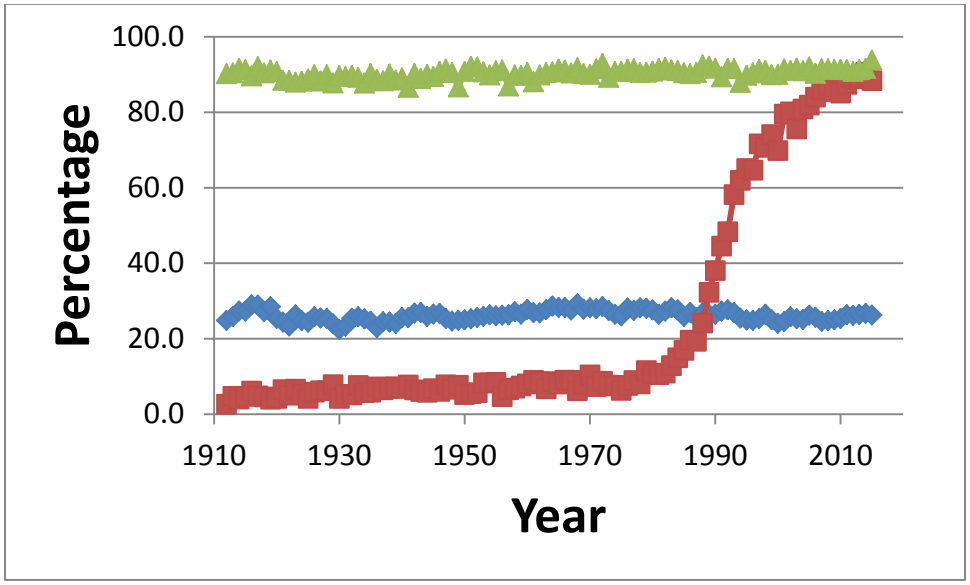
The Myth of the Closer

By David W. Smith
Presented July 29, 2016
SABR46, Miami, Florida

Every team spends much effort and money to select its closer, the pitcher who enters in the ninth inning to seal the deal and nail down a win. Of course, this now universal use of a closer only became the norm relatively recently and the reason given is simple: teams feel their chance of winning is increased by having this ace specialist. I decided to look at this assumption more closely and came to a probably startling conclusion: it isn't true that closers increase the chance for a team to win, or at least it is marginally true at best.

I will start by presenting a graph that looks a lot like a conclusion. Then I will work backward and take it apart to look for explanations. As a starting point, I will use the term "closer" to refer to a pitcher entering in the 9th inning in a save situation but we will return to examine that definition a little later. Here is the summary in Figure 1, covering 1912 to 2015.

Figure 1. Closer Entry and Relation to Winning, 1912-2015

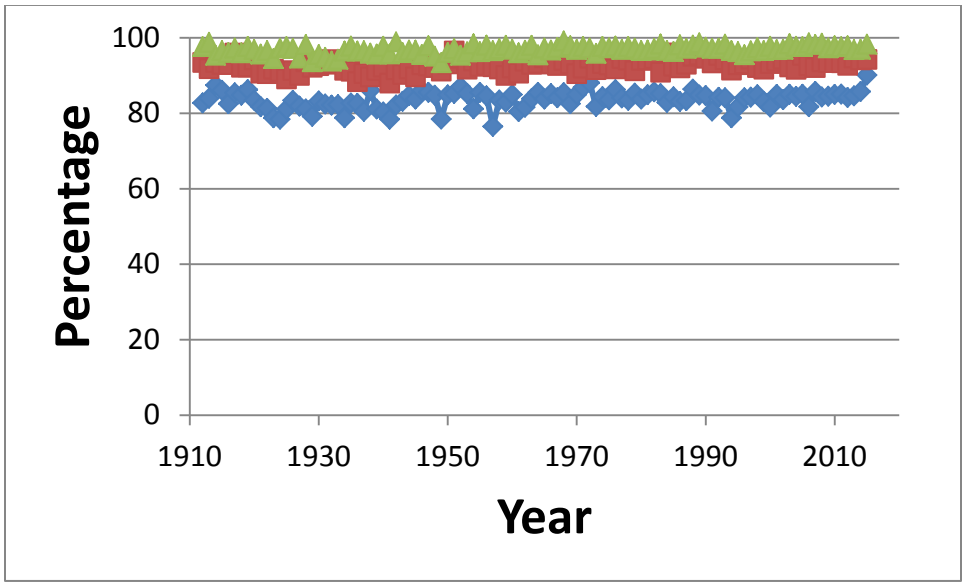


The blue line is the percentage of games in which a team entered the 9th inning with a lead of three runs or fewer, the modern definition of a save situation. The red line shows the percentage of those games in which a relief pitcher was brought in to start the 9th. Finally the green line is winning percentage in these games. There are four immediate conclusions. First, the chance of a team starting the 9th inning in a save situation (about 25%) is essentially unchanged over the past 104 seasons, covering 176791 games. Second, from 1912 to about 1980, there was very little change in the frequency of bringing in a new pitcher, being stable at about 10 %. Third, things started to change in 1980 and there is a dramatic increase in using a new pitcher and it

