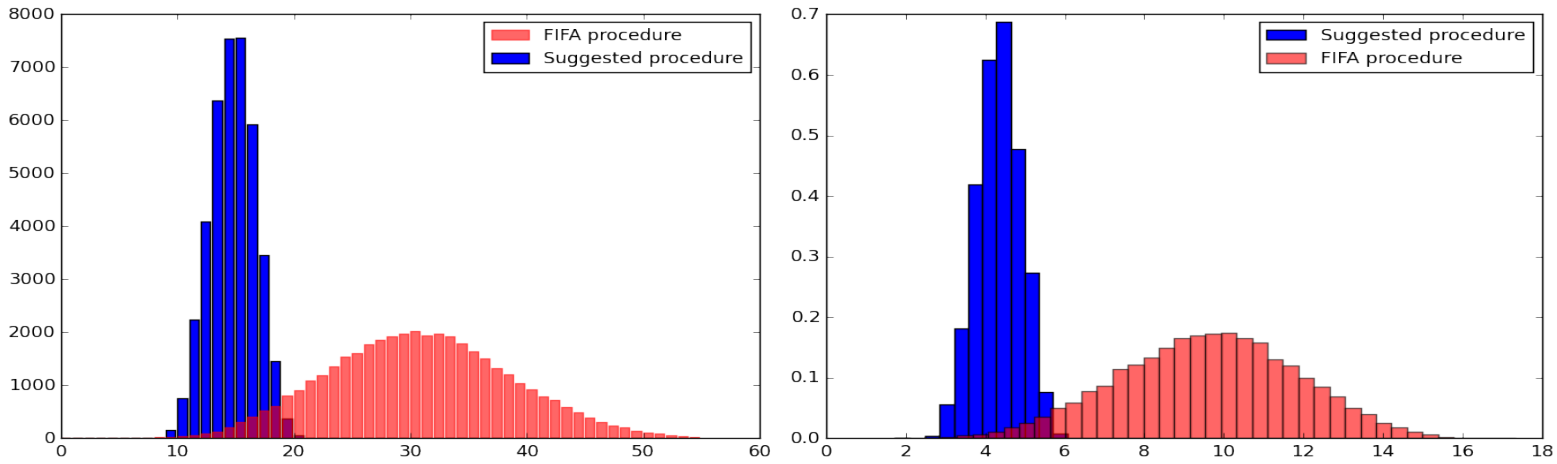


Figure : Distribution of the range (left) and standard deviation (right) of the eight sums of relative ranks, using FIFA rankings for seeding

