

## Do Good Teams Really Win More of the Close Games?

By David W. Smith

Presented July 10, 2003 at SABR33, Denver, Colorado

This topic first suggested itself to me last fall when the Angels were in the post-season and it was very common to hear how exceptional they had been in close games during 2002. Of course, many people find close games more exciting, but I also detected a different sort of praise for the Angels, an idea that their success represented some admirable “clutch” performance. I decided to pursue the phenomenon of close games in more detail and hopefully come up with some new perspectives along the way.

Let’s start with definitions, and there will be different ones before we are done. The simplest way to begin is to define a close game as a 1-run difference in the final score. The 2002 Angels played 53 regular season games decided by one run with a record of 31-22 (.585). Their record for the whole year was 99-63 (.611). Therefore, 32.7% of their games were 1-Run decisions. Was that exceptional? Was it normal? I examined these questions by looking at the scores of all Major League games from 1901 to 2002 and sorting the results by team. Table 1 summarizes the scope of the data used; the source was the game logs posted at [www.retrosheet.org](http://www.retrosheet.org).

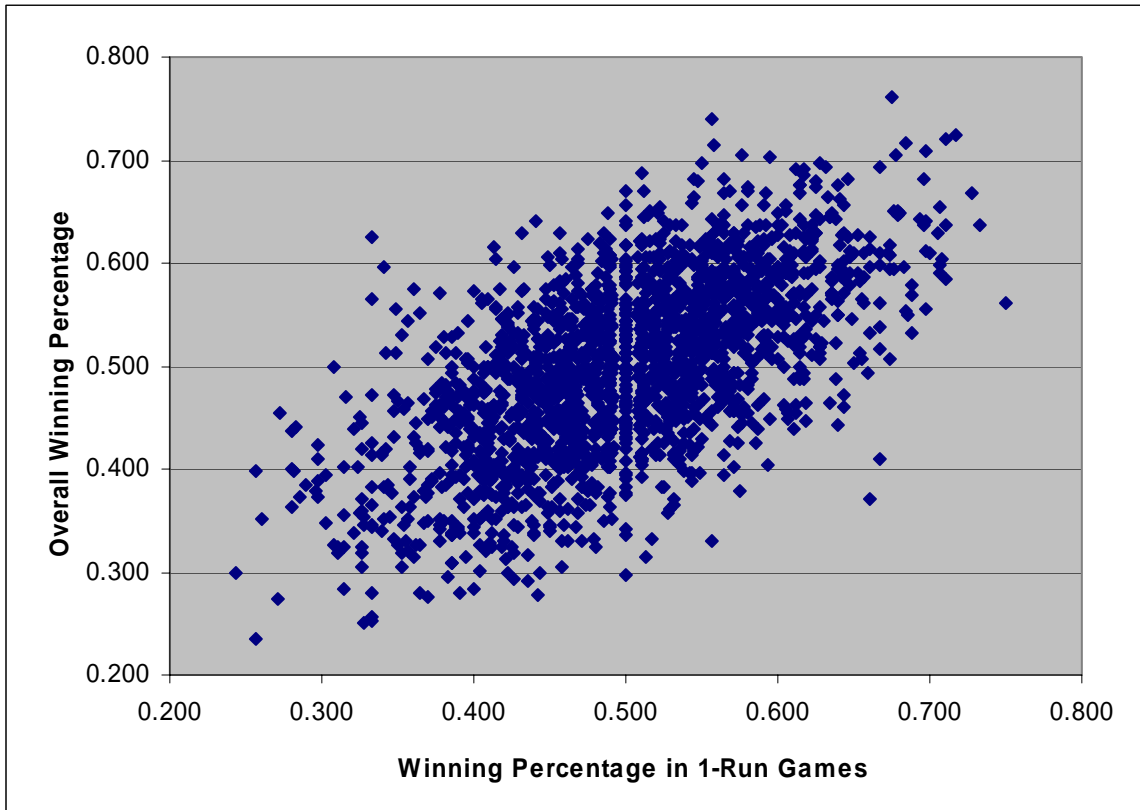
Table 1. Data analyzed in this study to investigate margin of victory.

