

Saavedra, Serguei, Scott Powers, Trent McCotter, Mason A. Porter, and Peter J. Mucha (2010). Mutually-antagonistic interactions in baseball networks. *Physica A*, Vol. 389, pages 1131-1141.

Saavedra, Powers, McCotter, Porter, and Mucha (2010) concocted a statistically-sophisticated evaluation system based on the run potential for specific batter-pitcher matchups. They presented findings using all Retrosheet data between 1954 and 2008. The results of their model correlated almost perfectly (.96) with an index based on overall run potential.

Seidman, Eric (2008). Ballad of the fatigued: The effects of long innings.

<https://www.baseballprospectus.com/news/article/7641/ballad-of-the-fatigued-the-effects-of-long-innings/>

Seidman, Eric (2008a). Ballad of the fatigued: Controlled results, time, and release points. <https://www.baseballprospectus.com/news/article/7702/ballad-of-the-fatigued-controlled-results-time-and-release-points/>

Eric Seidman (2008) used PITCHf/x data to examine the pitch velocity and vertical movement consequences for 30 starters who threw at least 40 pitches in the first inning for the 2007 season plus up to May 20th in 2008 (based on a list compiled by Dave Smith which most certainly originated with Retrosheet data). For the rest of that inning, average pitch velocity remained the same although vertical pitch movement decreased starting around pitch 20 and continued to do so for the rest of the inning. What happened during the subsequent second through sixth innings depended on the pitchers' fastball dependency. Those who threw fewer than 27 fastballs in the first did not lose velocity and added some vertical movement; those who threw 27 or more fastballs in the first lost about 1½ mph in the second but no more but lost significant vertical movement. In a follow-up copied-and-pasted table, Eric compared groupings of these pitchers based on average velocity with their performance in analogous starts in which they threw 24 or fewer pitches in the first:

	Slow		Medium		Fast	
IP	40+	Control	40+	Control	40+	Control
1	86.54	86.98	90.35	89.95	92.05	92.12
2	86.27	87.25	88.87	90.14	91.16	92.34
3	86.56	86.77	89.20	89.97	90.81	92.03
4	86.54	87.05	88.80	89.72	90.79	92.27
5	84.99	86.39	89.37	89.97	90.39	92.48
6	84.26	87.32	88.92	89.76	N/A	92.22

As for horizontal and vertical pitch movement:

	Slow		Medium		Fast	
IP	40+	Control	40+	Control	40+	Control

1	8.55/ 9.06	8.33/9.45	5.58/ 9.13	6.69/8.08	6.81/9.19	6.49/9.09
2	8.24/ 9.21	7.79/8.79	5.91/ 8.49	6.62/7.93	5.80/9.21	6.61/8.99
3	9.30/ 9.13	8.31/9.32	7.03/ 7.91	5.97/8.43	6.50/8.81	6.59/8.94
4	7.91/ 9.89	8.14/8.75	5.51/ 9.57	6.56/8.09	7.88/8.59	6.61/8.96
5	8.72/10.71	8.21/8.84	5.53/ 9.11	6.56/8.86	9.17/7.85	6.59/9.14
6	8.99/ 9.14	8.01/9.14	6.08/10.08	6.29/8.27	N/A	6.66/8.90

Seidman, Eric (2009). On the swing.

<https://www.baseballprospectus.com/news/article/9841/checking-the-numbers-on-the-swing/>

Eric Seidman (2009) examined a total of 897 seasons between 1974 and 2009 from 598 pitchers who both started and relieved at least ten times in those seasons to compare their performance at each. Overall, as relievers, their Fair Run Average, chosen because it sidesteps the problems with assigning run responsibility between starters and relievers that ERA has, was 0.68 lower and their strikeouts per plate appearance .023 higher, with no difference in walks per PA. Dividing the population into power pitchers (K + BB per PA greater than 28 percent), finesse (the same less than 24 percent), and neutral pitchers, the finesse group was a bit more advantaged as relievers (FRA 0.76 lower) than neutral (0.53) and power (0.52).

Seidman, Eric (2009). Attack of the finesse pitchers: Strategy [sic] and arms control.

<https://www.baseballprospectus.com/news/article/8525/attack-of-the-finesse-pitchers-strategy-and-arms-control/>

Eric Seidman used 2008 strikeout and walk data from Retrosheet to divide pitchers into finesse, power, and neutral categories (without detail on how the division was made, although in comments later he said finesse pitchers' K + BB per PA was lower than 24% of PAs); and slugging average to separate batters into power, contact, and average. These data were combined with what PITCHf/x data was then available. Here are some general numbers for each type

Pitcher	FB%	Velocity	Movement	FB%	CU%	SL%	CH%
Finesse	56.4	89.92	6.60/7.96	56.4	9.8	15.1	13.6
Neutral	55.3	90.53	6.27/8.38	55.3	11.9	14.1	12.1
Power	61.3	92.36	5.87/9.21	61.3	9.5	15.6	8.2

Power pitchers threw more fastballs of greater velocity, with more vertical but less horizontal movement than finesse pitchers, with neutral pitchers intermediate. On average, neutral pitchers had the greatest variety, and finesse pitchers threw more changeups and fewer fastballs than power pitchers.

Now look at pitcher/batter interactions:

Pitcher	Hitter	FB%	Velocity	Movement	C%	OOZ%
---------	--------	-----	----------	----------	----	------

Finesse	Contact	57.6	89.84	6.59/7.99	19.1	20.8
Finesse	Average	55.9	89.81	6.57/7.92	20.1	22.3
Finesse	Power	54.3	90.36	6.74/7.95	19.9	23.4

Neutral	Contact	57.5	90.49	6.32/8.39	18.2	21.3
Neutral	Average	55.3	90.52	6.27/8.43	19.1	23.2
Neutral	Power	53.8	90.96	6.19/8.35	19.6	23.8
Power	Contact	63.5	92.31	5.87/9.22	18.1	21.9
Power	Average	60.4	92.29	5.89/9.20	18.3	23.3
Power	Power	58.8	92.54	5.85/9.19	18.4	24.7

All pitches threw fewer but faster fastballs against power hitters. Pitch movement was not affected by batter type. Finesse pitchers hit the corners of the plate (C%) more often. All of them pitched outside of the rule book strike zone (OOZ%) more often against power hitters.

Seidman, Eric (2010). Drilling down on volatility and consistency.

<https://www.baseballprospectus.com/news/article/10005/checking-the-numbers-drilling-down-on-volatility-and-consistency/>

Based on 1974 to 2009 Retrosheet data, Eric Seidman (2010) noted that predicted ERAs for pitchers with at least 20 starts in four consecutive seasons who rated in the upper fourth and upper fifth in consistency in ERA across those seasons tended to be quite accurate, whereas those in lower fourth and fifth, i.e. the most volatile, tended to outperform their projections by about a tenth of a run.

Seidman, Eric and Russell A. Carleton (2010). Side effects on pitchers' hitting.

<https://www.baseballprospectus.com/news/article/9932/checking-the-numbers-side-effects-on-pitchers-hitting/>

Eric Seidman and Russell Carleton (2010) took on the question of whether a pitcher having to bat or run the bases results in worse pitching the next inning. 2008-2009 PITCHf/x data for pitchers with at least 30 PA revealed a drop-off of 2.6 percent in fastball usage and 0.11 in fastball velocity, with curveballs, sliders, and changeups all taking up the slack. In other words, pitch variety increased. Fastball movements decreased by 0.05 inch horizontally and 0.10 inch vertically; changeups lose 0.23 vertical inches. Having to run the bases had an analogous impact; 1.7 percent of fastballs becoming others pitches and an even slighter (.05) decrease in velocity. In some contrast with batting only, fastball movement drops more horizontally (0.15 inches) than vertically (0.04 inches), with curveballs and changeups losing as much as ¼ inch of movement. Unfortunately, this comparison appears to be against both pitchers who batted and did not get on base and pitchers who did not bat; it would be more informative to have been limited to the first of these groupings. Turning to outcomes and based on PA for seasons for batters with and pitchers facing at least 250 PA (which seasons are not mentioned, but the data surely is from Retrosheet), and controlling for pitcher and batter strength and handedness and pitch count, Eric Seidman and Russell Carleton (2010) only uncovered a .004 decrease in strikeouts per PA, with most of those K's lost becoming outs-in-play, and a slight increase in hits going for extra bases rather than singles. So there are batting and baserunning effects for pitchers,

particularly in terms of pitch movement, but they seem to have minimal impacts on outcomes.

Sela, Rebecca J., & Jeffrey S. Simonoff (2007). Does momentum exist in a baseball game? In Jim Albert and Ruud H. Koning (Eds.), *Statistical thinking in sports* (pages 135-151). Boca Raton, FL: Chapman & Hall/CRC

Time to examine another myth; that momentum effects exist within an inning such as when a team starts getting baserunners and scoring runs, it is likely to continue. Sela and Simonoff (2007) began with a standard Markov table of transition probabilities between different base-out situations but added sets of variables via logistic regression allowing for a series of more complicated models. The sets respectively incorporated:

1 – player quality effects; batter on-base and slugging averages and pitcher WHIP and strikeouts per nine innings, plus whether the home or away team is batting.

2 – situational effects; the number of batters faced and pitches thrown by the current pitcher in the game and the OBA and SLG for the next batter in case “protection” was real.

3 – momentum effects, the issue at hand; the result for the previous plate appearance unless the current batter began the inning, and the number of batters and runs scored since the last out.

Using Retrosheet data from 2003 and 2004 for establishing the models and 2005 for validating them, the authors noted that addition of the player quality and situational effects resulted in more accurate modeling. However, puncturing the relevant myth, the only momentum effects uncovered were negative; with two runners on base and either one or two outs, there is a slight increase in the odds that outs begat more outs. Consistent with this result, negative binomial regressions indicated that, in those situations, average runs for the remainder of the inning for the team at bat were lower than expected given base-out situation and current batter and pitcher quality. The authors did find support for one myth; double plays really were rally killers, decreasing subsequent run scoring more than expected; the authors did not consider whether this finding was responsible for the “anti-momentum” effects.

Shamsie, Jamal and Michael J. Mannor (2013). Looking inside the dream team: Probing into the contributions of tacit knowledge as an organizational resource. *Organization Science*, Vol. 24 No. 2, pages 513-529.