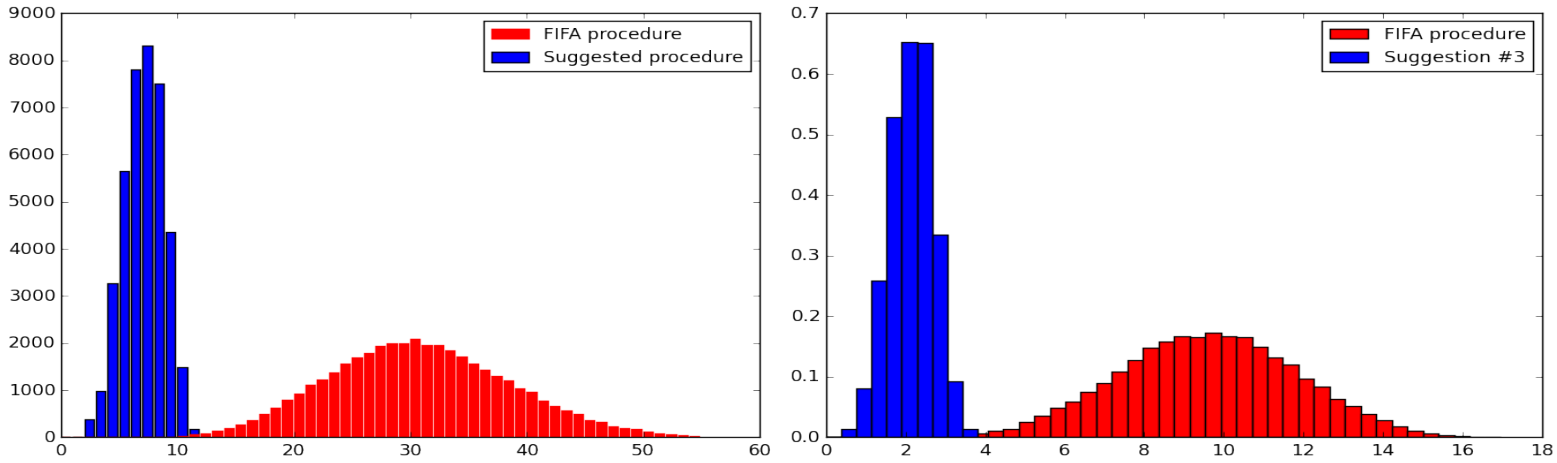
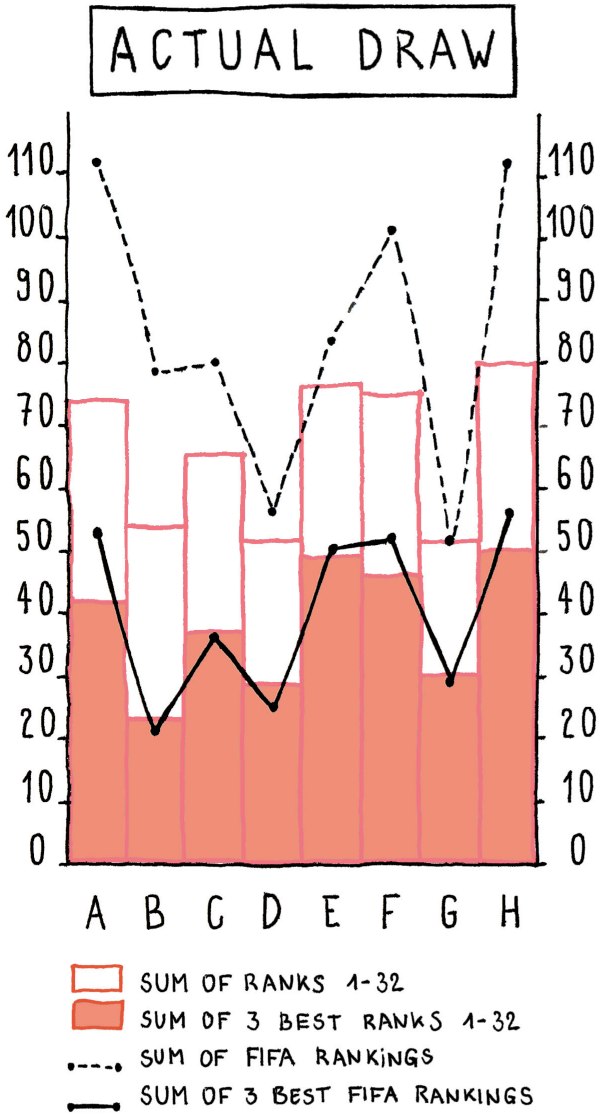
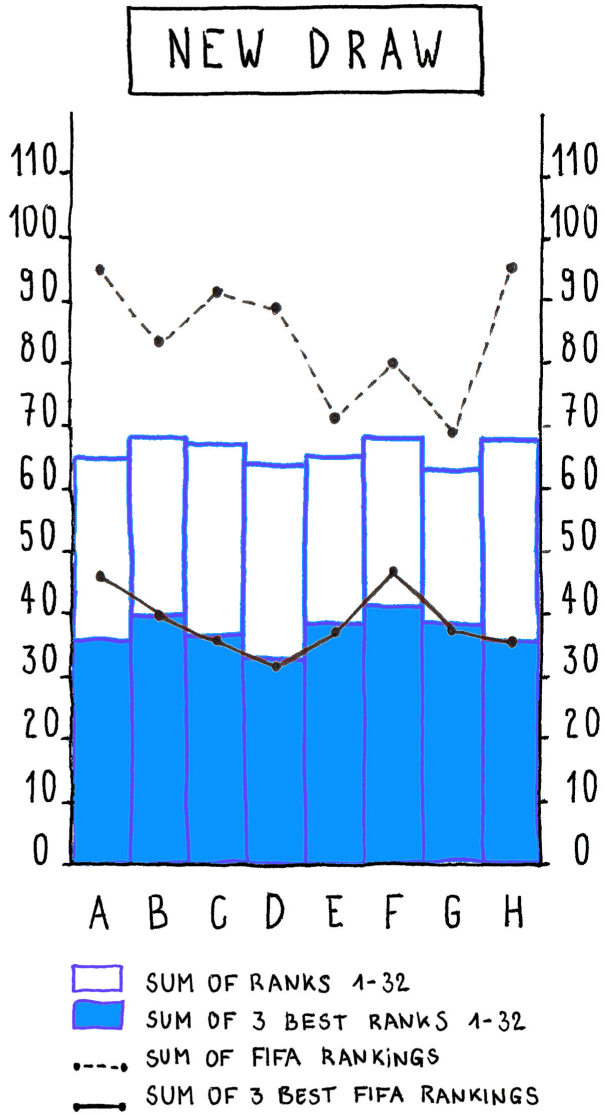