Team-Seasons	Total Games	1-Run Games	% 1-Run Games
2032	157653	48307	30.6

So, we can answer one question now. The Angels were a little above the 102 year average with their mark of 32.7%, but not much. For the record, the high value is 46.3% of games being 1-Run contests for the 1971 Astros and the low is 14.5% by the 1936 Browns.

Of course, the more important question is whether success in these games translates to success overall. That is addressed in Figure 1, which is a graph of the winning percentage for each team in each season vs the winning percentage in 1-Run games for that season. Therefore, there are 2032 data points here:

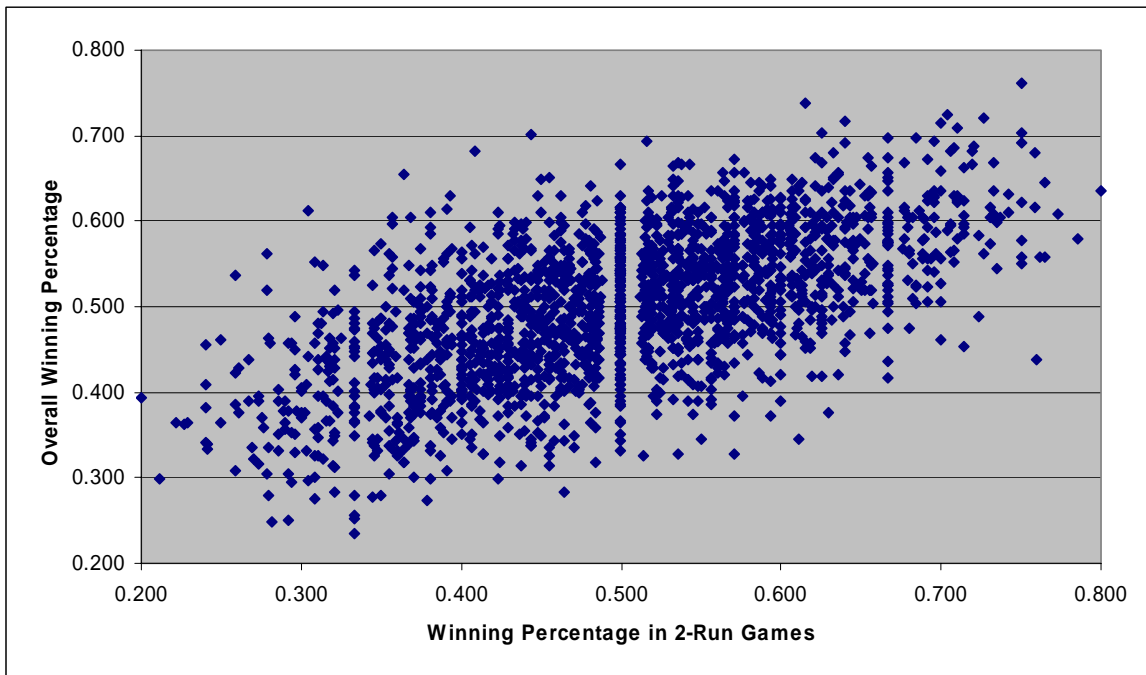
Figure 1. Winning percentage for each team-season in all games vs winning percentage in 1-Run games in that team-season.



Visual inspection indicates a positive relationship between the two values and this is confirmed statistically. The correlation coefficient for these data ( $r$ ) is 0.63, which is fairly strong. The Angels' performance in 2002 (.611 vs .585) puts them in the higher range of these results. Therefore, the first and simplest answer to the question in the title is yes: better teams do win more of their close games, at least of their 1-run games.

