On Any Given Sunday: May the better team have better odds

All of us have probably thought sometime during the football season that there must be a Rasch analysis in here somewhere. Every team in the National Football League plays a different selection of opponents but rankings are based, mostly[1], on a simple count of games won. We certainly know better than that. Here's my answer.

Table II.4: National Football League Ratings for 2012 Regular Season

Conference	Division	Team	Logit	Division	Conference
AFC	East	New England	0.53	-0.40	-2.44
		Miami	-0.16		
		NY Jets	-0.38		
		Buffalo	-0.39		
	North	Cincinnati	0.22	0.06	
		Baltimore	0.11		
		Pittsburgh	-0.02		
		Cleveland	-0.24		
	South	Houston	0.23	-1.15	
		Indianapolis	-0.21		
		Tennessee	-0.40		
		Jacksonville	-0.77		
	West	Denver	0.52	-0.96	
		San Diego	-0.24		
		Oakland	-0.51		
		Kansas City	-0.74		
NFC	East	NY Giants	0.29	0.00	2.44
		Washington	0.09		
		Dallas	0.02		
		Philadelphia	-0.40		
	North	Chicago	0.36	0.82	
		Green Bay	0.30		
		Minnesota	0.25		
		Detroit	-0.09		
	South	Atlanta	0.52	0.62	
		New Orleans	0.06		
		Carolina	0.03		
		Tampa Bay	0.00		
	West	San Francisco	0.61	1.00	
		Seattle	0.58		
		Arizona	-0.07		
		St. Louis	-0.12		

The readily available data consist of paired comparisons of points for and against, which can be arranged in a matrix with each team defining a row and a column. The row entries for each team are the points scored and the column entries are the points allowed. These data can be manipulated using reasonably straightforward arithmetic, to be disclosed later, to obtain a measure of each team's proficiency based on the teams they actually competed against and how they have fared. Teams earn high proficiency measures if they

[1]There are more rules for resolving ties in the rankings but these are designed more to create excitement and sell tickets than to ensure that the best team wins.

play quality opponents, score many points of their own, and allow few points by the opponents. Duh.

Evidence of the method's validity includes second-hand reports that it won several appetizers at a local bar by beating the in-house expert and that it finished second overall in a neighborhood pool.

Table II.4 and Figure II.2 show an analysis of the 2012 National Football League regular season. The teams are sorted by their Rasch proficiency rating within division within conference, which is not quite the same order as the order based simply on wins and loses. The *official* division winners are highlighted in yellow and wild card[2] teams in green. It is somewhat embarrassing that two teams (NY Giants and Chicago Bears) that were the highest rated in their divisions did not qualify for the post-season playoffs but I'm not sure if I should be embarrassed, if the Bears and Giants should embarrassed, or if the League that set the schedules and playoff rules should be.

This is another demonstration of the meaninglessness of the labels we give measures; measures have no meaning until we give them meaning. The Las Vegas odds makers are remarkably unimpressed by the information that Baltimore had a proficiency measure of 0.11 *logits*. Until we know a lot more about them, we have no way of knowing what logits are good for, what's a big number, or what values we would like to see. To begin to get a feel for what a big logit might be, the average logit in this analysis is 0.00; the highest in the table is 0.61 (San Francisco); and the lowest is -0.77 (Jacksonville). Anyone who paid any attention is pretty sure that San Francisco would defeat Jacksonville, even without the benefit of the logit measures. The logits do provide a method to quantify what *pretty sure* means.

The logit for a team is mathematically the *log odds* that the team will defeat the average team (which just happens to be Tampa Bay that year.) To get the odds, we need the *antilog* of the logit[3]. In this case, the odds that San Francisco ($logit = 0.61$) will defeat Tampa Bay ($logit = 0.0$) is $e^{0.61} = 1.84$ to 1 (or 11 to 6 or 11 times in 17 games, if you can only think in integers.)

We can generalize this expression and say that the odds that team A will defeat team B is the difference in the logits as a power of e. To get back to the question of what we mean by *pretty sure*, the odds that San Francisco will defeat Jacksonville are:

7. $Odds(SF\ over\ Jacksonville) = e^{[0.61 - (-0.77)]} = e^{1.38} = 3.97\ to\ 1.$

The worst team in the NFL should beat the best team once in every five games, which is why they play the game. On any given Sunday, any team can beat any other team, or so they say.