now happens over 90% of the time. Fourth and most importantly the chance of winning the game in these situations has shown amazingly little variation at a bit over 90%.
No matter when one wishes to claim the era of the closer began, it is clear from this graph that closers do not increase the chance of winning the game. Of course, this analysis does not explain individual games, but on the aggregate level, it is clear that in terms of wins, there is no long-term benefit to the closer strategy. One must wonder why it is employed so universally when there is no advantage to doing so. Although it would be tempting to stop the analysis here with these conclusions, I decided to go a bit further to see if I could deconstruct the results to elucidate underlying features that may have led to the current pattern.

The first thing I did is pretty obvious, which is to look at the three different kinds of save situation, namely leads of one, two, or three runs. These results are in Figure 2.

Figure 2. Winning Percentage as Function of 9th Inning Lead when Reliever Entered, 1912-2015



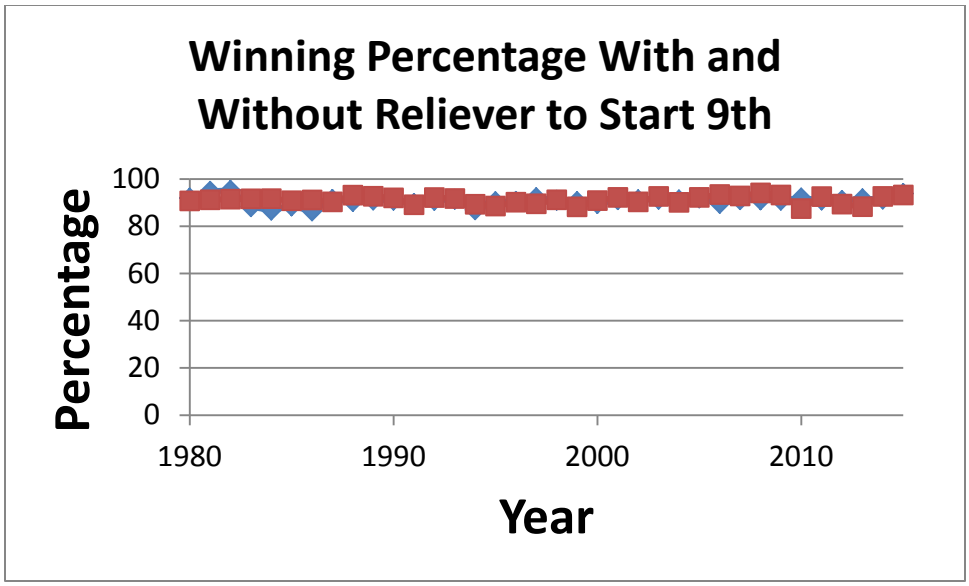
The scale is the same as in Figure 1 to facilitate comparisons. The expected pattern is observed, namely that entering with a three run lead translates to a win an average of 97% of the time, two runs, 93% and one run, 84%.

The winning percentages presented so far are for all 9th inning save situations. In order to explore the closer effect more explicitly, I subdivided these appearances as follows:

1. New relief pitcher.
2. Relief pitcher who was already in the game or starter still in.

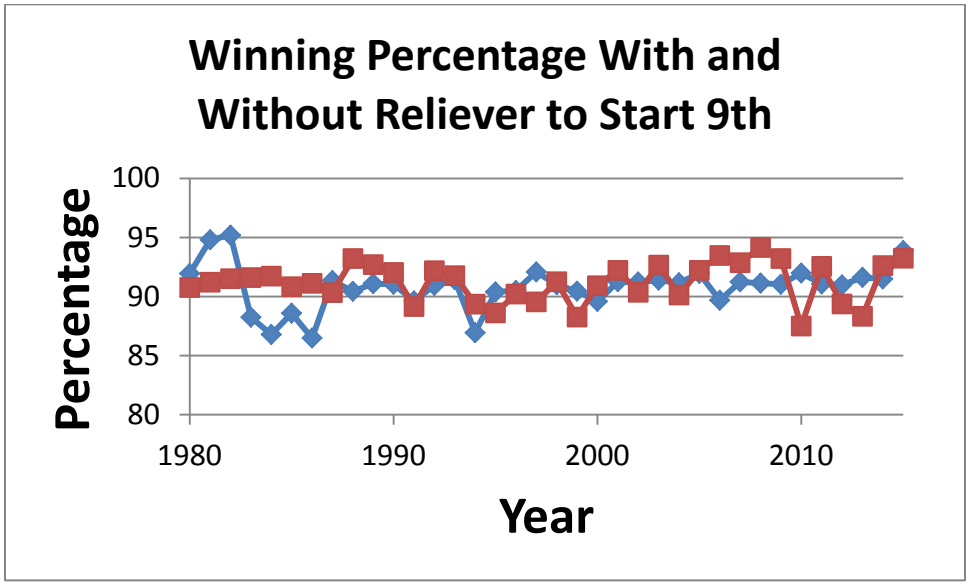
I present the results from 1980 through 2015, since it is clear from the first graph that this period is the time of rapid change in 9th inning relief pitcher usage and is the appropriate focus to address the effect of closers.