There has been a lot of academic studies (mostly quite poor) examining the relationship between player and team performance. Somewhat more interesting is Shamsie and Mannor’s (2013) attempt to measure the impact of factors over and above those related to sheer player skill, using data from 1985 gleaned from the Lahman Archive and Retrosheet. Although they did use one factor indirectly related to skill, the number of game appearances for a team’s roster, the others included managerial experience both overall and with the relevant team, past playoff experience for manager and players, and three measures of team stability: the number of players with the team

for at least three years, game-to-game consistency in starting lineups, and maintaining the same manager during a season. Every included factor has a significant, although in some cases small, impact on team winning percentage.

Shu, Pei Zhe (2016). Arsenal/Zone Rating: A PITCHf/x based pitcher projection system. *MIT Sports Analytics Conference*.

Shu (2016) proposed a pitcher projection system that combines PITCHf/x data on pitch speed, movement, and location with Retrosheet play-by-play data. Based on 2008 to 2014 data, the author claimed accuracy comparable to other projection methods and more success at predicting breakout and breakdown seasons as measured by 33 percent increases and decreases in performance.

Sidran, D. Ezra (n.d.). A method of analyzing a baseball pitcher's performance based on statistical data mining.

https://www.researchgate.net/publication/267918769_A_Method_of_Analyzing_a_Baseball_Pitcher's_Performance_Based_on_Statistical_Data_Mining/link/5489c6cf10cf214269f1abc7f/download

Sidran (n.d.) proposed a method for computing a running score of pitcher performance using Retrosheet pitch-by-pitch data and based on a point system assigning -4 for an opposition homer, -3 for a triple, -2 for a double, -1 for a single, walk, ball, or stolen base, and +1 for a strike, foul ball, and out on ball-in-play. This point system is obviously flawed, but given a correct weighting for the components the idea may have some value.

Silver, Nate (2003). Leading off.

<https://www.baseballprospectus.com/news/article/2149/lies-damned-lies-leading-off/>

Here are figures showing how OBA became more centralized to #3 and #4 batters over time.

Order #	1982-1989	1999-2000	2001-2002
1	.336	.349	.332
2	.333	.346	.331
3	.349	.384	.379
4	.345	.375	.368
5	.329	.356	.338
6	.322	.345	.327
7	.315	.326	.318
8	.308	.329	.312

Silver, Nate (2003). Hitting the wall.

<https://www.baseballprospectus.com/news/article/2128/lies-damned-lies-hitting-the-wall/>

Do batters do better or worse after July 1st (second half of season)? Does age impact on this? The following was based on 1999 through 2001:

Improvement (Decline) After July 1					
Age	n	BA	OBP	SLG	OPS
21	15	+.018	+.021	+.030	+.051
22	41	-.002	+.004	-.008	-.004
23	98	+.007	+.001	+.015	+.016
24	141	-.005	-.009	-.015	-.024
25	167	-.003	+.000	-.011	-.011
26	168	-.003	-.005	-.003	-.008
27	180	+.001	+.001	-.007	-.006
28	181	-.012	-.009	-.018	-.027
29	166	-.006	-.003	-.022	-.025
30	143	-.009	-.010	-.025	-.035
31	141	-.013	-.012	-.030	-.042
32	117	+.001	-.002	-.013	-.015
33	95	-.003	+.001	-.009	-.008
34	94	-.009	-.008	-.022	-.030
35	73	-.001	+.002	-.021	-.019
36	53	+.001	-.002	+.001	-.001
37	29	-.022	-.025	-.050	-.075
38	14	+.001	-.009	-.013	-.022
Young'uns	(21-24)	+.001	-.002	-.002	-.004
Mid-Career	(25-29)	-.006	-.004	-.015	-.019
Veterans	(30-33)	-.007	-.007	-.021	-.027
Old'uns	(34-38)	-.006	-.006	-.020	-.026

How about pitchers?

Improvement (Decline) after July 1						
Age	n	BA	OBP	SLG	OPS	K Rate
21	11	+.005	+.035	-.024	+.011	-3.3%
22	36	+.019	+.014	+.036	+.050	-1.3%
23	69	+.004	+.002	+.004	+.006	+0.8%
24	100	+.003	-.002	-.008	-.010	+0.0%
25	101	+.007	+.003	+.011	+.014	-0.1%
26	114	-.001	-.003	-.018	-.021	+0.4%
27	113	+.000	-.006	-.011	-.017	+0.7%
28	111	+.011	+.007	+.017	+.024	-0.2%
29	96	+.001	-.007	+.002	-.005	-0.2%
30	97	-.008	-.009	-.006	-.015	+0.4%
31	88	-.002	-.007	-.007	-.014	+0.1%
32	83	+.007	+.009	+.009	+.018	-0.7%
33	69	-.007	-.007	-.023	-.030	-0.6%
34	50	+.003	+.001	-.001	+.000	+0.0%
35	33	+.002	-.004	-.009	-.013	-1.3%
36	26	+.016	+.019	+.015	+.034	-0.6%
37	23	-.006	-.006	-.022	-.028	-0.3%
38	19	-.009	-.009	-.040	-.049	+0.2%

Young'uns	(21-24)	+ .006	+ .004	+ .002	+ .006	-0.1%
Mid-Career	(25-29)	+ .004	- .001	- .000	- .002	+0.2%
Veterans	(30-33)	- .003	- .004	- .006	- .010	-0.1%
Old'uns	(34-38)	+ .002	+ .001	- .008	- .007	-0.4%

In interpreting these figures, I suggest trusting the last four rows of each as the greater sample sizes iron out random variation.

Silver, Nate (2003). Redefining replacement level.

<https://www.baseballprospectus.com/news/article/2032/lies-damned-lies-redefining-replacement-level/>

Nate Silver argued that the concept of replacement level as commonly understood is problematic, because the quality of the player that a team would use as a replacement is dependent on how long the replacement needs to be used for. The longer the time the replacement player is needed, the better the player required. The following figures are for Batting Runs per PA based on career PA (1973-1992). Note that they increase very quickly until 50 PA and then ever more slowly afterward.

Min PA	Max PA	n	Mean PA	BR/PA
1	5	53	3.8	-0.1268
6	10	41	8.6	-0.0812
11	20	46	16.7	-0.0774
21	30	49	26.3	-0.0553
31	50	48	41.2	-0.0587
51	70	54	60.7	-0.0343
71	110	49	91.6	-0.0342
111	150	51	130	-0.0354
151	200	39	176	-0.0291
201	300	59	253	-0.0283
301	400	44	355	-0.0277
401	600	58	506	-0.0253
601	900	49	738	-0.0196
901	1200	50	1020	-0.0175
1201	1600	52	1409	-0.0096
1601	2200	49	1902	-0.0083
2201	3000	45	2595	+0.0015
3001	4200	48	3536	+0.0061
4201	5500	49	4835	+0.0102
5501	10184	24	7271	+0.0231

The relationship is thus curvilinear and can only be represented arithmetically by a logarithm. Here is his equation:

$$BR/PA = 0.0154 * \ln(PA) - 0.117$$

Here is an equation based on the one just above that defines a replacement level as a function of career PA. He called it Progressive Runs Above Replacement (PRAR).

$$\text{PRAR} = \text{BR} - \text{PA} * (.0154 \ln(\text{PA}) - .1324)$$

The replacement level player defined this way produces about 76 percent as many runs as the average player.

Silver, Nate (2003). Batter vs. pitcher matchups.

<https://www.baseballprospectus.com/news/article/1986/lies-damned-lies-batter-vs-pitcher-matchups/>

Nate Silver used 2002 Retrosheet data to break down batter/pitcher matchups by “power” vs. “finesse,” based on the square root of (walk rate X strikeout rate). Finesse was defined as .10 or less, power as .14 or more, and neutral as between the two, resulting in about one-third of the players in each of these three categories. First, the all batters vs. the two types of pitchers,

Pitcher	Power	Finesse	All Pitchers
BA	.241	.276	.261
OBA	.339	.327	.333
SLG	.387	.434	.417
KRate	21.9%	13.3%	17.1%

which shows that batters hit better and walk less (compare BA and OBA) against finesse pitchers, and all pitchers vs. the two types of batters,

Batter	Power	Finesse	All Batters
BA	.257	.265	.261
OBP	.361	.312	.333
SLG	.455	.390	.417
KRate	22.4%	13.6%	17.1%

which shows that pitchers give up more extra base hit and walks to power hitters. No surprises here. Now for some further breakdowns, comparing the data with matchup predictions based on (I am guessing) Dallas Adams' log5 method. First, power pitchers vs. finesse batters:

Power Pitcher v Finesse Batter, 2002

	Actual	Expected
BA	.244	.244
OBP	.311	.318
SLG	.362	.361
KRate	18.0%	17.7%

Finesse batters do poorly against power pitchers, but no worse than would be expected given their overall performance. Next, finesse pitchers vs. power batters

Finesse Pitcher v Power Batter, 2002

	Actual	Expected
BA	.278	.271
OBP	.356	.355
SLG	.487	.473
KRate	17.6%	17.7%

Better performance than just above, and perhaps a bit more production than what would be expected, Next, power pitchers vs. power batters

Power Pitcher v Power Batter, 2002

	Actual	Expected
BA	.239	.236
OBP	.371	.367
SLG	.429	.425
KRate	27.6%	28.2%

Nothing noteworthy here. Finally, finesse pitchers vs. finesse batters

Finesse Pitcher v Finesse Batter, 2002

	Actual	Expected
BA	.280	.279
OBP	.313	.306
SLG	.407	.407
KRate	10.2%	10.5%

The same. Overall, there is no evidence for platooning based on this sort of matchup.

Silver, Nate (2003). Pitcher vs. hitter matchups (Holes part deux).

<https://www.baseballprospectus.com/news/article/1936/lies-damned-lies-pitcher-vs-batter-matchups-holes-part-deux/>

A follow-up to the above, using 2000-2002 data. The question here is whether increasing the number of batter/pitcher matchups has any effect on long-time performance. There is a selection bias precluding a simple analysis of performance with number of times faced one another, because longer careers for each mean more matchups, and better players have longer careers. As a consequence, weaker players will drop out of the sample, meaning that overall means would wrongly imply that players improve as matchup PAs increase. Nate did a more sophisticated analysis comparing actual with expected outcome specific to that matchup. When you do, you get no consistent effect either way.

Silver, Nate (2004). Groundballs in the mix.

<https://www.baseballprospectus.com/news/article/2885/lies-damned-lies-groundballs-in-the-mix/>

The following is copied and pasted from the webpost:

This is why I assert ... that groundball ratio is a better predictor of home run rate than is home run rate itself. I looked at league- and park-adjusted statistics for all pitchers from 1975 onward who faced at least 500 batters in two consecutive seasons (1975 is the year in which reliable groundball-flyball data begins to be available from [Retrosheet](#)):

1. The correlation between home run rate in year N and home run rate in year N-1 is .326 (note that it is a little bit higher than in the previous example since we've increased the batters faced threshold).
2. The correlation between home run rate in year N and groundball rate in year N-1 is $-.345$. Though the sign proceeding the correlation figure is negative (since a *higher* groundball ratio tends to predict a *lower* home run rate), the magnitude of the correlation is a bit higher.