The odds that Baltimore would defeat San Francisco when they met in Super Bowl XLVII is done the same way, substituting Baltimore's logit:

8. $Odds(Balt\ over\ SF) = e^{(0.11 - 0.61)} = e^{-0.5} = 0.61\ to\ 1$

---

[2] The two teams with the best won/loss records but that did not win their divisions advanced to the post-season as "Wild Cards."

[3] Everything we do with log and antilog will use base e, so-called Naperian or natural logs. If $log_e\ x = y$, then x is the antilog of y and $x = e^y$. For most spreadsheet programs, this is the *exp()* function.

Or, about 3 to 5 in round numbers. This means that when two teams differ by half a logit, the weaker team will win three times in eight games and the stronger, five times. The actual Super Bowl XLVII played on the field was one of the three that Baltimore wins.

If two teams differ by one logit (e.g., Minnesota with a 0.25 logit and Kansas City with a -0.74), the odds will be:

9. $Odds(MN\ over\ KC) = e^{[0.25-(-0.74)]} = e^{0.99} = 2.7\ to\ 1$

A difference of one logit means Minnesota should win nearly three games out of four. If you prefer, odds can be changed to probability by dividing one odd by the sum of the two odds, i.e.,

10. $Prob(SF\ over\ TB) = \dfrac{e^{0.61}}{e^{0.61}+e^{0.0}} = \dfrac{1.84}{1.84+1} = 0.65.$

11. $Prob(MN\ over\ KC) = \dfrac{e^{0.25}}{e^{0.25}+e^{-0.74}} = \dfrac{2.7}{2.7+1} = 0.73.$

12. $Prob(SF\ over\ Jkv) = \dfrac{e^{0.61}}{e^{061}+e^{-0.77}} = \dfrac{3.98}{3.98+1} = 0.80.$

While all this has been fun things you can do with logits in your own home, I sincerely doubt if anything in this discussion has added much meaning to the logit scale for much of my audience. Truly understanding what a logit is comes from experience using them; this may entail watching a lot of football. Logits are very convenient for doing the algebra and arithmetic but not terribly attractive for parading before the public.

The values used on the left side of Figure II.2 are logits and, as labels, are arbitrary; we can replace them with any labels we like as long as we preserve the interval scale property. This just means that things that are equally spaced in logits must be equally spaced in the new metric. Legal transformations take the form *A + B*logit,* where *A* and *B* are up for grabs. We can transform the logits by multiplying by something and adding something. We are free to choose the two somethings.