Figure 3. Winning percentage with and without a reliever to start 9th inning, 1980-2015.



There really are two lines in this graph, with the blue line for games with a reliever to start the 9th and the red line when a new pitcher does not start the inning. The percentage scale is the same as in previous graphs for the sake of comparison. However, to see the picture more clearly, Figure 4 presents the same data, but with the percentage axis expanded.

Figure 4. Winning percentage with and without a reliever to start 9th inning, 1980-2015.



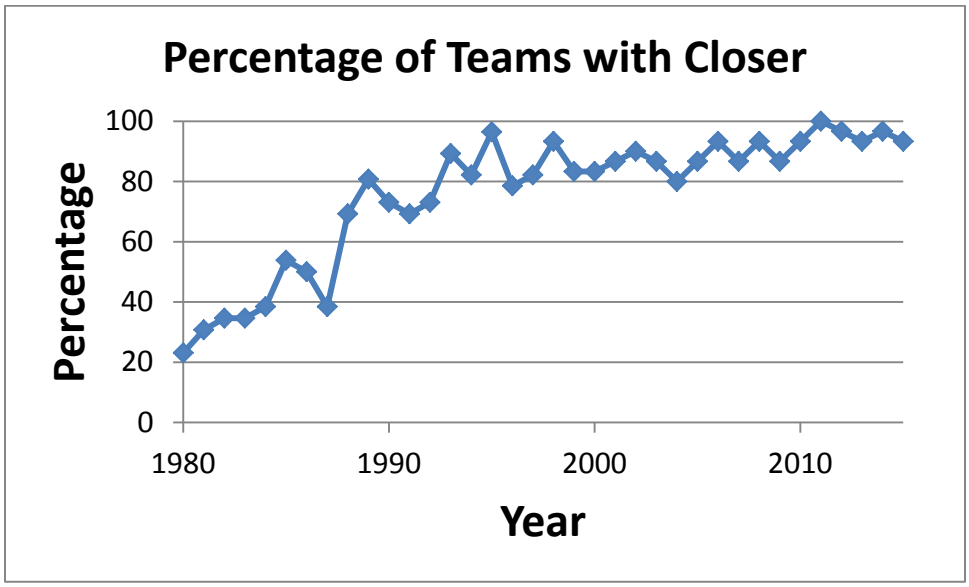
It is now clearer that there are two separate lines, with more variation in the games with the new pitcher. The average winning percentages over these years are 90.8 when a reliever starts the 9th inning and 91.2 when the same pitcher continues from the 8th, whether starter or reliever.

I expanded this result by separating the 9th inning relievers into "closers" and "non-closers". Of course, this requires a definition of closer and there are many ways to do that. I finally settled on the following two criteria:

1. Minimum of 40 relief appearances (pro-rated for strike years)
2. Minimum of 50% of appearances in save situations.

How many men met both of these criteria from 1980-2015?

Figure 5. Percentage of teams with closer, 1980-2015



The results are expressed as a percentage because there were different numbers of teams in this period, due to expansions. Based on these results, it looks like the "closer era" did not begin in earnest until around 1990 or even later. For example, in 1980, there were only six pitchers who met these criteria and 2011 is the only year that all the teams had a pitcher who did.