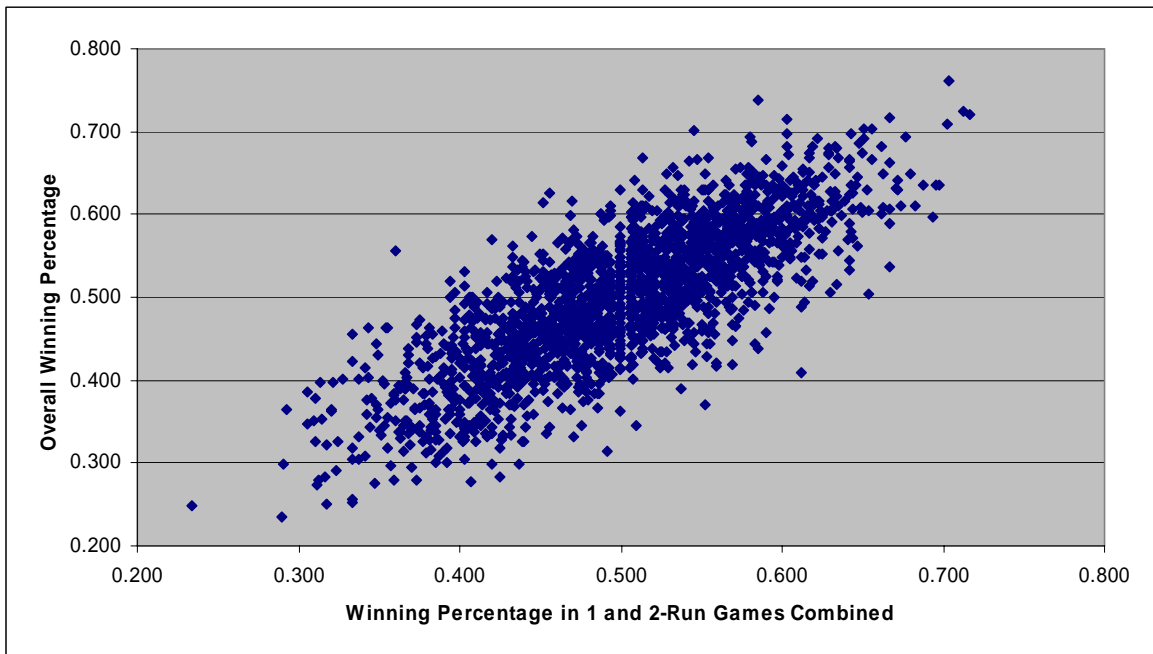
However, we must be sure that this is actually a meaningful result. For example, it may not be that good teams are exceptionally successful just in close games. Since by definition they win more games, perhaps they win more of all kinds of games, no matter the margin. Figure 2 presents the results for 2-Run games, since these contests could also reasonably be defined as close games.

Figure 2. Winning percentage for each team-season in all games vs winning percentage in 2-Run games in that team-season.



