Is there persistence in sequences of consecutive football results?

Stephen Dobson & John Goddard

A question of enduring fascination to sports fans concerns the nature of persistence in sequences of consecutive match results. Does a sequence of wins tend to build a team's confidence and morale, increasing the probability that the next match will also be won? Or does it tend to create pressures or breed complacency, increasing the likelihood that the next match will be drawn or lost? Does a sequence of losses tend to sap confidence or morale, increasing the probability of a further loss in the next match? Or does it tend to inspire greater effort, increasing the likelihood that the next match will be won or drawn? These questions are examined using English league match results data for seasons 1969-70 to 2008-09 (inclusive).

Table 1 reports the longest sequences of consecutive results in the league match results data set between the 1969-70 and 2008-09 seasons inclusive, based on four criteria: (i) matches without a win, (ii) matches without a loss, (iii) consecutive wins, and (iv) consecutive losses. Each of these criteria is applied to home matches, away matches, and to all matches (home and away). For the purposes of counting sequences of consecutive results, breaks between seasons are ignored.

Table 2 reports empirical unconditional and conditional match result probabilities, where conditioning is on the duration of various preceding sequences of consecutive similar results. The first row reports the unconditional home and away probabilities for a win (columns (1) and (2)), a win or draw ((3) and (4)), a loss ((5) and (6)) and a loss or draw ((7) and (8)). Of the 81,258 matches in the data set, 38,775 were home wins, 22,426 were draws, and 20,057 were away wins. Therefore 0.477 = 38775/81258 is the unconditional home win probability (column (1)), and so on.

Subsequent rows of Table 2 report the conditional probabilities that each result represents a 'reversal' of a previous sequence of consecutive 'identical' results, conditioning on the type and duration of the sequence. Two types of reversal are considered. First, a sequence of wins and draws is reversed by a loss; and a sequence of losses is reversed by a win or draw. Second, a sequence of wins is reversed by a draw or loss; and a sequence of draws and losses is reversed by a win. These two types of reversal are 'WD|L reversals' and 'W|DL reversals', respectively. Columns (3) to (6) of Table 2 show the conditional probabilities for WD|L reversals; and columns (1), (2), (7) and (8) show the conditional probabilities for W|DL reversals. For the purpose of calculating these probabilities, the results of any cup, European or friendly matches played within a sequence of league matches are ignored. So too are the venues (home or away) of the matches comprising the sequence of prior matches. The conditional probabilities themselves, however, are specific to the venue of match in question.

For example, the conditioning for the home win probabilities in column (1) is on the number of previous consecutive matches without a win (the number of previous matches drawn or lost). The home team had failed to win its four most recent matches in 15,260 of the 81,258 matches in the data set. This figure includes cases in which the sequence without a win was longer than four matches. In 6,610 of these 15,260 matches, the match result was a home win, implying a W|DL reversal. Therefore 0.433 = 6610/15260 is the home win probability conditional on the home team having played at least four consecutive matches without a win, prior to the match in question.[1]

Table 2 shows that the conditional probabilities of a good result (however defined) tend to decline with the duration of an unsuccessful spell, and the conditional probabilities of a poor result decline with the duration of a successful spell. Without further investigation, however, it would be incorrect to attribute this pattern to a positive persistence effect. The pattern in the conditional probabilities might be explained by the variation between teams in their relative quality or playing strength. The calculation of the win probability conditional on a long spell without a win, for example, is based mainly on the experience of weaker teams, whose win probability is below average because they are weak, but not specifically because they have not won recently. In other words, the pattern in the conditional probabilities reported in Table 2 might be explained by a team heterogeneity effect. Therefore any test for persistence in sequences of match results needs to control for heterogeneous team strengths.

Below, a Monte Carlo analysis is used to test for persistence effects. In the absence of any persistence effects, it is assumed that the following statistical model would accurately represent the distribution of match results in each season in each tier. According to this model, the result of the match between home team i and away team j is generated as follows:

Home win (k=2)	if	$\mu_2 < y^*_{i,j} + \varepsilon_{i,j}$	
Draw (k=1)	if	$\mu_1 < y^*_{i,j} + \varepsilon_{i,j} < \mu_2$	
Away win (k=0)	if	$y^*_{i,j} + \varepsilon_{i,j} < \mu_1$	[1]

where $y^*_{i,j} = \alpha_i - \alpha_j$; α_i and α_j are parameters reflecting the quality or playing strengths of team i and team j; μ_1 and μ_2 are additional parameters, known as 'cut-off parameters'; and $\varepsilon_{i,j} \sim N(0,1)$ is a random disturbance term, which follows a standard Normal distribution (with zero mean and variance of one). The disturbance term represents the unsystematic or random element in the result of the match between teams i and j.