Of course we can do better still if we account both for home run rate and for groundball rate in the previous season. A simple regression model that uses home run rate in year N as the dependent variable, and home run rate in year N-1 as the independent variable, is capable of explaining only about 11% of the variance in home run rate for the sample of pitchers we've taken above. If groundball rate in year N-1 is included as a second independent variable, the explanatory power increases sharply to 16%. We can get up closer to 20% if we include other factors like strikeout rate and walk rate (and do considerably better than that if we look at three years worth of previous seasons data, as PECOTA does)—but all the while, groundball rate maintains the largest influence on predicting home runs allowed.

Silver, Nate (2003). Solving a ninth inning quandary.

<https://www.baseballprospectus.com/news/article/1963/lies-damned-lies-solving-a-ninth-inning-quandary/>

The one time when an intentional walk might make sense is ninth inning, tie game, home team at bat, runner on third – in other words, when the runner on third scoring means a loss. Nate used 1982-1992 and 1999-2002 (most likely) Retrosheet data.

With less than 2 outs, here is the relevant data:

THNT, <2 Outs	H	BB	K	HBP	SF	DP	OUT
3rd	30.7%	10.6%	15.3%	2.1%	11.1%	0.5%	29.6%
1st/3rd	30.6%	7.0%	11.5%	2.2%	8.9%	6.4%	33.4%
Loaded	30.8%	4.6%	15.1%	0.9%	10.5%	8.4%	29.6%

“Outs” do not include double plays and or sac flies. Walks are unintentional.

Here are outcomes for walking a batter with a runner on third

THNT, Runner on 3rd, 1 out

Visiting win = 0.5% (DP)

Home win = 41.8% (H, SF)

Runners and 1st and 3rd, 1 out = 12.7% (BB, HBP)

Runner on third, 2 outs = 44.9% (K, OUT)

THNT, Runners on 1st and 3rd, 1 out

Visiting win = 6.4% (DP)

Home win = 39.5% (H, SF)

Bases Loaded, 1 out = 9.2% (BB, HBP)

Runners on 1st and 3rd, 2 out = 44.9% (K, OUT)

Looks like an intentional walk is a good idea – cuts down the odds of the home team winning. But walking a second batter looks bad; same for walking one with first and third already occupied:

THNT, Bases Loaded, 1 out

Visiting win = 8.4% (DP)

Home win = 46.8% (H, SF, BB, HBP)

Bases loaded, 2 outs = 44.7% (K, OUT)

The two out situation is different. Here are performance data:

THNT, 2 Outs	H	BB	K	HBP	OUT
3rd	18.4%	14.2%	18.8%	0.4%	48.3%
1st/3rd	23.0%	6.5%	14.9%	0.6%	55.1%
Loaded	20.7%	11.8%	15.1%	0.4%	52.0%

Overall, this has been much better for the away team. Here are the strategy outcomes:

THNT, Runner on 3rd, 2 outs

Visiting win = 67.1% (K, OUT)

Home win = 18.4% (H)

Runners on 1st and 3rd, 2 outs = 14.6% (BB, HBP)

THNT, Runners on 1st and 3rd, 2 outs

Visiting win = 70.0% (K, OUT)

Home win = 23.0% (H)

Bases Loaded, 2 outs = 7.1% (BB, HBP)

THNT, Bases Loaded, 2 outs

Visiting win = 67.1% (K, OUT)

Home win = 32.9% (H, BB, HBP)

Walks look bad for the visiting team.

Silver, Nate (2003). Estimating pitch counts.

<https://www.baseballprospectus.com/news/article/1823/lies-damned-lies-estimating-pitch-counts/>

Using 2001-2002 probably Retrosheet data, Nate's regression analysis led to the following pitch count estimator:

$$(3.17 \times \text{BFP}) + (3.44 \times \text{BB}) + (1.53 \times \text{K})$$

With a described impact by ground-ball/flyball ratio as higher ones tend to be associated with fewer pitches. But he didn't include it in the equation due to the difficulty at the time getting the data. Pitch counts should have increased over time because K rates had increased (Nate said BB rates also but that was false) and GB/FB ratio had decreased. Incidentally, a slight negative correlation occurred between pitches per batter and hit rate per balls in play, -0.10 .

Silver, Nate (2004). Making RBIs useful.

<https://www.baseballprospectus.com/news/article/2818/lies-damned-lies-making-rbis-useful/>

Nate proposed a useful RBI stand-in, which he called Context-Independent Run Batted In (CIRBI). It is as follows:

((Percentage of runners on third driven in multiplied by league average for that) +
((Percentage of runners on second drive in multiplied by league average for that) +
((Percentage of runners on first driven in multiplied by league average for that)
Multiplied by number of plate appearances
Plus homers

Note that it first provides a proportion of base runners driven in that is indeed independent of the presence or absence of the number of opportunities to do so, which is beyond the batter's control, weighted for batter opportunity as measured by PA. It then adds the run that homers contribute, which is under the batter's control.

The following was the 2003 leaderboard, which is instructive:

Player	CIRBI	RBI
Delgado_Carlos	138	145
Pujols_Albert	131	124
Sheffield_Gary	131	132
Rodriguez_Alex	128	118
Helton_Todd	124	117
Thome_Jim	124	131
Sexson_Richie	122	124
Wilson_Preston	120	141
Wells_Vernon	120	117

Lee_Carlos	117	113
Anderson_Garret	117	116

Note that the CIRBI can be interpreted the same way as RBI, and that a couple of players' figures, particularly Preston Wilson, were affected by the de-contextualization.

Silver, Nate (2004). Using the Golden Run Ratio.

<https://www.baseballprospectus.com/news/article/3559/lies-damned-lies-using-the-golden-run-ratio/>

Silver, Nate (2005). Introducing ORVY.

<https://www.baseballprospectus.com/news/article/4003/lies-damned-lies-introducing-orvy/>

Here are runs scored by one club in an inning for 2003 (copied and pasted):

Runs Scored	Frequency	Percent
0	30922	71.1%
1	6845	15.7%
2	3011	6.9%
3	1507	3.5%
4	670	1.5%
5	305	0.7%
6	117	0.3%
7	62	0.1%
8	12	0.0%
9	6	0.0%
10	6	0.0%
11	0	0.0%
12	1	0.0%
13	1	0.0%
14	1	0.0%

Note that it is a very neat exponential decay function such that the ratio between 5 and 6, 4 and 5, etc. down to 1 and 2 are pretty close, with that for 0 and 1 about twice as big as the others. Nate called the relationship the Golden Run Ratio (g), and learned that it is greater for lower-scoring teams. Nate computed a couple of g's; 4.33 for 5 runs per game, 5.64 for 3 runs per game. In his 2005c, Nate used these figures to compute win probabilities for given moments in the game. Here is, using Nate's example, the probabilities of the home team winning a game following the bottom of the seventh inning:

Score	Home Win %
+5 runs	98.2%
+4 runs	96.3%
+3 runs	92.6%
+2 runs	86.0%
+1 runs	74.1%

Tied	50.0%
-1 runs	25.9%
-2 runs	14.0%
-3 runs	7.3%
-4 runs	3.7%
-5 runs	1.8%

As you can see, the probabilities differ very little with large run surpluses or deficits but quite a bit with small ones, which reflects differences in leverage. One can use these figures to compute the change in win probability if a team scores a given number of runs in an inning. To continue the example, with the score tied, one extra run would increase win probability by 74.1 minus 50 or 24.1 percent, scoring a second run by 86 minus 74.1 or 11.9 percent, and so on. Nate then introduced One-Run Value Yield (ORVY). The ratio of the first increase by the second; in this case, 24.1 minus 11.9 or 2.02. The higher the ORVY, the more valuable one-run strategies (sacrifice bunts) are relative to multiple runs. In this circumstance, a one-run strategy would be a good choice. ORVY has the following implications:

- 1 – The later in the game, the higher the ORVY, so the more valuable one-run strategies are compared to multi-run.
- 2 – When a team is one run behind, its ORVY will always be 1. This is then the break-even point for scoring one run versus two.
- 3 – A team should never use a one-run strategy when trailing by more than one run, because the ORVY is too small. Again, using that chart, ORVY when two runs back would be 11.9 divided by 24.1, which is 0.49. Note that this is the inverse of the first example, which demonstrated that a one-run strategy would be good for a team two runs ahead.

Smith, David W. (2006). Does walking the leadoff batter lead to big innings? *Baseball Research Journal*, No. 35, pages 23-24.

Our fearless leader's papers are customarily posted on the Retrosheet research page, but this one is not. In answer to a baseball myth expounded often (and inspired by one of those expositions by Tim McCarver), Dave showed that, from 1974 to 2002, walks to leadoff batters have the same impact on scoring as any other way to get on base, in so doing adding another piece of evidence to others showing that all ways of getting on base have equivalent impacts.

Smith, Erin E. and Jon D. Groetzinger (2010). Do fans matter? The effect of attendance on the outcomes of Major League Baseball games. *Journal of Quantitative Analysis in Sports*, Vol. 6 Issue 1 Article 4.

The most strongly supported explanation for the consistent 54% home field advantage for baseball is the impact of fan support. In one piece of relevant evidence Smith and Groetzinger (2010) combined data for the years 1996 through 2005 from the Retrosheet and Baseball Archive databases with weather information from the National

Climatic Data Center, along with the Questec pitch monitoring system for 2001 and 2002. Overall, increasing attendance by one standard deviation (about 25 percent) resulted in what the authors say was .64 additional runs (I wonder if they really meant run differential) and an increase of 5.4% in the probability of a home team. Hits, doubles, and home runs were all weakly but positively related with attendance, and earned runs allowed negatively associated. In addition, there was a decrease in home team strikeouts as attendance rose, which could signal home plate umpire bias in calling balls and strikes. However, contrasting ballparks with and without the QuesTec system for reviewing umpire ball-strike calls under the questionable assumption that umpires are biased by fan support but the presence of the system would decrease that bias; they could not find any differences.

Smith, Sean (2009). Total Zone data. https://www.baseball-reference.com/about/total_zone.shtml

Sean Smith's (2009) TotalZone uses Retrosheet data to evaluate fielders, with the type of available data determining the exact method. When data on specific plays is missing, Sean would do the following:

Step 1 – Compute every batter's career proportion of batted balls for which plays were made at each fielding position. If the batter is a switchhitter, then do this separately for left- and righthanded plate appearances.

