Measuring Parity: Tying Into the Idealized Standard Deviation

Article *in* Journal of Sports Economics · August 2006

2 authors:

David D. Haddock
Independent Scholar

Louis P. Cain
Northwestern University

Measuring Parity

Tying Into the Idealized Standard Deviation

LOUIS P. CAIN
Loyola University and Northwestern University

DAVID D. HADDOCK
Northwestern University

We present a metric to calculate the idealized standard deviation (ISD) in the event that ties are possible and receive points in team-sport league standings. Our aim is to raise some issues regarding the customary application of that concept in leagues that do not employ a binomial system for determining rankings, to suggest that the reciprocal of the ISD may be a more intuitive construct if the intent is to represent parity, to provide scholars with a set of ISDs, and to suggest some interesting topics that arise from a quick perusal of the time series.

Keywords: *idealized standard deviation; parity; Major League Baseball; English Football League*

Beginning in 1934, the first professional football game each year pitted the defending National Football League (NFL) champion against a team of all-stars who had played for collegiate teams the prior season. In the early years that game was closely contested and widely followed; the first game ended in a scoreless tie. By the mid-1950s the professional teams became dominant, losing only once after 1960. As that dominance became obvious, interest waned. A heavy rainstorm forced the cancellation of the 1976 game, the final game in the series, late in the third quarter.[1]

Sporting contests are less interesting when the outcome is easily foreseen; the parity of the competitors is of considerable interest. The appropriate measure of parity, however, is not so readily obvious. Holding team characteristics constant,

AUTHORS' NOTE: We benefited from the contributions of Lynne Kiesling, John Panzar, Robert Porter, Ian Savage, and particularly David Surdam, who made extraordinary contributions to this piece. E. Kirk Ross and Brent Roth provided research assistance.

JOURNAL OF SPORTS ECONOMICS, Vol. 7 No. 3, August 2006 330–338
DOI: 10.1177/1527002504272937

shorter series, shorter seasons will more likely result in unbalanced records than will longer ones. Recognizing that, Quirk and Fort (1992) devised a measure—the idealized standard deviation (ISD)—to enable them to compare records in Major League Baseball (MLB) across seasons of different lengths and league composition. That measure has subsequently been adopted widely by other researchers.

As used by Quirk and Fort (1992; see also Fort & Quirk, 1995), the ISD seems unobjectionable as a measure of parity. A subtle difficulty arises, however, if their measure is used to examine parity within a league that awards points to ties, and it becomes especially problematic for cross-league comparisons where one league (such as soccer) experiences a large number of ties while another (such as baseball) plays each game until there is a winner. Researchers making cross-league comparisons ordinarily treat ties as a half-win/half-loss outcome, which is generally inappropriate. If baseball permitted tied games, and the games lasted only one or two innings, a league's best and worst teams would frequently tie. That would not mean that the teams were equal, merely that the game had been so severely truncated that the better team had too little opportunity to reveal itself. Because soccer and ice hockey tend to result in such low scoring affairs, an inferior team is quite often able to achieve a tie against a superior one. If those games were twice as long, there would be fewer ties, of course; however, more to the point the superior team would likely win more than one half of them.

Because baseball is accounted on a binomial basis, a win counting as 1 and a loss as 0, Fort and Quirk (Fort & Quirk, 1995; Quirk & Fort, 1992) made use of the binomial theorem to develop their parity measure. The ISD is the quotient of the actual standard deviation of results in a season divided by $.5/\sqrt{N}$, where N is equal to the number of games in a season, which works well for baseball and other sports that compare results on an one-zero basis.

The application of the Fort-Quirk ISD correction to sports such as North American hockey or English football raises some logical questions. Its use follows from an assumption that ties are one half a win and one half a loss.[2] Because the denominator of the Fort-Quirk ISD assumes that the probability of each team winning the game is .5, this would appear to be applicable to the case of a tie. In this context, however, .5 is a conditional probability: It is the probability that, were the contest to continue to resolution, the probability of each team winning in overtime is .5. The Fort-Quirk ISD does not take into account the probability that, before the fact, one of the outcomes of the competition is a tie.