Table 3 illustrates the correspondence between the final league table and the estimated parameters of [1] for the Premiership in the 2008-09 season. The ordering of the teams reflects the final league table, obtained by awarding three league points for a win and one for a draw. Table 3 also reports each team's win percent, obtained by awarding 1 for each win and 0.5 for each draw, and dividing the total by 38 (the number of matches played by each team). Table 3 also reports each team's $\hat{\alpha}_i$ for the 2008-09 season, with the parameter for the bottom-placed team, West Bromwich Albion, set to zero.

The estimates of μ_1 and μ_2 are shown at the foot of Table 3 together with illustrative fitted match result probabilities. The latter are calculated using:

$$P(\text{home win}) = 1 - \Phi(\hat{\mu}_2 - \hat{y}^*_{i,j})$$

$$P(\text{draw}) = \Phi(\hat{\mu}_2 - \hat{y}^*_{i,j}) - \Phi(\hat{\mu}_1 - \hat{y}^*_{i,j})$$

$$P(\text{away win}) = \Phi(\hat{\mu}_1 - \hat{y}^*_{i,j}) \qquad [2]$$

where Φ is the distribution function for the standard Normal distribution; and $\hat{y}^*_{i,j} = \hat{\alpha}_i - \hat{\alpha}_j$. The values of the cut-off parameters $\hat{\mu}_1$ and $\hat{\mu}_2$ allow for home-field advantage. The examples illustrate the implications of variations in $\hat{\alpha}_i$ and $\hat{\alpha}_j$ for the home win, draw and away win probabilities.

The Monte Carlo simulations enable comparisons to be drawn between the observed numbers of reversals in the data set (as defined above), and the numbers of reversals that should be obtained if [1] is the statistical model that describes correctly the distribution of match results if there is no persistence effect. Two test statistics are used to test for persistence effects: the first is based on the number of WD|L reversals; and the second is based on the number of W|DL reversals. In each case, the test statistic is τ = total number of match results divided by total number of reversals. If the observed value of τ is similar to its expected value obtained from the Monte Carlo simulations, the null hypothesis of no persistence cannot be rejected. If the observed τ is significantly higher than its expected value, reversals occur less frequently than they should occur if the null hypothesis is true. In this case the null hypothesis is rejected in favour of an alternative hypothesis of positive persistence. Conversely, if the observed τ is significantly lower than its expected value, reversals occur more frequently than they should when the null hypothesis is true. In this case the null hypothesis is rejected in favour of an alternative hypothesis of negative persistence.

The approach is similar in principle to the well-known runs test (Mood, 1940), which investigates the randomness of sequences of positive or negative increments to a time series. Mood used analytic methods to derive asymptotic sampling distributions for the numbers of 'runs' (sequences of consecutive positive or negative increments) expected under the randomness assumption, based on the binomial distribution. In the present case, simulation rather than analytic methods are used, because the expected numbers of reversals depend on the degree of inequality in the team strength parameters within each tier, and therefore vary between tiers and between seasons. The greater is the degree of inequality, the easier it is for a strong team to sustain a sequence of good results (and the harder it is for a weak team to break a sequence of poor results), and so the smaller is the expected number of reversals. Fort and Rosenman (1999) apply the runs test directly to sequences of match results in MLB to test for the presence of 'streaks'. Since there are no draws, match results are binary and the runs test is directly applicable. Each team is tested separately. The test controls for the quality of the team being tested (via the overall win percentage), but in contrast to the present analysis there is no control for variation in the quality of the opposition team.

To generate the expected mean durations of sequences of consecutive results under the null hypothesis of zero persistence, 120 sets of ordered probit estimates of the parameters of [1] are obtained. Using the actual fixture calendars as originally completed, a computer program then generates a complete set of simulated match results for the full 40-season period, under the assumption of zero persistence, by substituting randomly drawn values of $\varepsilon_{i,j} \sim N(0,1)$ into [1]. This exercise is repeated 5,000 times, in order to generate 5,000 sets of simulated match results each of which covers the entire 40-season period.

Table 4 reports match result probabilities conditional on each result representing a reversal of a previous sequence of consecutive results, calculated from the Monte Carlo simulations based on an assumption of no persistence. A comparison between these simulated conditional probabilities and the observed conditional probabilities reported in Table 2 (and allowing for occasional random variation in the latter) confirms that the actual probability of a reversal occurring is higher than the simulated probability under assumptions of no persistence.