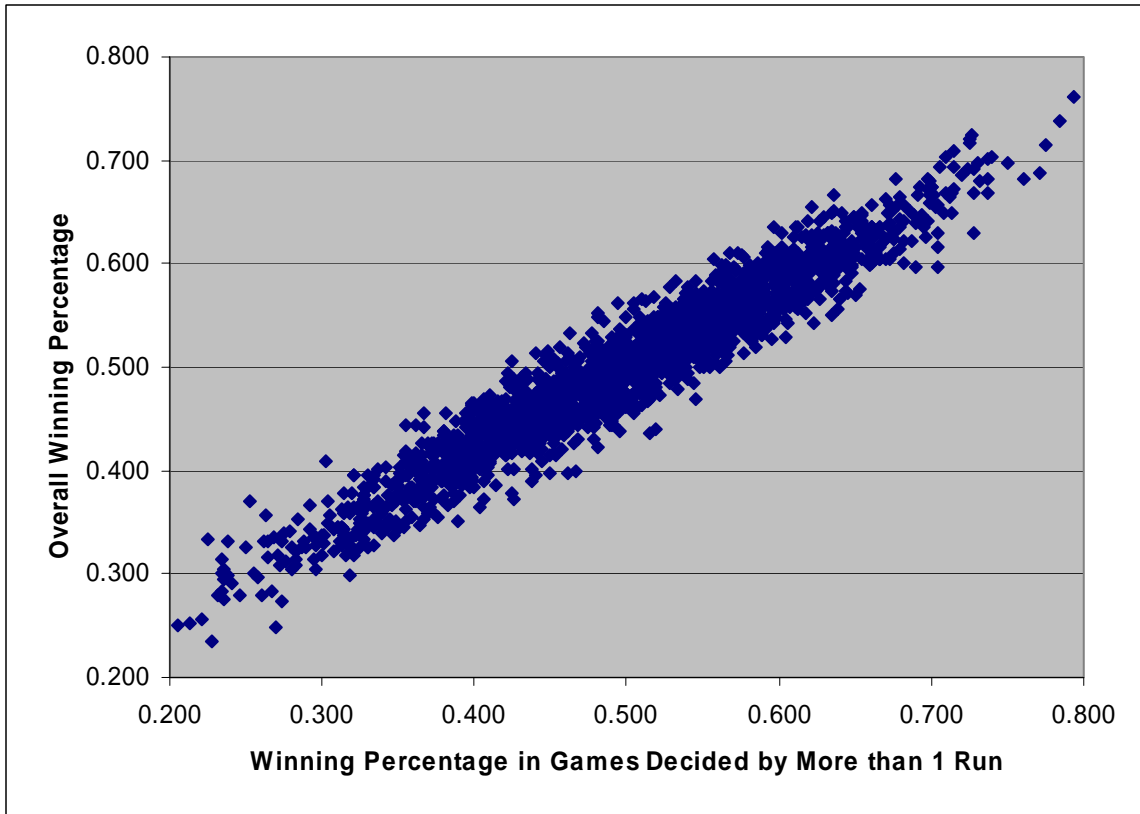
There is still a positive relation and the correlation coefficient is 0.62, important, but not overwhelming. I continued this analysis by looking at 1- and 2-Run games together, again under the assumption that they are both close results. Those numbers are in Figure 3.

Figure 3. Winning percentage for each team-season in all games vs winning percentage in games decided by one or two runs in that team-season.



This combination gives us a tighter relation, with a correlation coefficient of 0.79, which is a much stronger connection. At this point, I decided to take the flip side of close games and consider all games in which the winning margin was more than one run. The results, which were quite surprising to me, are in Figure 4, which is the same format as the first three graphs.

Figure 4. Winning percentage for each team-season in all games vs winning percentage in games decided by more than one run in that team-season.



The statistical test this time gives a correlation coefficient of 0.96, which is highly significant. This basically reverses my initial conclusion. That is, a win is a win, but in terms of relation to overall success, the 1-run victory is just not that special. The ability to win by a larger margin tells us much more. This conclusion is strengthened by the fact that 49% of all games are decided by more than two runs. How do the 2002 Angels stack up in this new analysis? Table 2 gives that answer.