What about individual pitchers? Who were closers for the most seasons? There are 11 men who had at least 10 seasons as a closer by my definition. They are:

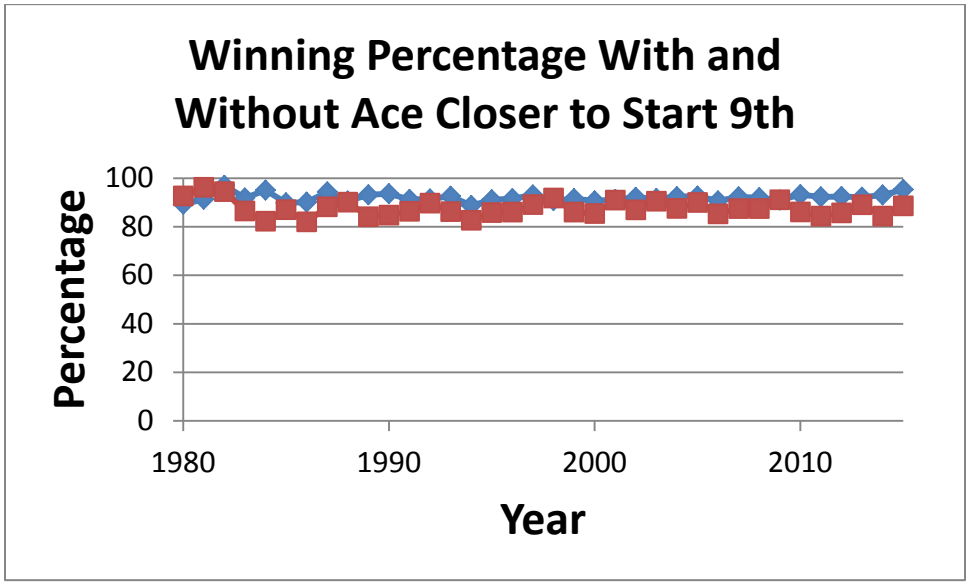
Table 1. Closers for most seasons, 1980-2015

Mariano Rivera	16
Trevor Hoffman	15
Lee Smith	13
John Franco	12
Billy Wagner	12
Dennis Eckersley	10
Roberto Hernandez	10
Troy Percival	10
Joe Nathan	10

Jonathan Papelbon 10
Huston Street 10

Using this definition of closer, I revisited the previous graph which showed a small benefit to not bringing in a new pitcher for the 9th inning. Figure 5 separates the 9th inning relievers into ace closers and others.

Figure 6. Winning percentage in relation to using an ace closer to start the 9th inning.



The blue line is the winning percentage when an ace closer enters and the red is the result when a someone else starts the 9th. There is an advantage to using the ace here, with a winning percentage difference of 92 vs 88.

All of my analysis to this point has addressed the effect of 9th inning pitcher usage on the chance of a team's winning the game, which is of course the ultimate purpose and should drive these decisions. However, there is another aspect to consider, which is the performance of pitchers in the 9th inning, which will also quite reasonably affect a manager's decision to change pitchers or not. I chose two parameters to examine the effectiveness of pitchers in the 9th inning

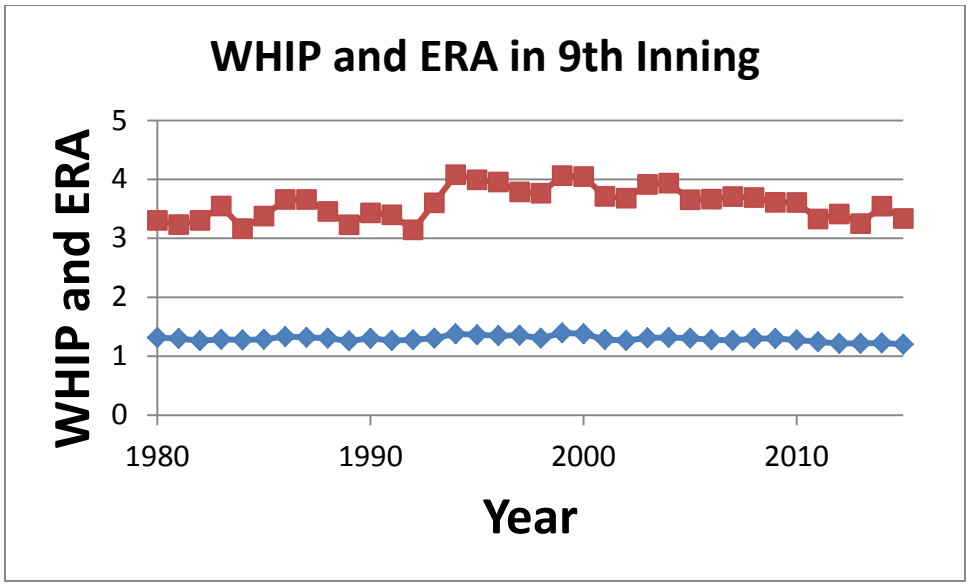
WHIP (walks plus hits per inning)
ERA

WHIP is more directly related to the individual pitcher, whereas ERA can depend on subsequent pitchers dealing with men left on base. These two measures were applied to three categories of pitcher:

1. Starters still in the game in the 9th.
2. Relievers not starting the 9th (may already be in game or enter during 9th).
3. Relievers starting the 9th.

First, let's look at the overall variation in WHIP (blue line) and ERA (red line) from 1980-2015, as shown in Figure 7.