Unfortunately, the theory of equal ignorance does not generate a probability before the fact for a win-loss versus a tie. That is an empirical question. We have addressed that question in the case of the English Football League. We have examined the record of each team for each season in today's Premier League (formerly the First Division) and Championship Division (formerly the Second Division). Between the 1888-1889 and 2003-2004 seasons, teams in the Premier League experienced wins 37.705% of the time, losses 37.705% of the time, and draws for the remaining 24.59%. The results from the Championship Division are similar:

37.64% wins, 24.72% draws, and 37.64% losses. Based on that, we calculated the standard deviation of the distribution where teams entering a contest had, before the fact, a 25% probability of leaving with a draw.

Such a calculation also requires knowledge of the scoring system because the standard deviation depends on the expected value, and the expected value depends on how many points are awarded for a win, a tie, and a loss. Assume, as was the case, that a win is worth 2 points, a tie is worth 1 point, and a loss is equal to 0. At first blush, because a tie is worth one half what a win is worth, it might appear that this leads one back to the binomial correction. That is an erroneous conclusion.

Under the assumption that the probability of a tie is 25%, the expected number of points earned in any contest is equal to 1 [$= .375 \times 2 + .25 \times 1 + .375 \times 0$]. The variance is equal to .75 [$= .375 \times (2 - 1)^2 + .25 \times (1 - 1)^2 + .375 \times (0 - 1)^2$]. In the binomial case, the variance is equal to .25 [$= .5 \times (1 - .5)^2 + .5 \times (0 \times .5)^2$].[3] Thus, for a season of N games, the variance is equal to $.75 \times N$, and the standard deviation is equal to $\sqrt{(.75 \times N)}$. Given that the standard deviation is calculated on the basis of an absolute number of points, not a percentage, there is no need to divide through by N to put the standard deviation in percentage terms.[4]

The differences between the approaches can be illustrated by looking at 1888, the inaugural season of the Football League. In Table 1, the first four columns report the results for each team (the names of the teams are omitted). In the next pair of columns, the results based on (2, 1, 0) system are reported as well as the percentage of points won, and in the final pair of columns, the results based on a (3, 1, 0) system are reported as well as the percentage of those points won. The calculation of the win percentage is based on the ratio of points earned to total possible points; hence, the win percentage in a 3-point system will be less than in a 2-point system. At the bottom of each of the final four columns is the actual standard deviation for that column and the ISD based on half-win/half-loss assumption for ties. It should be noted that the four idealized standard deviations are similar in size but different in magnitude. This raises the question of which ISD is ideal?

The answer to that question is of relatively little importance when the focus of one's research is within a league across seasons of identical lengths but becomes problematic when one attempts comparisons between leagues, especially between two leagues with different methodologies for determining the rank order of teams. We can only answer that question for ourselves, and we favor the first of those reported, where the base standard deviation is that calculated based on a win being worth 2 points; a draw, 1 point; and a loss, 0 points. We prefer this approach because (a) it is based on a point system used by leagues with draws around the world, (b) the probability of a draw is considered a priori, and (c) the denominator ($\sqrt{.75 \times N}$) is comparatively simple.

Others are free to select one of the others; and, to that end, we have calculated ISDs using each of these approaches for each year of the top two divisions in the English Football League. Those ISDs for selected years are reported in Appendix Table A1.[5] We have also calculated the ISDs, following the Fort-Quirk method only,

TABLE 1: Calculation of Idealized Standard Deviation

	1888 Season, Football League						
Games	*Win*	*Draw*	*Lose*	*2 Point Win*	*2 Point Win %*	*3 Point Win*	*3 Point Win %*
22	18	4	0	40	.909	58	.879
22	12	5	5	29	.659	41	.621
22	12	4	6	28	.636	40	.606
22	10	6	6	26	.591	36	.545
22	10	2	10	22	.500	32	.485
22	10	2	10	22	.500	32	.485
22	6	8	8	20	.455	26	.394
22	9	2	11	20	.455	29	.439
22	7	3	12	17	.386	24	.364
22	7	2	13	16	.364	23	.348
22	5	2	15	12	.273	17	.258
22	4	4	14	12	.273	16	.242
Actual *SD*				7.943	.181	11.676	.177
Idealized *SD*				1.955	1.693	1.890	1.660

for each season played by the two leagues in MLB, and for the two leagues taken together. These results for selected years appear in Appendix Table A2. We believe the similarity of the magnitudes and changes in the ISDs for the two leagues separately provides a justification for looking at a single ISD for each season of MLB.[6] The close comparability of the two baseball leagues can be seen in Figure 1. We have adopted a 7-year moving average to smooth some of the year-to-year fluctuations.