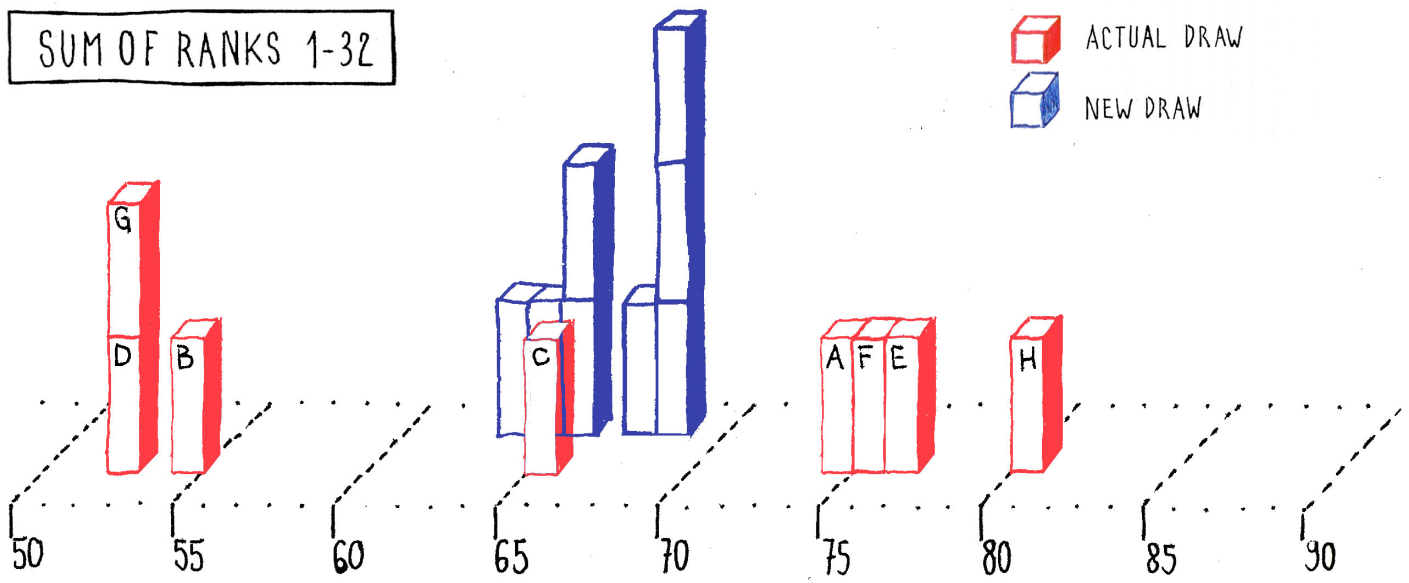
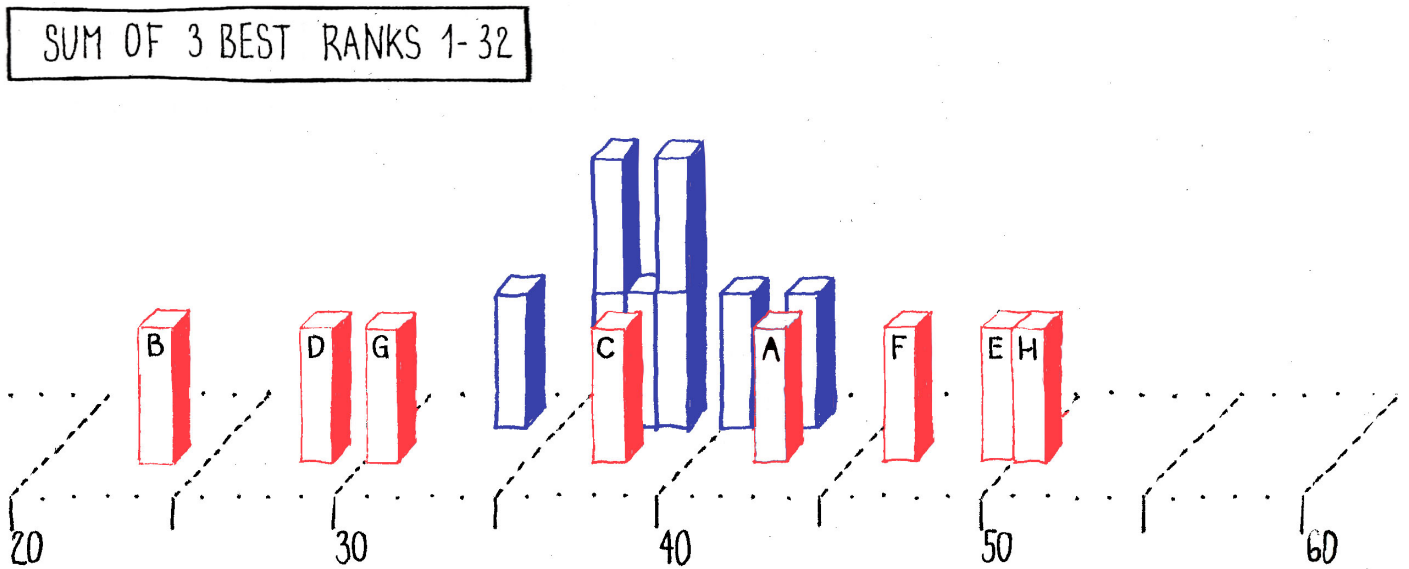