Step 2 – Assume that this proportion remains the same for hits, and based on how often the batter and each fielder play against one another, estimate how many hits each fielder should be assessed based on that proportion.

Step 2 – For every fielder, sum the result of Step 2 across all batters played against.

Step 3 – Divide the results of Step 2 by every fielder's total fielding chances, computed by

$$(\text{Total plays made}) + (\text{Errors}) + (\text{Result of Step 2})$$

Step 4 – Do park adjustments, and convert to runs responsible for (.75 per hit for middle infielders, .80 for the infield corners, and .85 for outfielders).

When batted ball type and fielder is available, one can estimate responsibility for hits somewhat more accurately, by giving third basemen 60 percent and shortstops 40 percent of the debit for singles to right, shortstops 52 percent and second basemen 48 percent of the charge for singles to center, and first basemen 55 percent and second basemen 45 percent of the deduction for singles to right. Groundball extra base hits are presumed to be down the lines and so totally given to the corner infielders. The plays that fielders make and do not make can be compared to league average for different batted ball types and pitcher/batter handedness. I assume that outfielders would be judged based on proportion of relevant plays made.

When actual hit location is available, one can use that without making any estimates. Outfielder arms, infielder double plays, and catcher performance are also evaluated; see the referenced article on these.

Smith, Sean (2010). Relievers yesterday and today. In Dave Studenmund (Producer), *The Hardball Times Baseball Annual 2010* (pages 176-182). Skokie, IL: Acta Sports.

Using Retrosheet data and limiting analysis to seasons in which they achieved a WAR of at least 3.0, Sean Smith (2010) examined changes in elite relief pitcher usage beginning about when relief specialists became prevalent (1954) and ending in 2008. First and foremost, although the average number of appearances for top relievers has stayed about constant at 65 during this period, the average number of innings pitched has substantially dipped from about 115 to 125 until 1984 down to the current 75 or so. Second, as greater workload allows for higher WAR, this decrease resulted in only one of the top twelve WAR seasons (Mariano Rivera, 1996, 5.4 WAR) occurring after 1986. Third, as fewer innings means less overuse and more staying power, the likelihood of a reliever following up a 3.0+ WAR season with one at least at 1.0 has increased from only 50 to 60 percent through 1980 and close to 70 percent since. Thus, several indicators suggest the early-mid 1980s as a breakpoint between the 2 and the 1 inning closer. Interestingly, Leverage Index only increased a bit, from for example 1.58 from 1954 to 1969 to 1.77 in 2005 to 2008.

Smith, Sean (2011). Do catchers have an ERA? In Dave Studenmund (Producer), *Hardball Times Baseball Annual 2011* (pages 143-146). Chicago, IL: Acta Sports.

In his book with Tom House, *Diamond Appraised*, Craig Wright introduced the idea of Catcher ERA, in which a catcher is evaluated according to whether the ERA of his team's pitching staff is better or worse when he is behind the plate as compared with his team's other catchers. Sean Smith (2011) examined the consistency across seasons using Retrosheet as part of the data source and analyzing these data via the "matched inning" prorating method Craig used. In order to neutralize differences in team fielding applying a DIPS-based bottom-up estimate of runs allowed that he concocted rather than the actual total to neutralize differences in fielding. He observed a .21 correlation across consecutive seasons starting with 2003 and ending with 2009 for 70 catchers with at least 2000 "matched" plate appearances (this would double count PAs for 2004 to 2008 as each of those seasons would be included twice). This implies some but not a lot of consistency across seasons in specific catcher's relative ranking. He also noted no staff ERA improvement as catchers gain experience, inconsistently with some who had found some (Tom Hanrahan, the Hirdt brothers in the 1981 *Baseball Analyst* book) and consistently with others (Keith Woolner in the 1999 *Baseball Prospectus*).

Song, Alex, Thomas Severini and Ravi Allada (2017, February 7). How jet lag impairs major league baseball performance. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 114 No. 6, pages 1407-1412.

Carleton, Russell A. (2017). Blame it on the plane.

<https://www.baseballprospectus.com/news/article/31079/baseball-therapy-blame-it-on-the-plane/>

Song, Severini and Allada (2017) replicated earlier claims about the impact of jet lag on home field advantage, based on home teams using 1992-2011 data (likely from Retrosheet). In fact, the home field advantage was nullified for teams returning home west-to-east through either two or three time zones when the visiting team had stayed in the same time zone; the analogous effect for home teams traveling east-to-west also occurred but more weakly. Home-team slugging average, and even more specifically number of doubles hit, were affected identically, as were slugging average by opposing team, runs allowed, and fielding-independent pitching, the latter two due to giving up more home runs. Visiting teams were also affected negatively by travel, although direction did not matter. Displeased with Song et al. averaging across seasons and players within teams in their analysis, Russell Carleton (2017) duplicated the study at the plate appearance level using 2012 through 2016, with time lags considered significant if either two or three hours. He got several significant findings across different types of game events but none were consistent throughout.

Soto Valero, César (2016). Predicting win-loss outcomes in MLB regular season games: A comparative study using data mining methods. *International Journal of Computer Science in Sport*, Vol. 15 No 2, Article 7.

Soto Valero (2016) compared the capability of data mining methods as predictors of game outcomes using Retrosheet data for 2005 through 2014.

Spearing, Harry, Jonathan Tawn, David Irons, and Tim Paulden (2023). Modeling intransitivity in pairwise comparisons with application to baseball data. *Journal of Computational and Graphical Statistics*, Vol. 32 No. 4, pages 1-19.

The authors propose a method for ranking teams. For testing it, they predicted seasonal team rankings based on the winner of each season series, using 2010-2018 Retrosheet data.

Stoll, Greg. Expected runs per inning.

<https://gregstoll.com/~gregstoll/baseball/runsperinning.html>

The title says it all, computed from Retrosheet data for 1957 through 2023, from 0 to 17 runs (apparently, 15, 16, and 17 were each achieved once during that period).

Swartz, Matt and Eric Seidman (2010). Introducing SIERA: Part 1.

<https://www.baseballprospectus.com/news/article/10027/introducing-siera-part-1/>

Swartz, Matt and Eric Seidman (2010). Introducing SIERA: Part 2.

<https://www.baseballprospectus.com/news/article/10032/introducing-siera-part-2/>

Swartz, Matt and Eric Seidman (2010). Introducing SIERA: Part 3.

<https://www.baseballprospectus.com/news/article/10037/introducing-siera-part-3/>

Swartz, Matt and Eric Seidman (2010). Introducing SIERA: Part 4.

<https://www.baseballprospectus.com/news/article/10042/introducing-siera-part-4/>

Using Retrosheet data from 2003 through 2009, Baseball Prospectus's Matt Swartz and Eric Seidman introduced SIERA (Skill-Interactive Earned Run Average), a very complicated pitching metric that they claim to be the most accurate predictor of them all. As with any metric based on regression analyses for specific seasons, that claim will only be true for those seasons. Nonetheless, the concept has been influential, and FanGraphs has its own version.

Tallavarjula, Surya (2026). A Monte Carlo simulation of baseball offense with speed-stratified baserunning and distributional validation. *Journal of Sports Analytics*, Vol. 12, Article 22150218251410737

Based on 2022 to 2024 Retrosheet data plus Statcast sprint speed data from 2015 through 2024, Tallavarjula (2026) proposed an offense simulation method. The simulation results for team runs per game were highly correlated, at about 0.9 by eyeball, with actual R/PA, and individual players RBI and runs scored correlated at perhaps 0.5. As part of an argument for including sprint speed as part of the analysis, the author showed substantial increases in the odds of extra base advancement of hits as baserunners became faster, particularly for scoring from first on a double (read off a diagram) from about 20 percent for the one-fifths slowest runners to almost 60 percent for the one-fifths fastest.

Tango, Tom aka Tangotiger (2006). Linear weights by run environment.

https://insidethebook.com/ee/index.php/site/article/linear_weights_by_run_environment

From Retrosheet data and based directly on work by Tom Ruane, here are run expectancies for different run environments:

RunsMin_	RunsMax_	n_	Rperl_	1B_	2B_	3B_	HR_	NIBB
0.000_	0.439_	15_	0.419_	0.436_	0.720_	0.986_	1.402_	0.279
0.440_	0.459_	21_	0.451_	0.448_	0.740_	1.017_	1.411_	0.292
0.460_	0.479_	8_	0.469_	0.454_	0.749_	1.005_	1.404_	0.295
0.480_	0.499_	11_	0.489_	0.460_	0.761_	1.052_	1.409_	0.305
0.500_	0.519_	17_	0.508_	0.459_	0.764_	1.060_	1.399_	0.303
0.520_	0.539_	4_	0.525_	0.471_	0.768_	1.078_	1.396_	0.314
0.540_	0.559_	7_	0.550_	0.474_	0.768_	1.043_	1.393_	0.324
0.560_	1.000_	7_	0.581_	0.489_	0.791_	1.049_	1.404_	0.337

“n” is the number of past seasons in each of the run environment groups; “Rperl” the average within each group. Note that doubles are consistently worth about 0.3 runs more than singles, triples 0.25 more than doubles, and walks 0.15 less than singles. Each of these increase by 0.06 or 0.07 from the lowest to the highest grouping. Home runs, on the other hand, do not increase, more evidence that homers are worth relatively more when runs are scarce than when runs are plentiful. This provides evidence that batting metrics should distinguish between homers and other offensive events types, as the metric BaseRuns does.

Tango, Tom M. (2008). With or without you. In Dave Studenmund (Producer), *The Hardball Times Baseball Annual* (pages 191-198). Skokie, IL: Acta Sports.

Tom Tango (2008) proposed a creative method that he called for With Or Without You (WOWY) for evaluating catcher ability to prevent passed balls and wild pitches, thwart attempted steals, and pickoff runners. For a given catcher:

- 1 – Choose a pitcher he caught.
- 2 – Count how many WPs, PBs, and innings occurred with that pitcher/catcher combination.
- 3 – Count how many WPs, PBs, and innings occurred with that pitcher and other catchers, and then use the ratio of WPs and PBs per inning to estimate the number that would have occurred if the other catchers had caught that pitcher the same number of innings and the catcher under examination.
- 4 – Comparing the results of steps 2 and 3 reveals how much better or worse the catcher under examination was than the other catchers for the given pitcher.
- 5 – Repeat these steps for all other pitchers the catcher under examination caught, and sum the results for an overall index.