More important, our conception of parity is that increasing parity should reflect an upward slope on such a diagram. Consequently, the numbers that underlie Figure 1 are the inverse of the ISDs for MLB. Given the definition of an ISD, increasing parity means counterintuitively that the ISD is decreasing because the actual standard deviation of that season's performance is in the numerator.

The inverse of the ISD as a 7-year moving average also underlies Figure 2 where the top two divisions of the English Football League are depicted along with a single inverse ISD for MLB. That baseball has generally moved toward greater parity over its history is clear from this diagram. There are some so-called blips along the way, particularly the Great Depression and World War II; however, the trend is significantly positive.[7]

Of much greater interest is the U-shaped nature of the parity curves for the top two divisions of the Football League. As yet we have no explanation for that shape. It is curious that the so-called peak for the Championship Division occurs just after the break for World War I, while that for the Premier Division occurs just before the break for World War II. Both divisions have relatively sharp upward movements at the time the number of competitors in the league increases. The Premier League increases rapidly just after the Championship Division began play. The Champion-

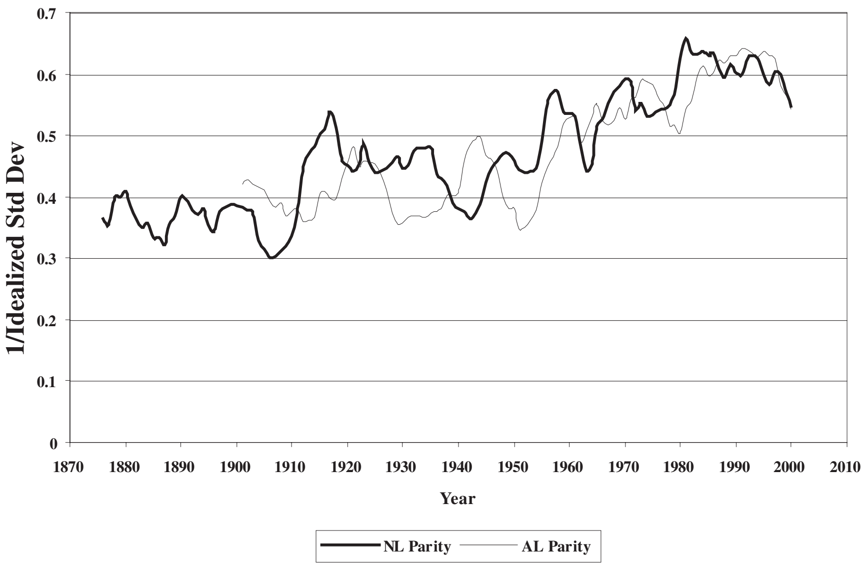

Figure 1: Parity in Major League Baseball: 7-Year Moving Average

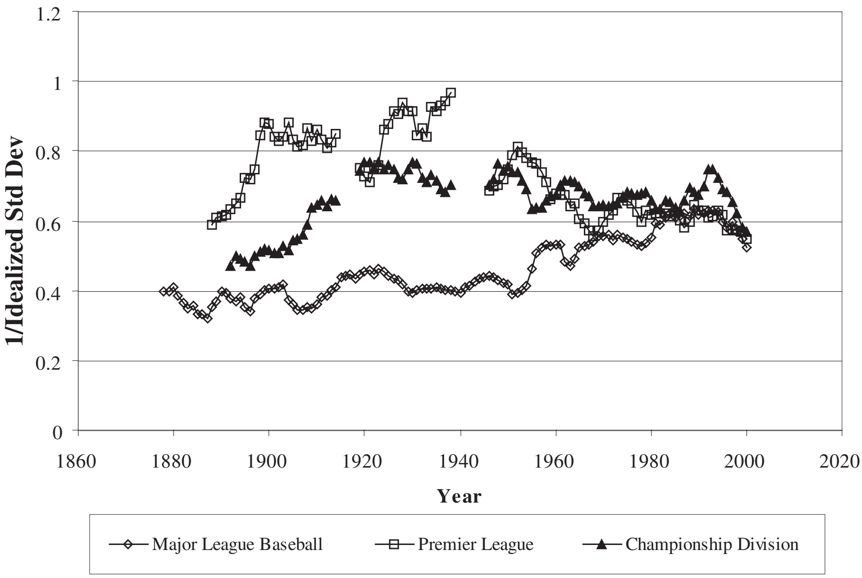

Figure 2: Major League Baseball, Premier League, and Championship Division Football: 7-Year Moving Average