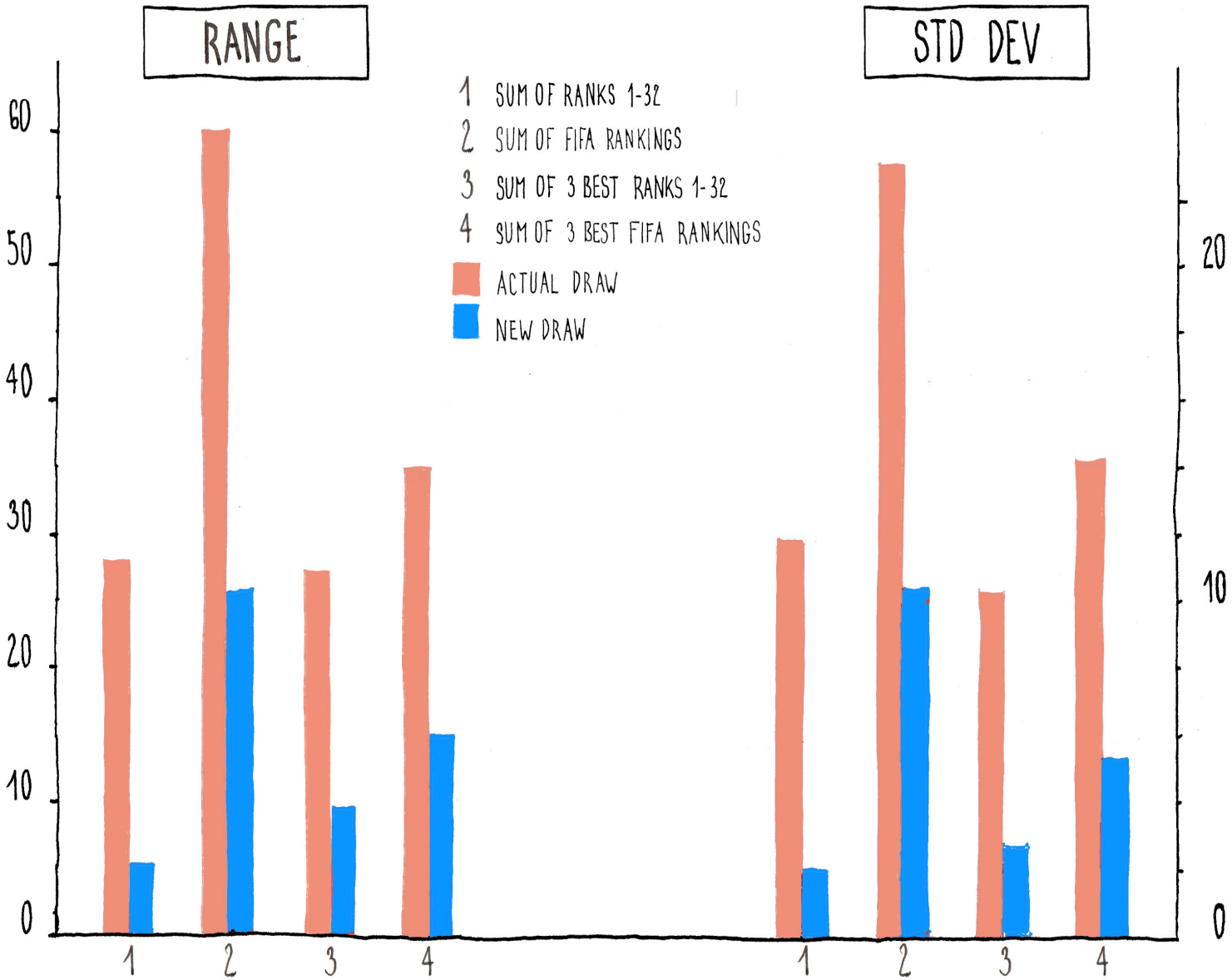
Figure : Using Elo ratings as of June 1, 2014 for seeding in our suggested procedure