Tom performed this task using Retrosheet data from 1972 through 1992. According to his chart displaying data for individuals during that era, the ones everyone thought were good (e.g., Jim Sundberg, Gary Carter) are indeed toward the top and those everyone thought were bad (e.g., Charlie Moore, Ron Hassey) toward the bottom. Tom noted that this method presumes that the other catchers to whom the catcher under examination is compared are league average; he tested the assumption and found it to be reasonably defensible. Incidentally, he noted that Tom Ruane had previously suggested this method. Michael Humphreys (2011) extended this idea to the evaluation of all fielders, by comparing a specific fielder's performance with those sharing his position on the same team in the same year.

Tango, Tom M. (2008) With or without...Derek Jeter. In Dave Studenmund (Producer), *The Hardball Times Baseball Annual* (pages 147-152). Skokie, IL: Acta Sports.

Tom Tango's With Or Without You also works for fielding in general. Tom described it in the context of Derek Jeter; Michael Humphreys (2011, pages 84-86) did a nice job of describing it in general, and I will use Michael's description. When

evaluating a particular fielder, the analyst uses relevant Retrosheet data to do the following:

- 1 – Choose a pitcher he fielded behind.
- 2 – When the fielder in question was playing, count how many batted balls in play that pitcher gave up, and how many of these batted balls were fielded by the fielder in question.
- 3 – When the fielder in question was not playing, count many of batted balls in play that pitcher gave up, and how many of these batter balls were fielded b others playing the same position as the fielder in question.
- 4 – Comparing the results of steps 2 and 3 reveals how many more or fewer balls the fielder in question would have successfully fielded than the “typical” other shortstop would have behind the same pitcher.
- 5 – Repeat these steps for all other pitchers the fielder in question played behind, and sum the results for an overall index.

Rather than the fielder’s team’s pitchers, one can do a WOWY analysis across opposition batters, different ballparks, and different baserunner situations to see if the results look any different.

Tango, Tom M. (2009). Catcher 911. In Dave Studenmund (Producer), *The Hardball Times Baseball Annual* (pages 191-198). Skokie, IL: Acta Sports.

Using Retrosheet data, Tom Tango (2009) examined every player who caught at least one game between 1956 and 2007 to compare the fielding performance of (1) those with at least half of game appearances as a catcher in a given season, (2) those who did not catch at least half of their game appearances in a given season but had in the past, (3) those who never caught at least half of their game appearances in a given season but did catcher at least ten times in their careers, and (4) those who caught fewer than ten games in their careers. Per 5000 batters (an approximate season of catching), those in the first three categories averaged 0.8, –8.5, –3.9, and a whopping –49.6 runs per season (measured as 0.5 runs gained for every caught stealing and pickoff and –0.25 runs for every stolen base, balk, wild pitch, and passed ball). In short, true emergency catchers were far worse fielders than even those who caught only occasionally. Keep in mind that the sample sizes for the last two categories were tiny.

In a second inquiry in the same book chapter, Tom compared the batting performance of catchers when playing on consecutive days versus having a day or two off between appearances, adjusted for relative playing time in each category. Contrary to expectation, there was absolutely no impact, with wOBAs of .323 for each. Finally, Tom compared the offensive performance of players before and after their 29th birthday. First basemen, other infielders, and outfielders produced about three runs per 650 plate appearances in the older category; catchers only 1.5 runs.

Tango, Tom aka Tangotiger (2011). How do great starting pitchers pitch the 4th time through the order?

https://insidethebook.com/ee/index.php/site/how_do_great_starting_pitchers_pitch_the_4th_time_through_the_order

Tom looked at the times through the other issue for “the top 15 starting pitchers” (what criteria of judgment not stated) between 1993 and 2010. Their combined wOBA was .293 first three times through and .313 fourth time, compared with .340 for all starters and .330 for relievers overall all times through.

Tango, Tom aka Tangotiger (2011). How do pitchers do each time through the order?
https://insidethebook.com/ee/index.php/site/comments/how_do_pitchers_do_each_time_through_the_order#comments

Continuing the analysis, Tom examined all 105 starters with at least 400 pitches fourth time through 1993 to 2010. I assume that he included every starter, not only those games going into fourth time, which is incorrect procedure. wOBA 1st time .320, 2nd time .331, 3rd time .341, 4th time .341.

In addition, there was no day versus night difference, contrary to an argument that pitching late at night results in biased data because of sleepy pitchers.

Tango, Tom aka Tangotiger (2011). DIPS: Nonpitchers as pitchers.
https://www.insidethebook.com/ee/index.php/site/dips_nonpitchers_as_pitchers

Tom Tango aka Tangotiger reported the following pitching performance comparison between pitchers and non-pitchers while batting between 1993 and 2011 (probably based on Retrosheet data):

	BABIP	HR/PA	K/PA	BB/PA
Pitchers	.295	0.027	0.168	0.096
Non-pitchers	.292	0.035	0.079	0.185

Note that pitchers' performance is far better than non-pitchers overall, but their batting average on balls in play hardly differs. This counts as evidence in favor of Vörös McCracken's claim that there is no difference among pitchers in BABIP. Tom also noted that pitchers as hitters had a .225 BABIP, far worse than non-pitchers as hitters.

Tango, Tom aka Tangotiger (2011). What happens to a pitcher's BABIP year to year?
https://www.insidethebook.com/ee/index.php/site/what_happens_to_a_pitchers_babip_year_to_year

The following table, offered by Tom Tango aka Tangotiger (2011x), consists of batting averages on balls in play for all pitchers with at least 400 balls in play in consecutive seasons 1993 through 2010 (likely based on Retrosheet data). The

rows are buckets centered around the posted figure for the first season, the columns the same for the next season, and column on the right are overall next-season BABIPs for each first-season bucket:

Note that as first-season BABIP increases, second-season BABIP does also, although at a much slower rate; Tom noted an 80 percent regression toward the mean slowing the increases. The cells in blue are median figures, which also increase. This is evidence that, contrary to a strong version of the McCracken thesis that pitcher BABIP is random, they are to at least a small extent not.

Tango, Tom aka Tangotiger (2011). What happens to a pitcher's HR rate year to year? https://www.insidethebook.com/ee/index.php/site/what_happens_to_a_pitchers_hr_rate_year_to_year

The same analysis idea as above, this time for home run rate for pitchers with at least 300 batters faced in the same home ball park in consecutive seasons. Home parks are first, followed by road parks:

They are quite similar, in each case regressing about 65 percent to the mean; as these are the same, specific home park must not bias the first table figures much. The implication is that pitchers have more control over homers than other hits, consistent with a weak formulation of the McCracken claim.

Tango, Tom aka Tangotiger (2011). Relief advantage on inherited/bequeathed runners. https://insidethebook.com/ee/index.php/site/tangos_lab_relief_advantage_on_inherited_bequeathed_runners

For all pitchers 1993-2010, Tom Tango (2011) used (I believe) Retrosheet data to make the following computations: for starters, an average BaseRuns of 4.98 per nine innings, compared with their average Runs Allowed of 5.05; for relievers, an average BaseRuns of 4.43, when their Runs Allowed was 4.32. So relievers had been assigned 0.07 too many and relievers 0.11 too little. The problem as Tom saw it was assignment of responsibility for runs between two pitchers splitting an inning. To remedy the problem, Tom suggested using the following formula for the first pitcher:

$$(3 - \text{outs responsible for}) \times (0.5 + [.08 \times \text{bases responsible for}])$$

His example: 1 out/runner on second:

$$(3 - 2) \times (0.5 + [0.8 \times 2])$$

which equals 0.42 runs charged to the first pitcher. If the baserunner scores, the second pitcher is charged for the rest of the run (0.58); if the baserunner is stranded, the second pitcher is credited with -0.42 runs.

Tango, Tom aka Tangotiger (2011). How do great starting pitchers pitch the 4th time through the order?

https://insidethebook.com/ee/index.php/site/how_do_great_starting_pitchers_pitch_the_4th_time_through_the_order

Tom looked at the times through the other issue for “the top 15 starting pitchers (what criteria of judgment not stated) between 1993 and 2010. Their combined wOBA was .293 first three times through and .313 fourth time, compared with .340 for all starters and .330 for relievers overall all times through.

Tango, Tom aka Tangotiger (2011). How do pitchers do each time through the order?

https://insidethebook.com/ee/index.php/site/comments/how_do_pitchers_do_each_time_through_the_order#comments

Continuing the analysis, Tom examined all 105 starters with at least 400 pitches fourth time through 1993 to 2010. I assume that he included every starter, not only those games going into fourth time, which is incorrect procedure.

1st time .320, 2nd time .331, 3rd time .341, 4th time .341.

In addition, there was no day versus night difference, contrary to an argument that pitching late at night results in biased data because of sleepy pitchers.

Looking more closely at the day versus night issue based on 76 starting pitchers with at least 100 plate appearances against during the day and 200 at night the fourth time through, in so doing insuring the correct analytic procedure, wOBAs were:

	First Time	Second Time	Third Time	Fourth Time
Day	.316	.326	.339	.334
Night	.319	.329	.341	.340

And the day versus night contrast for relievers, with the same minimum plate appearances against:

	7 th Inning	8 th Inning	9 th Inning
--	------------------------	------------------------	------------------------

Day	0.33	0.32	0.31
Night	0.33	0.32	0.31

Tango, Tom aka Tangotiger (2011). Is where you hit in the lineup conducive to better performance?

https://www.insidethebook.com/ee/index.php/site/is_where_you_hit_in_the_lineup_conducive_to_better_performance

Based on 1993 to 2010 data (the webpost ended “All hail Retrosheet!) for batters aged 25 through 29, Tom compared performance for a given player in different lineup positions, controlling for the relative number of plate appearances in each. The following table reveals the wOBA comparisons that were at least 1.7 standard deviation units different from one another:

Comparison			Comparison			Comparison		
Slots	1 st BA	2 nd BA	Slots	1 st BA	2 nd BA	Slots	1 st BA	2 nd BA
1 st vs. 4 th	.352	.368	3 rd vs. 6 th	.359	.354	6 th vs. 7 th	.340	.337
1 st vs. 6 th	.334	.344	4 th vs. 5 th	.365	.362	7 th vs. 8 th	.321	.326
2 nd vs. 3 rd	.362	.357	5 th vs. 9 th	.335	.323			

The ninth slot was usually lower than the others. Tom added this overall table:

lineup	woba1	woba2	diff	zScore
1	0.334	0.339	-0.004	-4.1
2	0.340	0.339	0.001	0.8
3	0.356	0.354	0.002	1.8
4	0.360	0.355	0.005	5.0
5	0.347	0.347	0.000	0.5
6	0.343	0.342	0.001	1.9
7	0.335	0.336	-0.001	-1.4
8	0.331	0.330	0.001	1.0
9	0.321	0.326	-0.005	-3.8

Tom realized that some of these disparities might be due to platoon effects, as batters are often placed in slots farther from the 3rd and 4th against same-sided than opposite-sided pitchers.