ship Division has a large increase during the World War I hiatus, just before League One (formerly the Third Division – South) was added. Even before the war, the teams that were being elected into the Championship Division were from large markets (e.g., London), while the teams that were excluded were from small markets (e.g., Lincoln). As Figure 2 reveals, parity in the Championship Division began to wane after World War I.

Parity in the Premier League continued to increase until World War II. That, in itself, is something of a puzzle. Indeed, during the 1930s, league parity is actually greater than what would be expected from a random draw from an a priori probability function. Then it starts down. The 7-year moving average reaches its postwar peak in 1952; actual parity, in 1954, immediately prior to the advent of two continent-wide competitions, the Champions League and the Union of European Football Associations (UEFA) Cup.[8] Other possibilities include the relaxation of salary rules and increased media revenue (especially in the larger markets). This is neither the time nor the place to resolve such issues; they are left for further research.

Our aim here has been merely to raise some issues regarding the concept of the ISD as applied to leagues that do not have a binomial system for determining rankings, to suggest that the reciprocal of the ISD may be a more intuitive construct if the intent is to represent parity, to provide scholars with a set of ISDs, and to suggest some interesting topics that arise from a quick perusal of the time series.

APPENDIX

TABLE A1: Idealized Standard Deviations

Season Year Begun	*Football League*							
	Premier League				*First Division*			
	W = 2		*W = 3*		*W = 2*		*W = 3*	
	Abs	*Prcnt*	*Abs*	*Prcnt*	*Abs*	*Prcnt*	*Abs*	*Prcnt*
1888	1.9554	1.6934	1.8902	1.6596				
1890	1.3484	1.1677	1.3373	1.1742				
1895	1.6257	1.4079	1.5611	1.3706	2.1773	1.8856	2.1652	1.9010
1900	1.2151	1.0523	1.1972	1.0512	1.5637	1.3542	1.5203	1.3348
1905	1.1863	1.0273	1.1888	1.0437	2.3102	2.0007	2.2705	1.9934
1910	1.3885	1.2025	1.3847	1.2157	1.4749	1.2773	1.4609	1.2826
1914	1.1040	0.9561	1.1427	1.0033	1.6159	1.3994	1.5868	1.3932
1915-1918	Play suspended during World War I							
1919	1.2706	1.1004	1.3462	1.1820	2.0263	1.7548	2.0189	1.7726
1920	1.5552	1.3469	1.5843	1.3910	1.3254	1.1478	1.3540	1.1888
1925	1.0198	0.8832	1.0106	0.8873	1.4816	1.2831	1.4968	1.3141
1930	1.6867	1.4607	1.6286	1.4299	1.4147	1.2252	1.4219	1.2484
1935	0.8951	0.7752	0.9014	0.7914	1.5298	1.3248	1.4954	1.3129
1938	1.3197	1.1429	1.3622	1.1959	1.6413	1.4214	1.6100	1.4135
1939-1945	Play suspended during World War II							
1946	1.7715	1.5342	1.7442	1.5314	1.6651	1.4420	1.5597	1.3694
1950	1.3367	1.1576	1.3040	1.1449	1.2919	1.1188	1.3100	1.1502
1955	1.0562	0.9147	1.0338	0.9076	1.3975	1.2103	1.4050	1.2336
1960	1.6651	1.4420	1.6951	1.4882	1.4631	1.2671	1.4647	1.2860
1965	1.6596	1.4373	1.6054	1.4095	1.4816	1.2831	1.4344	1.2593
1970	1.8760	1.6246	1.9093	1.6763	1.7379	1.5051	1.7158	1.5064
1975	1.7204	1.4900	1.6644	1.4613	1.4360	1.2436	1.4104	1.2383
1980	1.7257	1.4945	1.6079	1.4117	1.5218	1.3179	1.5116	1.3272
1985	2.1501	1.8620	2.1401	1.8789	1.4179	1.2280	1.3577	1.1920
1990	1.8343	1.5886	1.7847	1.5669	1.4003	1.2127	1.4038	1.2325
1995	1.8947	1.6409	1.8676	1.6397	1.1260	0.9751	1.0803	0.9485
2000	1.6962	1.4690	1.7210	1.5110	1.7976	1.5567	1.8184	1.5965
2003	1.8643	1.6145	1.8696	1.6414	1.7304	1.4986	1.6396	1.4395