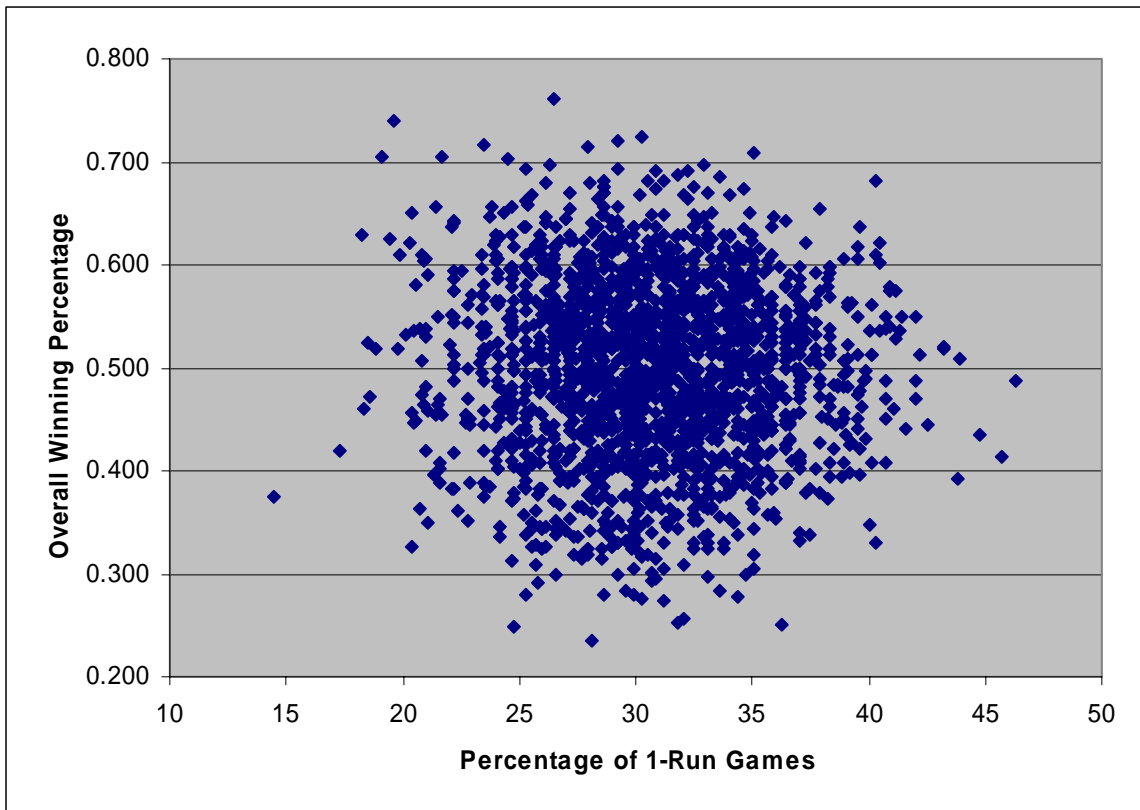
Table 2. Winning percentage of 2002 Angels in games of different final margins.

Final Score Margin	Wins	Losses	Winning Percentage
1-Run Games	31	22	.585
2-Run Games	11	11	.500
1&2-Run Games	42	33	.560
More than 1 Run	68	41	.624
Total	99	63	.611

There is a problem with small sample size, since the Angels only played 22 games that were decided by exactly two runs. However, when the large set of games decided by more than one run is examined, then once again the Angels are in the upper portions of this graph. The historical high value for winning percentage of games decided by more than two runs is .793 for the 1906 Cubs (overall percentage of .762) and the low is .192 for the 1942 Phillies (overall percentage of .278).

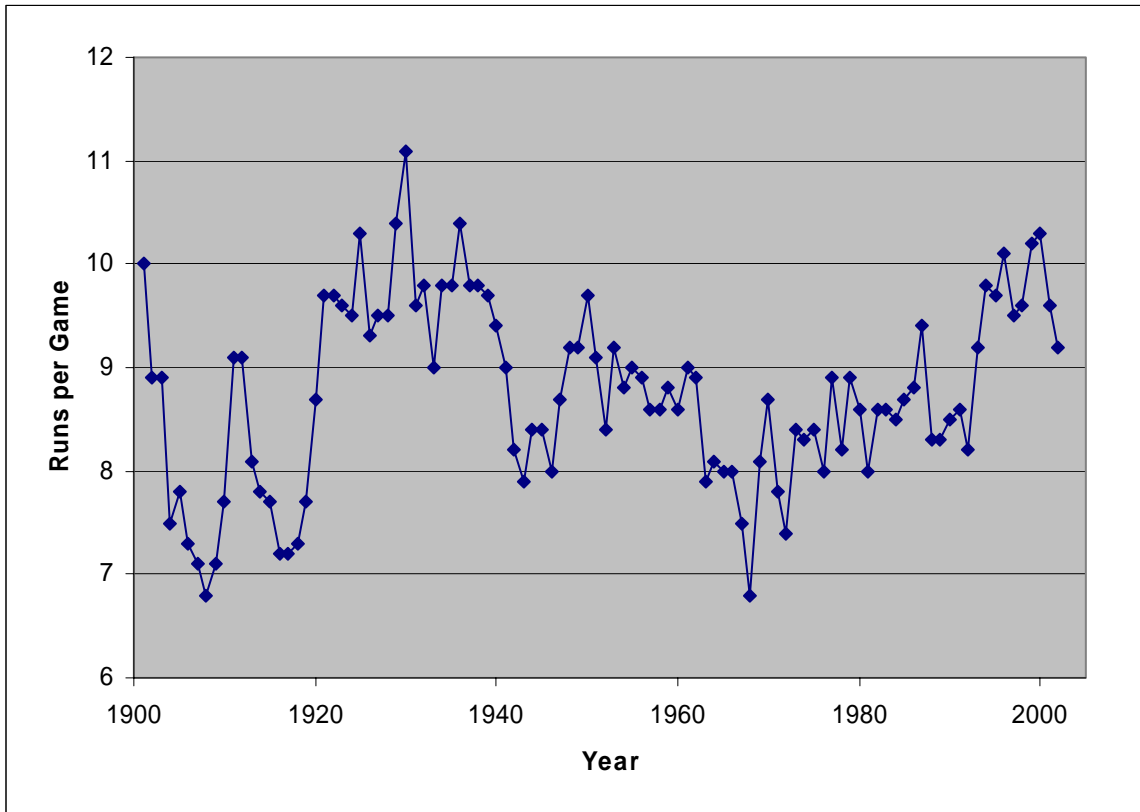
It is also interesting to ask if the good teams play more one run games. Figure 5 addresses this point. It is very clear that the chance of a close game has no relation to the quality of the team. The correlation coefficient for this relationship is  $-0.03$ , which shows almost no significance