Tango, Tom aka Tangotiger (2011). Chance of scoring from base, 1993-2007 (excluding 1999).

https://tangotiger.net/retrosheet/reports/chance_scoring_by_base_out_retrosheet_years_since1993.html

From Retrosheet data, the odds of scoring for all base/out situations, 1993 to 2007.

OUTS RATE_BAT RATE_1B RATE_2B RATE_3B

0	0.167	0.397	0.614	0.842
1	0.120	0.266	0.417	0.663
2	0.070	0.132	0.232	0.292

Tango, Tom aka Tangotiger (2011). Is where you play in the field conducive to better hitting performance?

https://www.insidethebook.com/ee/index.php/site/is_where_you_play_on_the_field_conducive_to_better_hitting_performance

NOT IN BIBLIOGRAPHY, IN REFEREMCES

Immediately after the batting slot webpost, Tom did the same for different fielding positions for batters aged 25 through 29. First, and foremost, batters performed worse as designated hitters than at all other positions, with the possible exception of shortstop (not enough DH/SS combinations to check):

fld1	fld2	woba1	woba2	diff	z
10	2	0.321	0.341	-0.020	-2.2
10	3	0.349	0.364	-0.014	-3.7
10	4	0.317	0.334	-0.017	-1.8
10	5	0.327	0.347	-0.021	-3.3
10	7	0.345	0.355	-0.010	-2.5
10	8	0.340	0.356	-0.017	-2.4
10	9	0.349	0.361	-0.013	-3.2

Other combinations differing by at least one standard deviation:

Combination			Combination		
Positions	1 st Position	2 nd Position(s)	Positions	1 st Position	2 nd Position(s)
3B vs. 2B	.320	.331	1B vs. others*	.347	.352
3B vs. SS	.321	.326	C vs. 1B	.334	.351
2B vs. OF	.320	.329	C vs. corners*	.323	.342
IF vs. OF	.323	.328			

* others refers to all but catcher; corners refers to 1B, 3B, LF, and RF.

Tango, Tom aka Tangotiger (2012). How bad are pitchers when they check out of MLB?

https://www.insidethebook.com/ee/index.php/site/how_bad_are_pitchers_when_they_check_out_of_mlb

Reported by Tom, Bill James made a conjecture about the last starts before pitching careers end that was supported in a reader's research findings that were posted on Bill's website. For all pitchers in Retrosheet data since 1911, the winning average for the last 30 decisions of the 20 pitchers with the most starts was .357; all others with at least 200 starts ended at .423. The proposed reason is a form of selection bias; given past excellence, teams will give the best pitchers more opportunity to right themselves with the hope that they will return to some semblance of their past excellence, with no analogous feeling for other pitchers.

Tango, Tom aka Tangotiger (2020). Run values by pitch count.
<http://tangotiger.com/index.php/site/run-values-by-pitch-count>

Includes charts based on Retrosheet data showing run expectancies, probabilities for at least one run, and proportions of plate appearance charts for each pitch count, divided into four intervals (1950-1968; 1969-1992; 1993-2009; 2010-2015).

Tango, Tom aka Tangotiger (2024). Complete historical run expectancy chart.
<http://tangotiger.com/index.php/site/complete-historical-run-expectancy-chart>

Expansion of 2020, including charts for the run value of homers, an extension of the earliest interval to 1947-1968 and the most recent to 2010-2023, and the addition of charts for 1900-1920 and 1921-1946.

Tango, Tom aka Tangotiger (no date). How are runs *really* created: Third installment.
<https://tangotiger.net/rc3.html>

As part of an argument that Bill James's Runs Created over(under)predicts scoring in extremely high(low) run scoring environments, due to the fact that home runs are more(less) valuable in low(high) run environments. This is because the higher the run environment, the more base runners there are and the more runs that will be scored, whereas the batter who hits a homer will be worth one run regardless (see <http://www.tangotiger.net/runscreated.html> for the argument). Here, Tom used 1974 to 1990 Retrosheet data to both calculate and simulate games in which between zero and ten home runs were hit. Here are his findings for actual games:

Runs Scored, breakdown by HR hit

HRclass	n	R	BsR	LWTS	RC
0	33,068	3.08	3.06	3.79	3.03
1	23,117	4.62	4.62	4.44	4.66
2	9,218	6.12	6.12	5.00	6.41
3	2,838	7.65	7.65	5.62	8.37
4	687	9.03	9.00	6.07	10.29
5	146	10.55	10.49	6.73	12.45

6	40	12.33	12.32	7.52	15.35
7	9	16.22	14.32	8.34	18.27
8	2	14.00	15.87	8.58	22.52
10	1	18.00	18.30	9.51	27.03

RC predicted well for games with zero and one homer, but began overpredicting with two, and the amount of overprediction increased as homes per game continued rising. Pete Palmer's run expectancy method also worked poorly, whereas David Smyth's BaseRuns (see *Baseball Analyst* No. 29), which separates the impact of homers from other methods to get on base, worked almost perfectly until sample sizes became tiny. And here is why:

Score rate, breakdown by HR hit

HRclass	n	SR	bsr_SR	lwts_SR	rc_SR
0	33,068	0.267	0.265	0.328	0.262
1	23,117	0.304	0.304	0.289	0.307
2	9,218	0.340	0.340	0.247	0.364
3	2,838	0.373	0.372	0.210	0.430
4	687	0.404	0.401	0.166	0.504
5	146	0.430	0.426	0.134	0.578
6	40	0.454	0.454	0.109	0.671
7	9	0.610	0.484	0.089	0.746
8	2	0.414	0.543	0.040	1.001
10	1	0.533	0.554	(0.033)	1.135

These are the percentages of baserunners that scored in these games. Note that the patterns for each were almost identical to above. I might add that the excellence of BaseRuns as an offensive measure has been demonstrated by other analysts, particularly Brandon Heipp aka Patriot.

Tango, Tom aka Tangotiger (no date). (untitled). <https://tangotiger.net/lwtschart.html>

Here, Tom posted charts showing the relationship between different Batting Runs components and runs per game for the 1919 to 2000 period, clearly based on Retrosheet data. The relationship was very strong for singles, walks, and outs, apparent but less strong for double and triples, and noticeable but weak for homers, showing where the metric's problem demonstrated in Tom's work just summarized lies.

Tango, Tom aka Tangotiger (n date). Actual wins, Retrosheet years.

<https://www.tangotiger.net/wins.html>

See this website for a huge table showing the odds of a win given inning (top versus bottom), bases occupied, outs, and score differentials from 5 behind to 5 ahead for 1957 through 2006 excepting 1999.

Tango, Tom M., Mitchel G. Lichtman and Andrew E. Dolphin. (2006). *The Book: Playing the Percentages in Baseball*. TMA Press.

I begin with an editorial comment: This book belongs on the shelf of anybody who seriously studies quantitative baseball data. The entire book is based on sophisticated analysis using Retrosheet data (different seasons for different analyses, so I will skip the details on what seasons were employed). I will only list the themes, as describing all the findings would take too long:

In Chapter 1, entitled Toolshed, the authors explain the basics of run expectancy tables and their interpretation, and compute the “run value” of 20 possible events occurring during games, lists as demonstrations the run value of home runs at each base-out situation and the odds of scoring different numbers of runs at each base-out situation given an average of 3.2 or 5 runs per game. They also include the odds of a team winning the game given every base-out situation in every half inning (top of first through bottom of ninth) for every increment from being ahead by four runs to behind by four runs and the “win value” of the 20 events, which tells you how critical the situation is in which the event occurs on average. Finally, they define Tango’s measure of offensive performance, weighted on-base average, which in a linear weights-type formula but calibrated to be interpreted as one interprets OBA.

Chapter 2 takes on the issue of batting and pitching streaks, this time using 2000-2003 Retrosheet data. They note tiny but discernible tendencies for batters who have been hot or cold for five games to stay that way for a few more games, and the same for pitchers who have been hot over their last four appearances (but not for cold). However, as they did not correct for strength of opponent or ballpark, one should not read too much into this.

Chapter 3 is on batter/pitcher matchups and notes that specific player/player matchups probably are meaningless, replicates previous findings for lefty/righty and groundball/ flyball tendency matchups, finds no interaction effects between batters/pitchers good at controlling the strike zone or at making contact, and not much evidence that good pitching stops good hitting.

Chapter 4 addresses various situational issues. Contrary to all other research, the authors do find consistent clutch hitting tendencies for batters, but they are tiny and practically meaningless. They note no analogous clutch pitching effect for relievers. Pinchhitting indeed does lead to worse performance than being in the lineup, and it is not because pinchhitters tend to face fresh relievers in the late innings. There is no performance difference between hitting with runners on versus base empty.

Chapter 5 turns to the lineup. Here they weight run value by lineup opportunity (i.e., each lineup position has about .11 more plate appearances than the next and differing proportions across the base/out situations, i.e. leadoff batter comes up with fewer base runners than any other), and conclude consistently with received wisdom

that the leadoff batter should indeed be the best on-base average player and the last four slots (with an exception to be noted below) should have the team's worst hitters in descending order of run production. In contrast, the number 3 slot should have a weaker hitter than #s 2, 4, and 5. Again consistent with tradition, good basestealers/baserunners ought to be before batters who hit singles and don't strike out, and the "pitcher bats eighth/pre-leadoff hitter bats ninth idea does work if the pitcher is an average or better hitter for the position.

Chapter 6 considers the standard platoon differential. Most of what is here replicates the findings of several others concerning batters, but there is one useful addition: the platoon differential is not in general large enough to counteract the performance decrement for pinchhitters, such that one should only pinchhit for platoon advantage if the pinchhitter is considerably better than the batter replaced.

Chapter 7 features the starting pitcher, mostly concerning workload issues. Pitchers do perform a bit worse as the game continues on average. Across games, they perform best with five days rest, but the argument for a six-man rotation falters considering the (absence of) quality one's sixth starter would likely possess. Pitchers who blow through the first nine hitters tend to return to normal for the next nine, whereas pitchers who are hammered by the first nine batters still tend to struggle with the next nine and likely are having a bad day. Finally, pitchers usually perform better as relievers as starters, with the possible exception of starters pitchers with little or no experience as relievers at all.