Figure 7. WHIP and ERA for all pitchers in 9th inning, 1980-2015



ERA clearly varies more over these 36 seasons than WHIP does. Table 2 summarizes these results and compares them to overall performance in all innings.

Table 2. WHIP and ERA, 1980-2015

	WHIP	ERA
Overall	1.37	4.16
9th inning	1.29	3.60

Therefore, the 9th inning performance in these two measures has been is a bit better than the rest of the game, especially in ERA.

What about the three categories of pitchers I defined a moment ago? Table 3 summarizes these breakdowns.

Table 3. WHIP and ERA in 9th inning, 1980-2015

	WHIP	ERA
Starter	1.27	3.90
Reliever not entering to begin ninth	1.33	3.56
Reliever entering to begin ninth	1.28	3.54

There is a big variation in the results for the starters in recent years because so few starters are still pitching in the 9th and the sample size is therefore very small, leading to more noise in the data.

So both WHIP and ERA are lower for the pitchers selected to start the 9th compared to other relievers, but not dramatically so. In terms of WHIP the relievers are only slightly different from the starters, but there is a larger difference in ERA. I suspect the ERA difference arises in part from starters leaving the game with men on base who later score.

At this point, it is clear to me that the modern use of a closer to increase the chance of winning is a myth. Of course, it is actually much more complicated than that. The modern era of ubiquitous closer usage has other consequences. The well-known pattern for many teams is now:

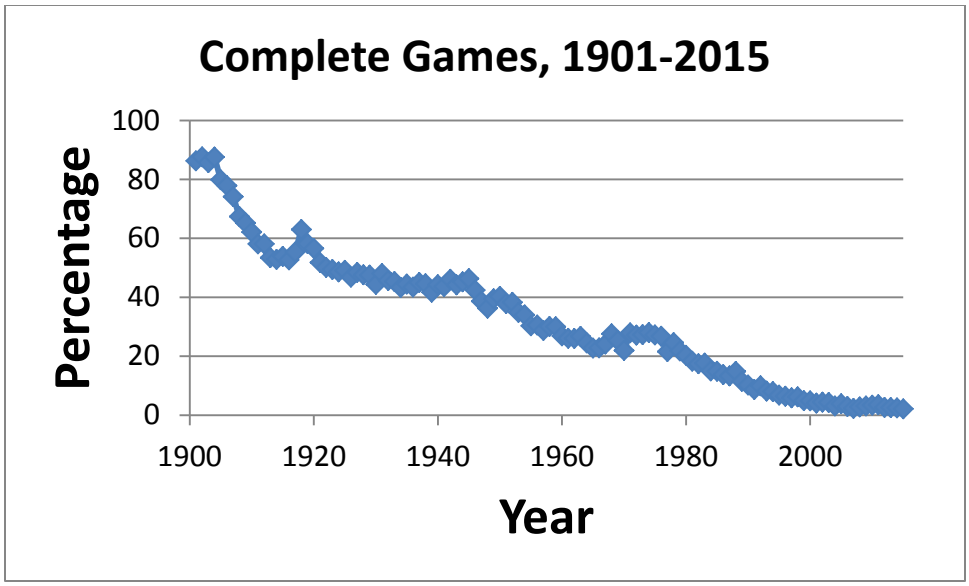
Starter pitchers 6 innings, hopefully leaving with a lead
7th Inning specialist enters
8th Inning specialist enters, also termed the "setup" man
Closer finishes off the win

Any given game will likely not follow that pattern, but it is enough of a template that teams organize their staffs around it. It is common now to refer to "the back end of the bullpen" which means all three of these specialists, not just the closer. There are two main reasons I hear offered for this new reality: 1) it results in fewer innings and pitches, especially for starters; and 2) each man in the bullpen "knows his role" and will presumably maximize his success with the appropriate mental preparation. I cannot speak to the mental aspect, although I am rather skeptical about it, but distributing the workload in a predictable way to lessen the physical strain makes sense to me.

How did we get to this pattern? After all, as I showed above, modern closer usage is a relatively recent phenomenon (20-25 years) and not all teams got there at the same time. I decided to review some long term trends which I have reported on before, as have others.