Figure 5. Winning percentage for each team-season in all games vs percentage of 1-Run games.



Another variable to consider is the level of scoring. During major league history, there have certainly been periods of high offense and periods of less offense. It is reasonable to ask if the probability of a 1-run game is related to these changes. First, Figure 6 has a brief review of runs per game for both teams combined from 1901 to 2002.

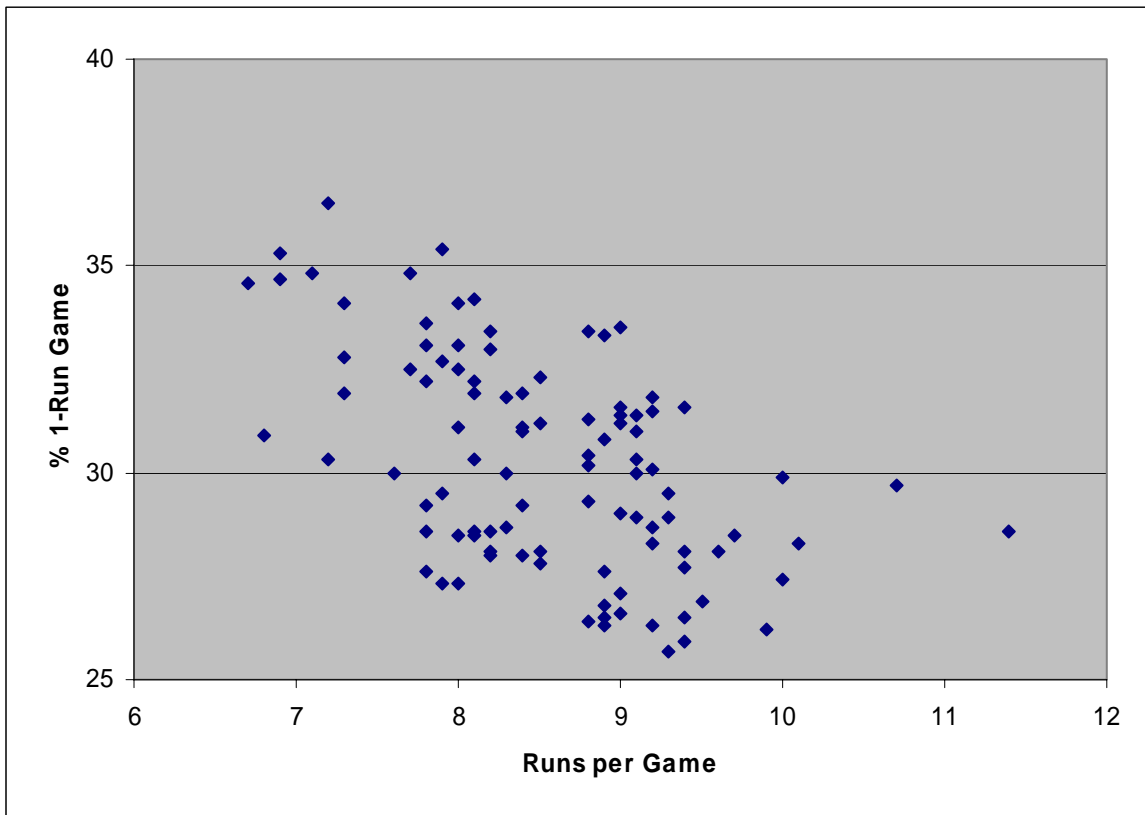
Figure 6. Runs per game, combined for both teams, from 1901 to 2002.