ACTUAL DRAW
110
100
90
80
70
60
50
40
30
20
10
0
A B C D E F G H
SUM OF RANKS 1-32
SUM OF 3 BEST RANKS 1-32
SUM OF FIFA RANKINGS
SUM OF 3 BEST FIFA RANKINGS


NEW DRAW
110
100
90
80
70
60
50
40
30
20
10
0
A B C D E F G H
SUM OF RANKS 1-32
SUM OF 3 BEST RANKS 1-32
SUM OF FIFA RANKINGS
SUM OF 3 BEST FIFA RANKINGS


SUM OF RANKS 1-32
ACTUAL DRAW
NEW DRAW
G
D
B
C
A F E
H
50
55
60
65
70
75
80
85
90


SUM OF 3 BEST RANKS 1-32
B
D G
C
A
F
E H
20
30
40
50
60


RANGE
STD DEV
1 SUM OF RANKS 1-32
2 SUM OF FIFA RANKINGS
3 SUM OF 3 BEST RANKS 1-32
4 SUM OF 3 BEST FIFA RANKINGS
ACTUAL DRAW
NEW DRAW
60
50
40
30
20
10
0
20
10
0
1
2
3
4
1
2
3
4

Lucky/unlucky teams

- From our analysis, one can also quantify how lucky/unlucky a team was during the day of the draw.
- Luckiest team: Mexico. Unluckiest: Australia.
- Lucky teams: Algeria, Argentina, Belgium, Croatia, Ecuador, Russia, South Korea...
- Unlucky teams: Costa Rica, England, Germany, Ghana, Netherlands, Portugal, Uruguay...
- By looking at the p-value of their draw, one can rank teams from the luckiest to the unluckiest; see Aisch and Leonhardt (2014).