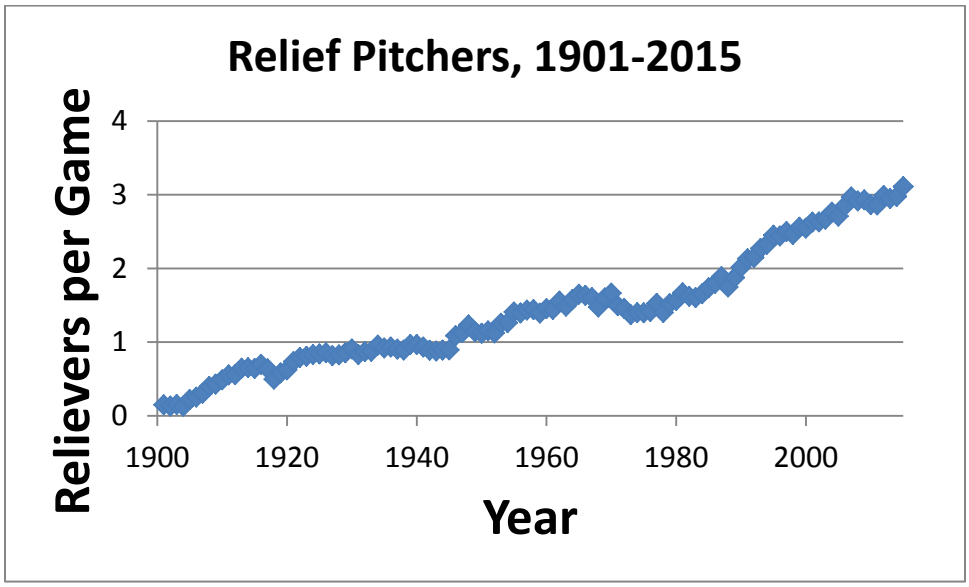
The percentage of complete games has dropped steadily since 1901, as shown in Figure 8.

Figure 8. Percentage of Complete Games, 1901-2015



There are a couple of interesting blips during both World Wars and the advent of the DH. The high point was 87.6% complete games in 1902 and the low is 2.1%, set in 2015. The corresponding piece of information is the number of relievers per game, as shown in Figure 9.

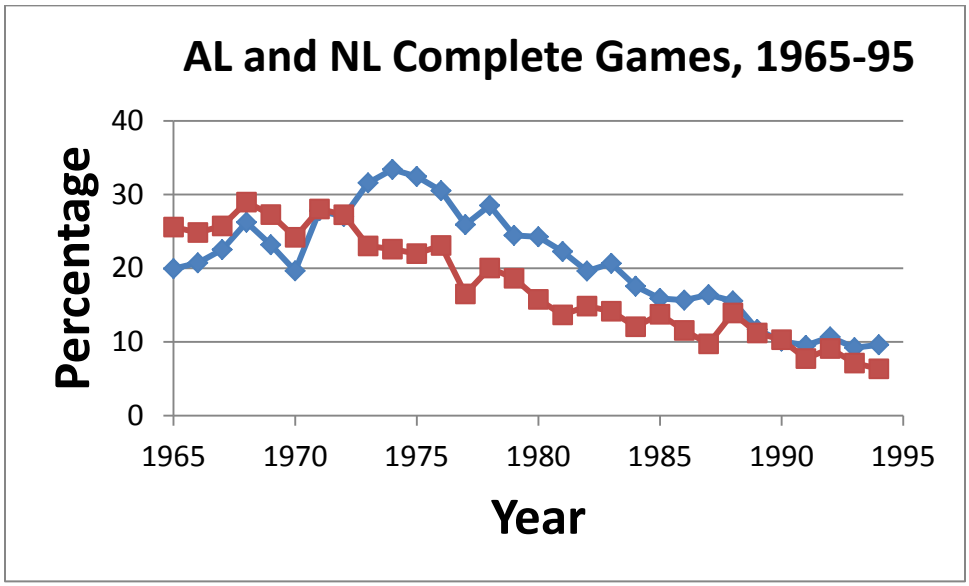
Figure 9. Relief Pitchers pre Game, 1901-2015



As expected, this graph is essentially the mirror image of the previous one, including the blips. The low value was 0.15 relievers per game in 1901, which is about one reliever for every 7 starts and the high value was 2015, when it reached 3.1 per game, the all-time high.

One of the fascinating aspects of this pattern concerns the advent of the DH in 1973. To illustrate this in more detail, I focused on the years 1965 to 1995. The percentages for each league are shown in Figure 10.