This basic graph has been presented in many places, but it is useful to review the main features briefly. The leagues have been combined here to make the graph less messy. There are indeed differences between the leagues, but in most cases, the yearly changes occur in a parallel way. Note that the rate of run scoring has dropped in the last two years, back to where it was in 1993. By the way, so far in 2003, through June 30, the value is back up to 9.6, which is the same as it was in 2001. I tried to separate my analysis by era, but it wasn't clear where the dividing lines should be: the deadball era before 1920? In 1947 when integration began? In 1961 when integration is generally considered to have been completed? In 1973 when the DH began? I saw no empirical basis for any of these boundaries and the data I have already presented had all seasons combined and some very clear trends emerged. Therefore, I chose to consider all years equally.

The question then is: is there a relation between level of run scoring and number of close games? The information to address this is in Figure 7.

Figure 7. Percentage of 1-Run games as a function of runs scored per game. Each data point is all the games for one season, all teams combined, 102 total data points.

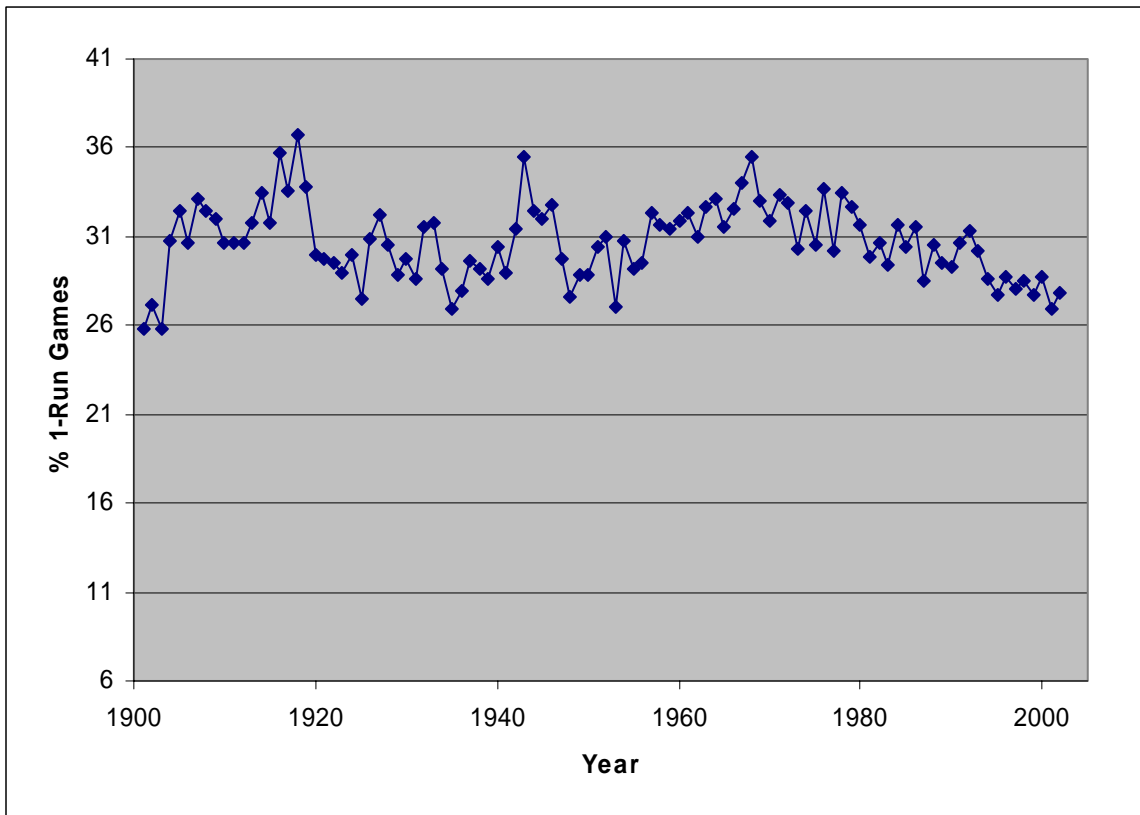


There is a definite relation, with the correlation coefficient being  $-0.73$ . This does make sense, since there are more chances for wide margins of victory when there are more total runs being scored in the average game.

The meaning of this result in light of the earlier data is that the basic patterns I have found seem to be independent of the level of offense. That is, the frequency of 1-run games changes, but its predictive value in overall team success is not affected. The final demonstration of this point is seen in Figure 8, which presents the percentage of 1-run games each year. The variation here is much less than that seen in runs per game (Fig. 6), showing the relative constancy of occurrence of 1-run games.



Figure 8. Percentage of 1-run games each year. Each data point is all the games for one season, all teams combined.