Table II-5: Transforming Logits: Y = A + B*logit

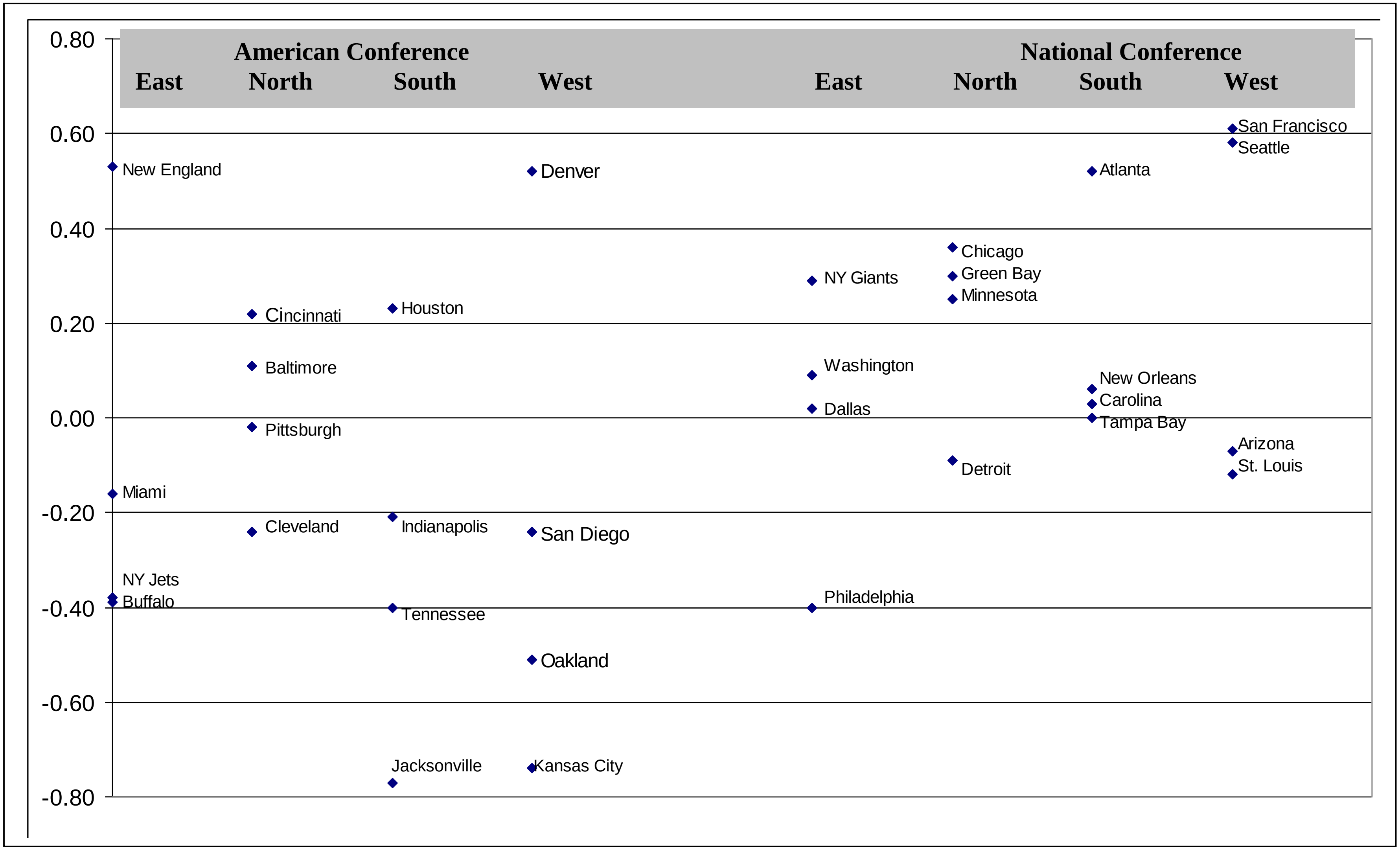
B=>	**5**	**50**	**40**	**100**	**200**	
A=>	**4**	**50**	**100**	**250**	**500**	
Logit	Line	Findex	FQ	Fit	FAT	Team
1.0	9	100	140	350	700	
0.8	8	90	132	330	660	
0.6	7	80	124	310	620	49ers
0.4	6	70	116	290	580	
0.2	5	60	108	270	540	Ravens
0.0	4	50	100	250	500	Bucs
-0.2	3	40	92	230	460	
-0.4	2	30	84	210	420	
-0.6	1	20	76	190	380	
-0.8	0	10	68	170	340	Jaguars

Table II.5 gives some arbitrary choices for *A* and *B* and the resulting scales. They are all equally correct. The first transformation just re-labels the horizontal lines in Figure II.2

from zero to eight. The second looks like, but isn’t, percent correct; the third like IQs, but isn’t; the fourth like WITs, but isn’t; the last like SAT, but isn’t. Use one of these or pick your own *A* and *B*; it’s your choice.

Replacing the labels on the left side of Figure II.2 with any other legal set won’t change the picture in any way any more than changing to team nicknames in the last column of Table II.5 changed anything. But using nicer labels might make some of your audience less hostile and San Francisco should still have won the Super Bowl and Baltimore shouldn’t even have been in it.

Figure II-2:


0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
American Conference
East
North
South
West
National Conference
East
North
South
West
New England
Denver
San Francisco
Seattle
Atlanta
Chicago
Green Bay
Minnesota
NY Giants
Cincinnati
Houston
Baltimore
Washington
New Orleans
Carolina
Tampa Bay
Dallas
Pittsburgh
Arizona
St. Louis
Detroit
Miami
Cleveland
Indianapolis
San Diego
NY Jets
Buffalo
Tennessee
Philadelphia
Oakland
Jacksonville
Kansas City