NOTE: W = the number of points awarded for a win; Abs = the idealized standard deviation based on the absolute number of points awarded; Prcnt = the idealized standard deviation based on the percentage of the points awarded.

Table A2: Idealized Standard Deviation

Season	Major League Baseball (MLB)		
	National League	*American League*	*MLB*
1876	3.7400		3.7400
1880	3.1301		3.1301
1885	3.7282		3.7282
1890	3.8805		3.8805
1895	2.9955		2.9955
1900	1.4711		1.4711
1901	2.5660	2.1135	2.2686
1905	3.5142	2.1688	2.8179
1910	2.6758	2.7921	2.6421
1915	1.2181	3.4688	2.5115
1920	1.7135	2.9226	2.3144
1925	1.6430	2.4376	2.0082
1930	2.4228	2.7472	2.5023
1935	2.0639	2.5070	2.6365
1940	2.4344	2.1810	2.2332
1945	2.8488	1.8076	2.3008
1950	2.0303	3.1381	2.5524
1955	1.8619	2.7499	2.2683
1960	2.1691	2.1119	2.0681
1965	2.3060	2.2975	2.2404
1970	1.5596	2.4037	1.9814
1975	1.8761	1.8618	1.8279
1980	1.5837	2.0375	1.8068
1985	2.2058	1.8454	1.9778
1990	1.4479	1.4534	1.4216
1995	1.4330	1.9922	1.7026
2000	1.7574	1.4587	1.5707
2001	1.6461	2.6124	2.0478
2002	2.0569	2.6024	2.3275
2003	1.7903	2.3518	2.1017

NOTES

1. In addition to the waning interest, increasing salaries, and the fear of injury were factors in the cancellation.

2. See, for example, Szymanski and Valletti (2003).

3. The expected value in a (1,0) binomial system is 0.5 (Szymanski (2003, p. 1154).

4. If 3 points are awarded for a win, 1 for a draw, and 0 for a loss, as is now the standard in the English Football League, the expected value is 1.375 and the standard deviation is equal to $\sqrt{(1.734375 \times N)}$. If the percentage of total possible points is used to calculate the result of each team, then $0.5/\sqrt{N}$ is the appropriate correction.

5. The complete set of years is available from the authors.

6. When the inverse ISDs are regressed against time, the estimated value for the National League in 1905 is .394 and parity is increasing at a rate of .0022 per year. For the American League, the estimated value in 1905 is .365, and parity is increasing at a rate of .0024 per year.

7. A dummy variable for the advent of free agency was included in a regression of the ISD versus time. One would accept the hypothesis that the coefficient on this variable was essentially zero.

8. Regression analysis suggests that the UEFA Cup has had a stronger influence than the Champions League on the change in the ISD.

REFERENCES

Fort, R., & Quirk, J. (1995). Cross-subsidization, incentives and outcomes in professional team sports leagues. *Journal of Economic Literature*, *33*(3), 1265-1299.

Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature*, *16*(4), 1137-1187.

Szymanski, S., & Valletti, T. M. (2003, June). *Promotion and relegation in sporting contests* Available at www1.imperial.ac.uk/_Resources/(3FD6F572-7C30-4955-9F42-182250A05EE9)/promotionandrelegationinsportingcontestsfinal.doc

Quirk, J., & Fort, R. (1992). *Pay dirt: The business of professional team sports* Princeton, NJ: Princeton University Press.

Louis P. Cain is a professor of economics at Loyola University Chicago and an adjunct professor of economics at Northwestern University.

David D. Haddock is a professor of economics and of law at Northwestern University.