The final topic I will address is the change over the years in the occurrence of specific final scores. That is, what is the most common score of a major league game and how has it changed over the years? During the 1901-2002 seasons, there were 279 different final scores in Major League games. The five most frequent, along with the number of occurrences and percentage of each are shown in Table 3.

Table 3. Most frequent final scores, 1901-2002

Final Score	Number of Occurrences	Percentage of All Games
3-2	9274	5.9
4-3	8809	5.6
2-1	7671	4.8
5-4	6778	4.3
4-2	5787	3.7

The top four are all 1-run games, comprising some 21% of all games, which fits with what we have seen before. As to the question of changes in this pattern, I offer Table 4 which presents a series of snapshots at 20 year intervals.

Table 4. Most frequent final scores, each 20 years, 1901-2002

1901		1920		1940	
Score	%	Score	%	Score	%
4-3	5.1	3-2	6.2	4-3	5.7
5-4	4.3	2-1	5.3	3-2	5.4
5-3	4.1	4-3	5.0	3-1	4.4
3-2	3.6	5-4	4.0	2-1	4.1
4-2	3.5	3-1	3.5	5-4	4.1
1960		1980		2000	
Score	%	Score	%	Score	%
3-2	6.7	3-2	6.5	4-3	5.0
4-3	5.2	4-3	6.0	3-2	4.4
2-1	4.9	2-1	4.9	5-4	4.2
5-3	4.6	5-4	4.8	2-1	3.3
5-4	4.4	5-3	4.4	6-5	3.3

Four different scores occupied the top spot on this list in these 102 seasons, as summarized in Table 5.

Table 5. Number of times each score was most frequent in a season, 1901-2002

Final Score	Number of Occurrences
3-2	53
4-3	36
2-1	11
5-4	2 (1930 and 1935)

Remember that 1-Run games occur just over 30% of the time, even though the single most frequent result for each year is a 1-Run decision. The greatest proportion in this category was in 1916, when 9.0 % of all games ended up 3-2. The lowest proportion came in 1935, when the most frequent game was 5-4, but that only happened 4.3% of the time in that era of very high offense.

To revisit the question posed in the title, it seems reasonable to conclude that, while 1-Run games may provide good theater, they are not especially good predictors of ultimate team success.

**SUMMARY:**

- Success in 1-Run Games is correlated with overall success
- Success in Games won by a larger margin is a better predictor
- Frequency of 1-Run Games varies with level of scoring
- Most frequent final scores vary widely across the years

**SUGGESTIONS FOR FURTHER STUDY**

- Patterns within a game:
  - Value of scoring first
  - Come from behind wins
  - Wins in last at bat