Chapter 8 is the relief pitcher's turn. Conceptually, they compared the generic very good relief pitcher (analogous to one who would win 68% of their games) to the generic average one (50%). The 18% difference between the two breaks down to 2% an inning. In theory one would always do better with the very good reliever, but in practice you don't want to overwork him and so look for situations in which you don't lose much using the average reliever. Assuming long-term equal usage, the strategic implication is that a very good relief pitcher is worth bringing in a game rather than an average one if the odds of the good reliever winning is more than 2% more than the average reliever in a given base/out/inning situation and not if the odds are less than 2%. Using Retrosheet data from 1999-2002, they determined, for example, that the very good reliever need only be used in the ninth inning/three run lead situation (the easiest possible save given today's scoring procedures) if there is a baserunner with no outs or two baserunners with no or one out. Using historic data, they also argue that very good relievers can be trusted to not lose effectiveness up to about 25 pitches, which on average allows bringing them in during the eighth inning. Finally, they claim (and present evidence) that relievers in general do not lose effectiveness if used two or three days in a row. I am less confident in the last of these claims is defensible given that such usage is rare for the typical pitcher, and their data may not represent what would happen long-term if such usage became commonplace.

Chapter 9 is the most detailed analysis of the sacrifice bunt as a strategic tool thus far presented, taking up more than 50 pages of their book. They used Retrosheet data from 2000 through 2004 throughout, and, using Palmer's method, showed that the runner on first/zero outs sacrifice was overall even more harmful than in Pete's findings,

likely due to the overall increase in offense. In general, however, they applied a different and very useful method. For example, rather than comparing expected runs between runner on first/no out and runner on second/one out, they compared runs scored for the rest of the inning between runner on first/no outs when sacrifices were attempted and runner on first/no outs when sacrifices were not attempted. Note the term attempted: one can attempt to sacrifice, foul the pitch off, and then hit a home run on the next pitch; and these successful at bats ought to be included as well as the failures. Anyway, their wealth of findings are too numerous and complicated to describe in detail, and interested reader should consult *The Book*. In summary, the value of the sacrifice is affected by strength of the batter and of the batter on deck (the lower the on-deck's OBA, the better the bunt is), infield alignment (better if the infield is playing back), inning (better earlier in the game as infielders are less likely to be playing in for it), run environment (better when runs are more scarce), bunter skill, and baserunner speed. In addition, one should not use the same strategy all of the time as the other teams will respond accordingly with their defensive alignment, so randomly placed variation to decrease predictability will help.

Chapter 10 considers the intentional walk. Based on 2000-2004 Retrosheet data, there were no base-out situations in which the IBB decreased expected runs for the opposition overall. This was true even when the batter in question is much better than the batter on deck, including the #8 batter with the pitcher expected to come to the plate. There are a couple (second and third / one out, third / one out) in which it increases the defensive team's odds of winning, but by less than one percent. Interestingly, these are among the situations in which managers used it the most during those years, implying some intuitive understanding of the situation. Other exceptions are tied games in the bottom on the ninth when the IBB helps if it doesn't advance the lead runner, and when you have reached a 3-0 count against a very good hitter.

Chapter 11 is the stolen bases' turn. Overall success in basestealing during the 1999 through 2002 period of time, about 68%, was in general below the breakeven rate of 72%. The latter rate was dependent on game score (75.4% when three runs ahead and 66.9% when three runs behind) and inning (as the game progresses, the breakeven worsens when the team at bat is behind but improves when the team at bat is ahead). Interestingly, the data also provided evidence consistent with the platitude that baserunners disrupt the defense and improve the fortunes of hitters. Mean wOBA, .358 overall, was .372 with runners on first and less than two outs. Again not surprisingly, that broke down to .378 for lefthanded hitters and .368 for righties.

Finishing in Chapter 12 with the pitchout, the odds of success following a pitchout dropped to 47%. The implication that pitching out is a good strategy must be tempered by the fact that it adds a ball to the count, aiding the batter. That aid is highly dependent on the count. The TMA group (they were uncharacteristically silent on which years they used; I would guess 1999 to 2002) calculated a tiny increase in wOBA from .222 to .245 (corresponding to a scant .03 rise in runs scored) with a pitchout at an 0-2 count, but a large increase of .116 (equivalent to .15 runs) pitching out at 2-1. Combining the two, they estimated the breakeven point for pitchouts when the count is 0-2 and the opposing team believes the odds of an attempted steal are a scant 18

percent (in other words, it's a good strategy at 0-2), but this changes to 54% with a 2-1 count and one out (meaning that the opposing team has to feel that an attempt is more likely than not).

Teeter, Chris (2015). Swinging at 3-0 pitches: A high-risk decision.

<https://www.beyondthefoxscore.com/2015/3/5/8151763/baseball-swinging-count-pitches-balls-strikes-strikezone-sluggers>

This is a detailed study of batter swing tendencies on 3-0 counts using Retrosheet data. Between 2009 and 2014, batters swung only at 3-0 only 7.6 percent of the time, in fact only 11.7 percent of the time when the pitch was in the middle of the plate. Between 2009 and 2013, 3-0 swings were more likely to occur innings 1 to 6 (an average proportional occurrence per inning of 12 2/3%) than 7 to 9 (8%), and more often when the batter's team was ahead. Here are breakdowns for different baserunner situations.

Baserunners	Swing%	Baserunners	Swing%
0-0-0	4.9	1-2-0	10.5
1-0-0	7.9	1-0-3	13.0
0-2-0	4.8	0-2-3	3.4
0-0-3	6.0	1-2-3	6.3

Temple, David G. (2013). Some observations and questions on handedness.

<https://blogs.fangraphs.com/some-observations-and-questions-on-handedness/>

David Temple's work was based on Retrosheet data on plate appearances between 1961 and 2013. In 1961, only about 2 1/2 percent of plate appearances went to switchhitters, but that rose to close to 10 percent from the mid-1960's to mid-1970's, then to almost 20 percent from the late 1980's to the mid-1990's, but down to about 15 percent between 2000 and 2013. PAs to righties were at about 60 percent from 1961 to the mid-1970's and between 50 and 55 through 2013. As for lefties, they received about 35 percent in the early 1960's, 30 percent from the mid-1960's to the mid-2000's, and then back to about 35 percent through 2013.

Pitchers have consistently been about 70 percent righty, 30 percent lefty throughout the half-century examined here.

Until the mid-1970's, 60 percent of right-handed batters threw right; this was down to about 50 percent in the 1990's and between 50 and 55 percent afterward. As for lefty batters, over 20 percent threw right-handed in 1961, then down to around 15 percent from the early 1970's to the mid-2000's, and then back up to 20 percentage. Fewer than 5 percent of switch-hitters threw right-handed in 1961, with this proportion going as high as almost 20 percent mid-1980s to mid-1990s, then back to around 15 percent in 2010.

Thornton-Lugo, Meghan A., Matthew W. McCarter, Jonathan R. Clark, William Luse, Steven J. Hyde, Zahra Heydarifard, and Lulu S. R. Huang (2023). Makeup calls in organizations: An application of justice to the study of bad calls. *Journal of Applied Psychology*, Vol. 108 No. 3, pages 374-402.

These authors claimed to have evidence that home plate umpires having been making make-up calls after mistakes, based on 2008-2014 data from BaseballSavant plus contextual info from Retrosheet. Their results suggest that, after missing a ball, umpers were 23 percent more likely to call a ball in the next five pitches than otherwise to the same batter and 10 percent more likely to do the same on other batters from the same team. After missing a strike, umpers were 15 percent more likely to call a strike in the next five pitches to the same batter. Adding context, the latter finding favoring pitchers was actually less likely to occur as leverage increased, the opposite of the authors' relevant hypothesis and evidence that the ump was more interested in getting things right. There was no analogous impact on either directions for batters. Their major findings are in line with those reported by Moskowitz and Wertheim in their book *Scorecasting*.

Thress, Tom (2012). Beyond Player Win Average. *Baseball Research Journal*, Vol. 41 No. 2, pages 22-29.

This to all extents and purpose is an updating of Mills and Mill's Player Win Averages analysis, providing ratings for prominent players beginning with 1948 and using Retrosheet data.

Timmerman, Bob (2008). The world of catcher interference.

http://baseballanalysts.com/archives/2008/08/the_world_of_ca.php

Includes Retrosheet data Phil Birnbaum examined on catcher's interference between 1956 and 2007 for this webpost. Reading off Phil's graph between 4 and 10 CI per year from the late 1950's to the early 1960's, total went way up afterward but was extremely variable year-by-year, averaging maybe 20 but ranging between 10 and 30.

Timmerman, Thomas A. (2007). "It was a thought pitch": Personal, situational, and target influences on hit-by-pitch events across time. *Journal of Applied Psychology*, Vol. 92 No. 3, pages 876-884.

Are Black players more susceptible to being hit by pitches? Earlier evidence implied that this may have been true in the 1950s but not anymore. Timmerman (2007) examined whether pitchers from the southern U.S.A. were more likely to hit Black batters than White batters immediately after a home run, after that batter had previously hit a home runs, or one of their own teammates were hit. Using Retrosheet data from 1960 to 1992 and 2000 to 2004 and controlling for batter OPS, whether a DH was used in the game, differential in team scores (assuming the losing team's pitcher would be

more likely to hit a batter), and pitcher walks per plate appearance, Timmerman noted overall increases in HBP in all three circumstances. However, opposite to what he expected, White batters were most likely to be hit by southern pitchers after they had homered and after the pitcher's teammate had been hit, with Blacks second and Hispanics last. Interestingly, pitchers not born in the south were more likely to hit Blacks than Whites and Hispanics in those circumstances.

Tollison, Robert D., and Octavian Vasilescu (2011). The designated hitter rule and the distribution of pitching talent across leagues. *Journal of Sports Economics*, Vol. 12 No. 4, pages 448-463.

It stands to reason that good hitting pitchers are a less valuable commodity and poor hitting pitchers less of a problem in a league with a designated hitter than a league without. It follows that a bias toward trading good hitting pitchers from the A.L. to the N.L. and poor hitting pitchers from the N.L. to the A.L. should have occurred around the time of the DH's imposition. Tollison and Vasilescu used the Retrosheet transaction file for trades. Examining (non-Retrosheet) data from 1960 through 1985, and controlling for pitcher quality as measured by ERA, age, and usage patterns as measured by IP, there appeared to be such a bias in 1972 and 1973 but not before and after. A second type of analysis found the same for 1970 (perhaps imagining the coming of the rule change) and 1972.