In order to test the null hypothesis that there is no persistence effect, for each of the 5,000 sets of simulated match results, the test statistic τ (=number of matches ÷ number of reversals) is calculated for each of the two types of reversal. By examining the sampling distributions of the two sets of 5,000 simulated τ, critical values are established, leading to the acceptance or rejection of the null hypothesis of no persistence.

The persistence tests are carried out using the data for all 40 seasons from 1969-70 to 2008-09, and using the same data subdivided into eight sub periods of five seasons each: seasons 1970-1974, 1975-1979 and so on through to 2005-2009. Table 5 reports the results of these tests. The upper panel shows the results for WD|L reversals, and the lower panel shows the results for W|DL reversals. The columns headed p0.5, p2.5, p5.0, p95.0, p97.5, p99.5 show the 0.5, 2.5, 5, 95, 97.5 and 99.5 percentiles of the sampling distributions of the test statistic τ under the null hypothesis of no persistence, obtained from the Monte Carlo simulations.[2]

Accordingly, a 95 per cent confidence interval for τ under the null hypothesis of no persistence, based on the results for WD|L reversals, is given by (2.197, 2.215). The null hypothesis is rejected at a significance level of 5 per cent if τ falls outside this range. Similarly, a 99 per cent confidence interval for τ is given by (2.194, 2.217). The null hypothesis is rejected at a significance level of 1 per cent if τ falls outside this range.

The final two columns of Table 5 report the actual values of τ, and the corresponding p-values. The p-value is the minimum significance level at which the null hypothesis of no persistence can be rejected. In the results based on computations over the entire 40-season period, for WD|L reversals, τ=2.197 falls just inside the lower bound of the 95 per cent confidence interval, but outside the lower bound of the 90 per cent confidence interval (p-value=.0556). The null hypothesis of no persistence cannot be rejected at the 5 per cent significance level; but the null is rejected at the 10 per cent level. For W|DL reversals, however, τ=2.182 falls outside the lower bounds of both the 95 per cent and the 99 per cent confidence intervals (p-value=.0000). In this case the null hypothesis of zero persistence is rejected, in favour of an alternative of negative persistence, at any significance level. W|DL reversals occur more frequently than is expected if the null hypothesis is true.

In the results based on computations for five-season sub periods, the pattern is similar. For WD|L reversals, the null hypothesis of no persistence is not rejected at the 5 per cent level for any of the eight five-year sub periods. The null hypothesis is rejected at the 10 per cent level for two of the eight periods. For W|DL reversals, in contrast, the null hypothesis of no persistence is not rejected at the 5 per cent level for four of the eight five-year sub periods. The null hypothesis is rejected at the 10 per cent level for seven of the eight sub periods. In all cases where the null is rejected, τ is lower than is expected if the null hypothesis is true. The number of reversals is therefore higher than is expected if the null hypothesis is true.

Overall the results indicate that sequences of match results are subject to statistically significant, negative persistence effects. On average, sequences of consecutive wins and sequences of consecutive matches without a win tend to end sooner than they would if there were no statistical association between the results of consecutive matches after controlling for heterogeneous team strengths. There is, however, an element of asymmetry in the pattern. The average duration of sequences of matches unbeaten is higher than the average duration of sequences without a win; and the average duration of sequences of losses is higher than the average duration of sequences of wins. Accordingly, the evidence of a negative persistence effect in the data on W|DL reversals is stronger than it is in the data on WD|L reversals.

Finally, the procedure described above is valid if the assumption of no variation in the team strength parameters within each season is correct. A difficulty arises, however, if this assumption is incorrect, because the actual and expected numbers of reversals are sensitive to these parameters. If there is within-season variation in the team strength parameters α_I in [1], the expected numbers of reversals are somewhat lower than in the case where there is no variation, and the persistence test described above is biased towards detection of a positive persistence effect. This suggests that rejection of the null hypothesis of no persistence in favour of an alternative of negative persistence is a particularly strong result. If there is within-season variation in the team strength parameters, this test tends to be biased in the opposite direction.

References

Fort, R. and Rosenman, R. (1999). Winning and management for streaks. Proceedings of the Joint Statistical Meetings: Section on Statistics in Sports. Alexandria, VA: American Statistical Association.

Mood, A. (1940). The distribution theory of runs. *Annals of Mathematical Statistics* 11, 367-392.