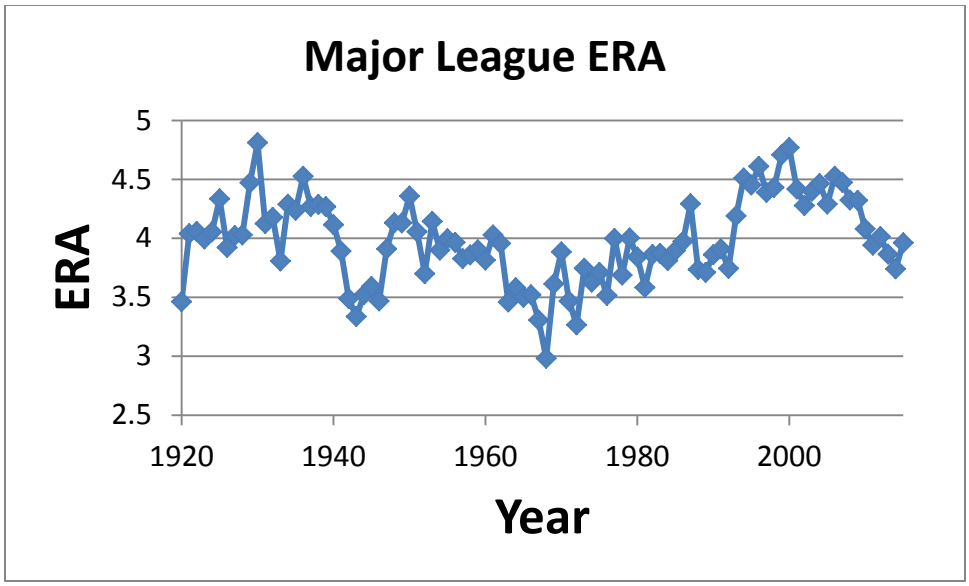
Figure 10. Complete Game Percentage by League, 1965-1995.



The AL values are in blue and the NL in red. From 1965-1972, the two leagues were similar in complete game percentage, but in 1973 the AL percentage jumped way up for the first time in over 20 years, while the NL dropped. It is easy to explain this as the AL managers were adjusting to the realities of having the DH, which was implemented in 1973. However over the next15 years, the gap continually narrowed and the two leagues are now nearly identical. This correlates nicely with the increased relief pitcher usage so that the different world for starting pitchers when there is a DH no longer carries over to the end of the game.

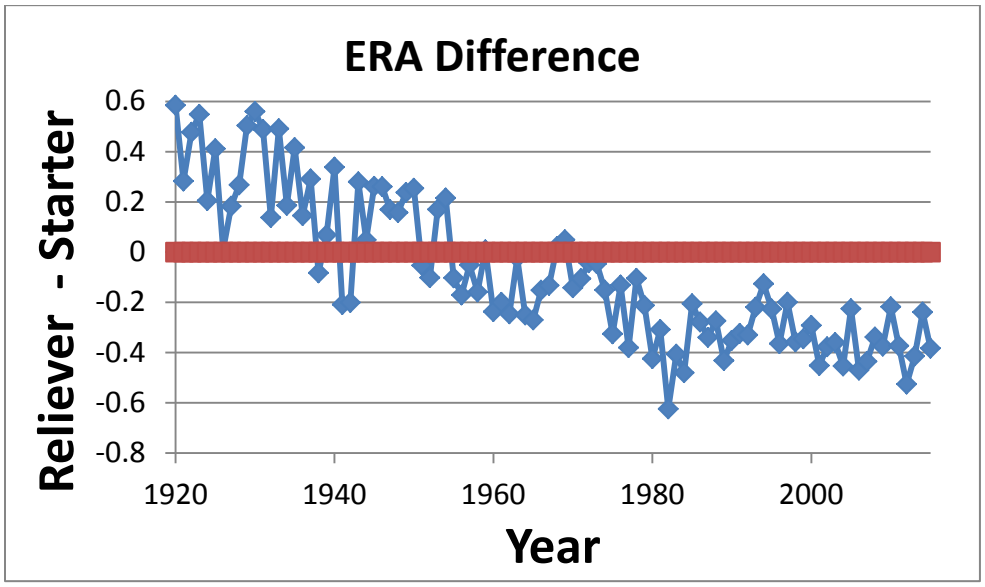
There are other meaningful measures for the shifting balance between starters and relievers, including ERA, innings pitched, and batters faced per game. ERA has varied a great deal during the last 96 seasons, as shown in Figure 11.

Figure 11. Major League ERA, 1920-2015



The high value of 4.81 came in 1930 and the low point of 2.98 occurred in 1968, the "Year of the Pitcher". The 2015 value of 3.96 is almost identical to the 96-year average of 3.98. Of course, my focus here is on relief pitchers so I calculated relief and starter ERA separately. Rather than present a graph with two lines that can be noisy and hard to interpret, I present in Figure 12, the difference between them, with the reported value being reliever ERA minus starter ERA.