Number of admissible continental distributions since 1998

	1, 2 Europ. team per group		0, 1, 2 Europ. team per group	
	N_I	N_{II}	N_I	N_{II}
2014	6	24	6	24
2014 (Elo)	60	108	60	108
2010	252	24	428	24
2006 (OFC)	18	338	18	410
2006 (AFC)	18	110	18	126
2002	32	0	60	0
2002 (rebalanced)	48	48	48	48
1998	60	0	84	0
1998 (rebalanced)	108	9	108	9

Benefits of imposing the S-curve constraint:

1. From 315,360 admissible continental distributions to only 6 and 24!
2. Even more balanced groups

Conclusion

- Flaws of the current rules of the FIFA World Cup final draw: unbalanced groups; unfair to some teams; possible outcomes not equally likely. These flaws result from the way FIFA has decided to enforce the geographic constraints that they put on the draw.
- We suggest a new tractable procedure that produces eight random, balanced, and geographically diverse groups, is fair to all teams, and produces equally likely outcomes.
- Devising fair rules guarantees a fair competition, in which no team feels aggrieved.
- Group balance ensures that the teams that advance to the round of 16 owe it to their sporting merit, not to a lucky draw, and that the ones that fail to advance must blame themselves, not an unlucky draw $\Longrightarrow$ Better knockout stage, better tournament.
- The suggested procedure is actually very simple to put in place and lends itself to a nice and eagerly anticipated TV show.
- We believe that this is how the teams should be drawn, and we hope to convince both FIFA and football fans throughout the world.

Selected references


Aisch, G. and D. Leonhardt. June 5, 2014. "Mexico, the World Cup's Luckiest Country." *The New York Times*. nytimes.com/2014/06/06/upshot/mexicos-run-of-world-cup-luck-has-continued.html.


Guyon, J. 2014. "Rethinking the FIFA World Cup final draw." *To appear in Journ. Quant. Anal. Sports*. Long version: ssrn.com/abstract=2424376.


Guyon, J. June 4, 2014. "The World Cup Is Unfair. Here's A Better Way" *The New York Times*. nytimes.com/2014/06/05/upshot/the-world-cup-draw-is-unfair-heres-a-better-way.html.


Guyon, J. June 13, 2014. "A Better Way to Rank Soccer Teams in a Fairer World Cup." *The New York Times*. nytimes.com/2014/06/14/upshot/a-better-way-to-rank-soccer-teams-in-a-fairer-world-cup.html.


Jones, M.C. 1990. "The World Cup draw's flaws." *Mathematical Gazette*. 74(470):335–338.


Kloessner, S., and M. Becker. 2013. "Odd Odds: The UEFA Champions League™ Round of Sixteen Draw." *Journ. Quant. Anal. Sports*. 9(3):249–270.


Rathgeber, A. and H. Rathgeber. 2007. "Why Germany Was Supposed To Be Drawn in the Group of Death and Why It Escaped." *Chance*. 20(2):22–24.

2010

Pot 1	Pot 4	Pot 5	Pot 8
1 South Africa (85)	*16* Mexico (18)	*17* Côte d'Ivoire (19)	*32* North Korea (91)
2 Brazil (1)	*15* Chile (17)	*18* Serbia (20)	*31* New Zealand (83)
3 Spain (2)	*14* Greece (16)	*19* Paraguay (21)	*30* Slovenia (49)
4 Netherlands (3)	*13* Cameroon (14)	*20* Australia (24)	*29* South Korea (48)
Pot 2	**Pot 3**	**Pot 6**	**Pot 7**
5 Italy (4)	*12* Switzerland (13)	*21* Uruguay (25)	*28* Japan (40)
6 Germany (5)	*11* USA (11)	*22* Denmark (27)	*27* Ghana (38)
7 Argentina (6)	*10* Portugal (10)	*23* Algeria (29)	*26* Honduras (35)
8 England (7)	*9* France (9)	*24* Nigeria (32)	*25* Slovakia (33)