Notes

1. The conditional probabilities in the other columns of the upper panel of Table 2 are calculated in the same way. The probabilities are not reported in cases where there were fewer than 50 sequences of the required duration on which to base the calculation. This limit is breached for sequences of consecutive wins and consecutive losses (columns (3), (4), (7) and (8)) of duration greater than (about) eight matches.

2. In the simulations for WD|L reversals for the period 1969-70 to 2008-09, for example, the first row of Table 5 shows that 0.5 per cent of the simulated τ were below 2.194 (and 99.5 per cent of the simulated τ were above 2.194); 2.5 per cent of the simulated τ were below 2.197 (and 97.5 per cent were above); and so on. At the opposite end of the range of values for τ, 97.5 per cent of the simulated τ were below 2.215 (and 2.5 per cent were above); 99.5 per cent of the simulated τ were below 2.217 (and 0.5 per cent were above); and so on.

Table 1: Longest runs of consecutive results, 1969-70 to 2008-09

Matches unbeaten		End-month	Consecutive wins		End-month
Arsenal	49	Oct-04	Arsenal	14	Aug-02
Nottm Forest	42	Nov-78	Newcastle Utd	13	Oct-92
Chelsea	40	Oct-05	Reading	13	Oct-85
Reading	33	Feb-06	Charlton Athletic	12	Mar-00
Bristol Rovers	32	Jan-74	Fulham	12	Oct-00
Liverpool	31	Mar-88	Liverpool	12	Oct-90
Arsenal	30	Oct-02	Luton Town	12	Apr-02
Leeds Utd	30	Feb-74	Manchester Utd	12	Aug-00

Consecutive defeats			Matches without a win		
Sunderland	17	Aug-03	Derby County	36	Aug-08
Walsall	15	Feb-89	Cambridge Utd	31	Apr-84
Brighton & Hove Albion	12	Jan-73	Hull City	27	Nov-89
Brighton & Hove Albion	12	Oct-02	Oxford Utd	27	Aug-88
Carlisle Utd	12	Dec-03	Newport County	25	Jan-71
Barnet	11	Oct-93	Rochdale	25	Aug-74
MK Dons	11	Mar-04			
Stoke City	11	Aug-85			
West Bromwich Albion	11	Dec-95			

Table 2 Empirical unconditional and conditional match result probabilities

n	Probability of a win, conditional on n = number of previous consecutive matches without a win		Probability of a win or draw, conditional on n = number of previous consecutive losses		Probability of a loss, conditional on n = number of previous consecutive matches without a loss		Probability of a loss or draw, conditional on n = number of previous consecutive wins	
	Home (1)	Away (2)	Home (3)	Away (4)	Home (5)	Away (6)	Home (7)	Away (8)
0	0.477	0.247	0.753	0.523	0.247	0.477	0.523	0.753
1	0.465	0.236	0.731	0.494	0.232	0.464	0.496	0.738
2	0.451	0.228	0.715	0.474	0.222	0.449	0.478	0.716
3	0.438	0.219	0.695	0.456	0.210	0.438	0.460	0.693
4	0.433	0.216	0.686	0.442	0.199	0.423	0.451	0.663
5	0.422	0.212	0.670	0.415	0.191	0.406	0.424	0.646
7	0.409	0.202	0.611	0.413	0.168	0.387	0.382	0.583
10	0.396	0.190	-	-	0.132	0.344	-	-
15	0.344	0.177	-	-	0.114	0.303	-	-
20	0.346	0.163	-	-	0.086	0.294	-	-

Table 3 Premiership table 2008-09 season and ordered probit team quality parameter estimates

	Won	Drawn	Lost	League points	Win ratio	$\hat{\alpha}_i$
Manchester United	28	6	4	90	.8158	1.6534
Liverpool	25	11	2	86	.8026	1.4863
Chelsea	25	8	5	83	.7632	1.3943
Arsenal	20	12	6	72	.6842	1.1348
Everton	17	12	9	63	.6053	.8941
Aston Villa	17	11	10	62	.5921	.8372
Fulham	14	11	13	53	.5132	.6379
Tottenham Hotspur	14	9	15	51	.4868	.5685
West Ham United	14	9	15	51	.4868	.5622
Manchester City	15	5	18	50	.4605	.4771
Wigan Athletic	12	9	17	45	.4342	.3757
Stoke City	12	9	17	45	.4342	.3944
Bolton Wanderers	11	8	19	41	.3947	.2242
Portsmouth	10	11	17	41	.4079	.3004
Blackburn Rovers	10	11	17	41	.4079	.2744
Sunderland	9	9	20	36	.3553	.1556
Hull City	8	11	19	35	.3553	.1808
Newcastle United	7	13	18	34	.3553	.1809
Middlesbrough	7	11	20	32	.3289	.0685
West Bromwich Albion	8	8	22	32	.3158	0