Figure 12. Difference in Starter and Reliever ERA, 1920-2015

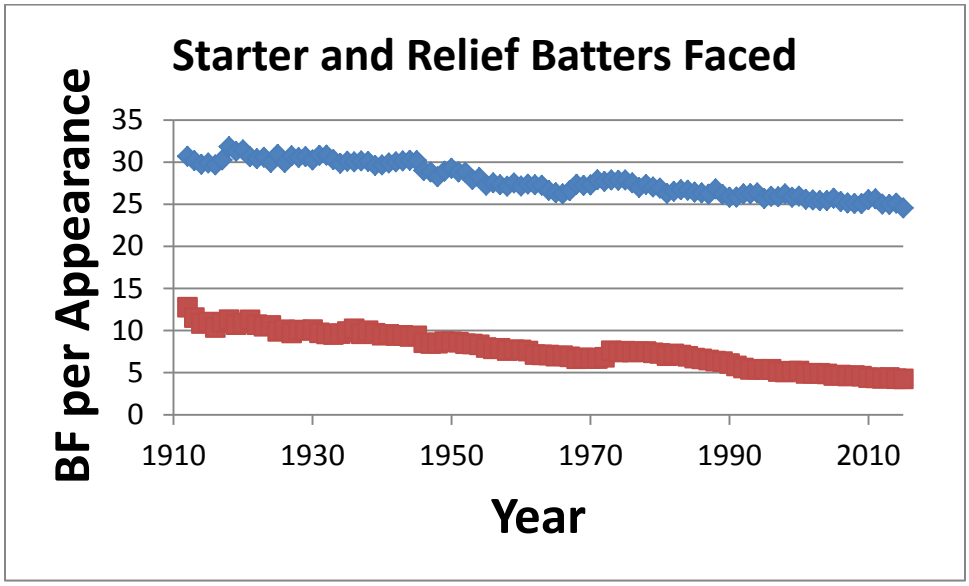


The heavy red line in the middle is at zero which would occur if starters and relievers had the same ERA. In fact, we see that with very few exceptions (1938-1941 and 1952-1953), from 1920 through 1953, the starters had the better ERA. Since 1954, the relievers have had better ERA every year except 1968 and 1969, when starters were lower by very small amounts: 0.02 and 0.05, respectively. This is a remarkable result and goes to the question of how individual

pitchers are chosen to be starters or relievers. My summary is that from 1920 to 1953, the pattern was that the starters had better performance by this measure and since then the balance has switched and relievers now have lower ERA, with recent values in the vicinity of 0.4 runs better, or nearly half a run a game. The biggest difference was 0.62 runs in 1982. This fits with the notion that relievers used to be those pitchers who weren't good enough to start, but now relievers are groomed (not just closers) for these roles. Perhaps they do well because they know they will have shorter stints, perhaps leading to higher velocity on a more consistent basis.

This pattern of progressively shorter outings is shown clearly in Figure 13.

Figure 13. Batters Faced per Appearance by Starters and Relievers, 1912-2015



The continuous decline is obvious with starters dropping from a high of 30.7 to the current low of 24.5 and relievers decreasing from 12.8 to 4.3. These lines are not strictly additive since as we saw in Figure 9, there are over three times as many relief pitchers per game in 2015 than in 1912 and the data in Figure 13 are calculated on a per appearance basis. Presumably the shorter outings may save wear and tear on pitchers' arms.

In addition, as I have reported before (Figure 11 in http://retrosheet.org/Research/SmithD/Batting%20Order%20Lineup2006.pdf), pitchers have a significant advantage against batters in their first matchup of each game. That is, batters "learn" during the game. In the last 25 years, the number of times a relief pitcher faces a given batter twice is vanishingly small, especially for those who pitch at the end of the game. This is a major contributing factor to the improved ERA for relievers.

Finally, I must address the statistical category of saves. It is common to hear announcers in the 8th inning explain that the closer will only enter in a save situation and that if the lead were to grow to four runs, then someone else would pitch the ninth. This situation is often accompanied by a view of the bullpen showing two pitchers warming up side by side. Of all the aspects of closer usage, this is the one that troubles me the most. The decision of which pitcher to bring in

is strongly determined by a statistical category, the save. I can think of no other example in which a decision is driven by something other than the chance to win the game. Cardinal manager Mike Matheny said: "The save stat cannot be ignored; there are contracts involved."

Others have addressed the use of closers: Bill James suggests that using an ace in the 7th inning in a crucial situation may be more important; I agree. Mike Emeigh believes that closers are more important in extra innings. Wayne Towers has a presentation tomorrow afternoon on the possible changing meaning of high save totals.

Conclusions

1. The entry of a new pitcher to start the 9th inning has increased dramatically since 1980.
2. The presence of this new pitcher has had almost no effect on a team's chances to win.
3. Ace closers bring slightly more wins than other 9th inning pitchers (92% vs 88%)
4. Performance of 9th inning pitchers is almost indistinguishable between closers and others.
5. Increased use 9th inning pitchers correlates with overall increase of relief pitchers.
6. Pitchers have had progressively shorter stints for over 100 years.
7. Current pattern of closer usage is not justified by their contributions to team wins.