Table : Pots by level for the teams which qualified to the 2010 FIFA World Cup South Africa^{TM}. The number in brackets is the October 2009 FIFA ranking. The S-curve follows increasing FIFA rankings, except for the host country, which is protected and put in first position of Pot 1. The italicized number indicates the position in the S-curve, from *1* to *32*

2006

Pot 1	Pot 4	Pot 5	Pot 8
1 Germany	*16* Czech Rep.	*17* Portugal	*32* Togo
2 Brazil	*15* Paraguay	*18* Costa Rica	*31* Angola
3 England	*14* Croatia	*19* Saudi Arabia	*30* Ghana
4 Spain	*13* Sweden	*20* Poland	*29* Trinidad and Tobago
Pot 2	**Pot 3**	**Pot 6**	**Pot 7**
5 Mexico	*12* Japan	*21* Iran	*28* Australia
6 France	*11* South Korea	*22* Tunisia	*27* Côte d'Ivoire
7 Italy	*10* Netherlands	*23* Ecuador	*26* Ukraine
8 Argentina	*9* USA	*24* Serbia & Montenegro	*25* Switzerland

Table : Pots by level for the teams which qualified to the 2006 FIFA World Cup Germany^{TM}. The italicized number indicates the position in the S-curve, from *1* to *32*

2002

Pot 1	Pot 4	Pot 5	Pot 8	Pot 8 (rebalanced)
1 South Korea	*16* Belgium	*17* Portugal	*32* Senegal	*32* Senegal
2 Japan	*15* USA	*18* Ireland	*31* China	*31* China
3 Brazil	*14* Sweden	*19* Russia	*30* Ecuador	*27* **Slovenia**
4 Argentina	*13* Paraguay	*20* Nigeria	*29* Costa Rica	*28* **Poland**
Pot 2	**Pot 3**	**Pot 6**	**Pot 7**	**Pot 7 (rebalanced)**
5 Italy	*12* Denmark	*21* Saudi Arabia	*28* Poland	*29* **Costa Rica**
6 Germany	*11* Croatia	*22* South Africa	*27* Slovenia	*30* **Ecuador**
7 France	*10* England	*23* Tunisia	*26* Uruguay	*26* Uruguay
8 Spain	*9* Mexico	*24* Cameroon	*25* Turkey	*25* Turkey

Table : Pots by level for the teams which qualified to the 2002 FIFA World Cup Korea/Japan^{TM}. Note that the lower part of the S-curve has 10 European teams (out of 15). Right: Pots 7 and 8 after using the S-curve rebalancing algorithm. The italicized number indicates the initial position in the S-curve, from *1* to *32*

1998

Pot 1	Pot 4	Pot 5	Pot 8	Pot 8 (rebalanced)
1 France	*16* Norway	*17* Morocco	*32* Iran	*32* Iran
2 Germany	*15* Denmark	*18* Cameroon	*31* Jamaica	*31* Jamaica
3 Brazil	*14* USA	*19* Nigeria	*30* South Afr.	*27* **Austria**
4 Italy	*13* Colombia	*20* Saudi Arabia	*29* Paraguay	*28* **Croatia**
Pot 2	**Pot 3**	**Pot 6**	**Pot 7**	**Pot 7 (rebalanced)**
5 Spain	*12* Belgium	*21* Yugoslavia	*28* Croatia	*29* **Paraguay**
6 Argentina	*11* England	*22* South Korea	*27* Austria	*30* **South Afr.**
7 Romania	*10* Bulgaria	*23* Scotland	*26* Chile	*26* Chile
8 Netherlands	*9* Mexico	*24* Japan	*25* Tunisia	*25* Tunisia

Table : Pots by level for the teams which qualified to the 1998 FIFA World Cup France^{TM}. Note that the lower part of the S-curve has 10 European teams (out of 15). Right: Pots 7 and 8 after using the S-curve rebalancing algorithm. The italicized number indicates the initial position in the S-curve, from *1* to *32*