Cut-off parameters: $\hat{\mu}_1 = -.7119$ $\hat{\mu}_2 = .0250$

Illustrative fitted match result probabilities:

	Home win	Draw	Away win
Liverpool v Middlesbrough	0.673	0.209	0.118
Middlesbrough v Liverpool	0.309	0.285	0.406
Aston Villa v Blackburn Rovers	0.563	0.252	0.185
Blackburn Rovers v Aston Villa	0.417	0.284	0.299
Manchester City v Wigan Athletic	0.501	0.269	0.230
Wigan Athletic v Manchester City	0.480	0.274	0.246
All matches (average)	0.455	0.255	0.290

Table 4 Simulated unconditional and conditional match result probabilities

n	Probability of a win, conditional on n = number of previous consecutive matches without a win		Probability of a win or draw, conditional on n = number of previous consecutive losses		Probability of a loss, conditional on n = number of previous consecutive matches without a loss		Probability of a loss or draw, conditional on n = number of previous consecutive wins	
	Home (1)	Away (2)	Home (3)	Away (4)	Home (5)	Away (6)	Home (7)	Away (8)
0	0.477	0.247	0.753	0.523	0.247	0.477	0.523	0.753
1	0.461	0.230	0.731	0.489	0.231	0.462	0.488	0.730
2	0.444	0.217	0.707	0.462	0.218	0.444	0.455	0.701
3	0.428	0.206	0.684	0.439	0.206	0.427	0.425	0.671
4	0.415	0.197	0.663	0.417	0.195	0.411	0.397	0.644
5	0.403	0.189	0.642	0.396	0.185	0.396	0.373	0.618
7	0.382	0.175	0.598	0.357	0.167	0.368	0.333	0.573
10	0.356	0.158	0.537	0.303	0.145	0.333	0.288	0.518
15	0.317	0.135	-	-	0.120	0.290	-	-
20	0.284	0.116	-	-	0.102	0.258	-	-

Table 5 Tests for persistence in sequences of consecutive match results

	Monte Carlo simulations						Actual	
	p0.5	p2.5	p5.0	p95.0	p97.5	p99.5	τ	p-value
Sequences without a loss or sequences of losses: ratio of matches played to WD/L reversals								
1970-2009	2.194	2.197	2.198	2.213	2.215	2.217	2.197	.0556
1970-1974	2.191	2.198	2.202	2.245	2.250	2.258	2.228	.7284
1975-1979	2.147	2.155	2.159	2.201	2.205	2.213	2.184	.7480
1980-1984	2.135	2.142	2.146	2.187	2.191	2.198	2.153	.3152
1985-1989	2.160	2.168	2.171	2.213	2.217	2.225	2.203	.3988
1990-1994	2.163	2.171	2.175	2.216	2.220	2.228	2.173	.0764
1995-1999	2.187	2.196	2.200	2.242	2.246	2.255	2.205	.2216
2000-2004	2.198	2.205	2.208	2.252	2.257	2.265	2.216	.2964
2005-2009	2.207	2.215	2.219	2.262	2.266	2.275	2.215	.0528
Sequences of wins or sequences without a win: ratio of matches played to W/DL reversals								
1970-2009	2.199	2.203	2.204	2.219	2.220	2.223	2.182	.0000
1970-1974	2.193	2.201	2.205	2.248	2.251	2.258	2.200	.0428
1975-1979	2.151	2.158	2.162	2.202	2.206	2.215	2.151	.0104
1980-1984	2.138	2.147	2.151	2.192	2.196	2.205	2.149	.0836
1985-1989	2.162	2.170	2.174	2.216	2.221	2.228	2.175	.1112
1990-1994	2.166	2.174	2.178	2.220	2.225	2.232	2.178	.0948
1995-1999	2.192	2.201	2.206	2.249	2.253	2.262	2.165	.0000
2000-2004	2.204	2.213	2.217	2.261	2.266	2.274	2.211	.0368
2005-2009	2.217	2.227	2.230	2.276	2.281	2.289	2.229	.0792

1970-2009=1969-70 to 2008-09 seasons; 1970-1974=1969-70 to 1973-74 seasons, and so on.