Tourtellotte, Shane (2012a). The hangover effect. <https://tlt.fangraphs.com/the-hangover-effect/>

Tourtellotte, Shane (2012c). The double-header hangover effect. <https://tlt.fangraphs.com/the-double-header-hangover-effect/>

Tourtellotte, Shane (2012d). An incomplete history of the double-header. <https://tlt.fangraphs.com/an-incomplete-history-of-the-double-header/>

Tourtellotte, Shane (2014a). Beyond the ninth inning. <https://tlt.fangraphs.com/beyond-the-ninth-inning/>

Tourtellotte, Shane (2014b). Hair of the dog: The hangover effect revisited. <https://tlt.fangraphs.com/hair-of-the-dog-the-hangover-effect-revisited/>

These webposts included a wealth of data concerning extra inning games plus what Shane called the "hangover effect"; the tendency for teams to perform relatively poorly after a very long game. I only present some of the data here; there is more at the webposts. First, Shane provided the raw number of extra inning games lasting given numbers of innings; I turned these into the following proportions:

Inning	10th	11th	12th	13th	14th	15th	16th	17th	18th	19th	20th	21st	22nd
Games	.4703	.2528	.1350	.0709	.0358	.0171	.0088	.0046	.0026	.0012	.0005	.0002	.0002

Here are the winning averages of 46 teams playing 18 innings or longer in subsequent days between 1990 and 2011 (overall team inning average was .5013, with winners at .5027 and losers at .4999).

Group of days	1-3	4-7	8-10	11-15	16-20	21-30
Winners	.5424	.5250	.4746	.4632	.4368	.4489
Losers	.4516	.3289	.5088	.4900	.4659	.4643

Shane (2014b) continued this project by looking at specific offensive and defensive hangover effects. There were none for offense...

Offensive Changes for “Hangover” Teams			
<i>Batting Stats</i>	<i>“Hangover”</i>	<i>Full Season</i>	<i>Difference</i>
H/9	9.06	9.052	0.008
R/9	4.648	4.623	0.025
BA/OB/SL	.2644/.334/.414	.2625/.3308/.4125	.0019/.0032/.0015
OPS	0.748	0.7433	0.0047
BB/9	3.385	3.339	0.046
SO/9	6.41	6.475	-0.065
HR/9	0.998	1.009	-0.011
HBP/9	0.332	0.32	0.012

...but significant dropoffs for defense:

Defensive Changes for “Hangover” Teams			
<i>Batting Stats</i>	<i>“Hangover”</i>	<i>Full Season</i>	<i>Difference</i>
H/9	9.297	9.025	0.272
R/9	4.839	4.589	0.25
ERA	4.381	4.193	0.178
UERA	0.458	0.396	0.062
HR/9	1	0.985	0.015
BB/9	3.372	3.334	0.038
HBP/9	0.32	0.327	-0.007
SO/9	6.542	6.681	-0.139
FIP*	4.321	4.258	0.063

From 2012c: The following tables from include 74 double-headers from 2004 to 2011. They show the “hangover effect” doubleheaders seem to cause for subsequent games. The included teams had combined seasons record of relevant teams .5056.

The first table is for all games, not counting “rematches.”

Group of days	1-3	4-7	8-10	11-15	16-20	21-30
D-H players' record	109-125	226-253	141-163	233-275	259-259	433-455
Winning percentage	.4658	.4718	.4638	.4587	.5000	.4876
Cumulative record	109-125	335-378	476-541	709-816	968-1075	1401-1530
Cumulative win pct.	.4658	.4698	.4680	.4649	.4738	.4780

“Teams that swept their double-headers had an average season record of .5264; those who split played at .5123 over the full season; losers of their double-headers had season marks of .4723.” Record for subsequent days.

Group of days	1-3	4-7	8-10	11-15	16-20	21-30
Winners' record	55-40	57-70	36-46	65-70	71-60	102-101
Winning percentage	.5789	.4488	.4390	.4815	.5420	.5025
Cumulative record	55-40	112-110	148-156	213-226	284-286	386-387
Cumulative win pct.	.5789	.5045	.4868	.4852	.4982	.4994
Splititters' record	88-88	107-128	84-78	116-128	122-123	218-237
Winning percentage	.5000	.4553	.5185	.4754	.4980	.4791
Cumulative record	88-88	195-216	279-294	395-422	517-543	735-780
Cumulative win pct.	.5000	.4745	.4869	.4835	.4877	.4851
Losers' record	35-66	63-61	34-48	50-80	56-72	107-105
Winning percentage	.3465	.5081	.4146	.3846	.4375	.5047
Cumulative record	35-66	98-127	132-175	182-255	238-327	345-432
Cumulative win pct.	.3465	.4356	.4300	.4165	.4212	.4440

More on double-headers from 2012d; data from 1912-1915 including Federal League, 1937-1941, 1955-1959. More doubleheaders correlated with worse team record, r of $-.32$. This is at least partly due to worse teams having more scheduled double-headers in the 1955-1959 stretch, with a regression line running from 13 DHs for 1st place teams to 15 for the 8th place teams. Shane believed this accounted for about one-third of the correlation above. Due to poorer teams trying to attract fans of course. Also, the correlation was almost all due to unscheduled DHs (rainouts made up).

Double-headers accounted for only 2 to 4 percent of scheduled games through about 1941, when it jumped to 16 percent during the war years. Down to 10 percent through mid 1950s, and then generally went down until virtually gone by late 1980s.

Tourtellotte, Shane (2013). Is there a no-hitter letdown? <https://tbt.fangraphs.com/is-there-a-no-hitter-letdown/>

To see if there was a performance letdown in the start after a no-hitter, Steve looked at 135 no-hitters between 1950 and 2012 using Retrosheet data:

	IP	R/9	ER/9	H/9	BB/9	K/9
No-hitter season	26623.2	3.889	3.489	8.204	3.077	6.526
Game after no-hitter	925	3.853	3.444	7.881	2.724	6.645

No letdown – if anything, performance was a bit better next game. Further, starters went as far in the following game as any other start:

	Starts	IP	IP/G
No-hitter season	2032	13644.1	6.715
Game after no-hitter	67	449	6.701

Tourtellotte, Shane (2016). Steals of home: The millennium so far.

<https://tth.fangraphs.com/steals-of-home-the-millennium-so-far/>

Turkenkopf, Dan (2009). Stealing a run. <https://tth.fangraphs.com/stealing-a-run/>

Shane Tourtellotte (2016) used Retrosheet plus game recaps from 2000-2015 to examine the state of the attempted steal of home. In total, there were 190 successful steals in 720 attempts (26.4%). With one out, there were 79 successful out of 408 (19.4%), many of which were failed squeeze bunts. With two outs, there were 104 successful out of 291; the 35.7 percent rate approximated break-even for runner on third/two outs. In addition, there were 162 pickoff attempts, 68 of which resulted in errors and 66 runs scored on those errors. I assume that the other 94 were successful pickoffs. Finally, 400 balks resulting in runs scoring occurred. The average yearly run expectancy was +8.30 for successful steals and +8.88 for balks, but –21.95 for caught stealing and –2.56 for pickoffs, for a total of –7.32, meaning an average of about a tenth a run lost per attempt. The number of attempts varied from 8 (five of the seasons) to 18, 17, and 15 twice.

Dan Turkenkopf (2009) also used Retrosheet data plus game recaps for 2000 through early 2009. His results are suspect, as he found 15 successful attempts of stealing home against only 10 unsuccessful attempts. The useful part of his study was his notice that 13 were attempted against lefty pitchers with 9 successful versus 10 against righties with 4 successful. Something is wrong here (it sums to only 23 attempts), but it shows that as lefties are facing away from the runner they are more susceptible to attempts.

Tourtellotte, Shane (2016). History delayed. <https://tth.fangraphs.com/history-delayed/>

Shane studied this question based on 561 games in which no-hitters were taken into the seventh inning between 1994 and 2015 using Retrosheet data. Overall, no hitters after

6th inning survived the 7th 39.4 percent of the time, after the 7th survived the 8th 42.5 percent of the time, after the 8th survived the 9th 60.6 percent of the time. The results showed that all told in the 6th through 8th, pitchers who lost no-hitter attempts in the 7th through 9th saw an average of 4.36 batters with 15.78 pitches thrown. Those who maintained it through an inning faced an average of 4.35 batters and 16.60 pitches thrown. In conclusion, there is no evidence in favor of the myth, with the pitches thrown actually slightly opposite to it. Shane attempted more statistical analyses which did not change this conclusion.

Tourtellotte, Shane (2017). How umpires' ejection rates change with age and experience. <https://tbt.fangraphs.com/how-umpires-ejection-rates-change-with-age-and-experience/>

Shane Tourtellotte (2017) examined umpire ejections for 2012 to 2016 using data from Retrosheet and a now-defunct website called the Umpire Ejection Fantasy League. In total, there were 743 by the home plate, 114 by the first base, 47 by the second base, and 62 by the third base ump, total 966 with very little year-to-year differences (range 216 to 179). Managers were tossed in 431 cases, non-managers in 538. Shane's real interest was in seeing if older/more experienced umps were more tolerant. Overall, age was unrelated to ejection rate: correlations of – 0.03 by games and –0.06 by age. Dividing ages into five-year buckets there was a downward trend from 26-30 years old (24 percent higher than average) through 56-60 (11½ percent lower than average) before a final jump back to 23 percent higher for 61-65; but no clear corresponding tendency for career games. The evidence, if any, was weak.

Treder, Steve (2006). Examining the relief of relieving.
<https://tbt.fangraphs.com/examining-the-relief-of-relieving/>
 IN BIBLIOGRAPH, IN REFERENCES, NEEDS RETROSHEET ENTRY

From Retrosheet data:

Overall MLB Starter vs. Relief ERA

Years	Starter	Relief	Delta
2001-05	4.50	4.10	– 9%
1996-00	4.68	4.35	– 7%
1991-95	4.24	3.96	– 7%
1986-90	4.02	3.66	– 9%
1981-85	3.93	3.50	– 11%
1976-80	3.88	3.60	– 7%
1971-75	3.60	3.43	– 5%
1966-70	3.48	3.40	– 2%
1961-65	3.76	3.56	– 5%
1957-60	3.88	3.77	– 3%

But this is a biased comparison, because the very best starters rarely relieve and the pitchers around replacement level are usually relievers. So to counteract that problem, the following is based on 291 pitchers who both started and relieved at least 75 times:

Median ERA+: 98

96 LHP, 195 RHP (33%, 67%)

	G	SV/G	IP/G	H/9	HR/9	UBB/9	SO/9	ERA
As Starter (50%)	56656	n/a	6.30	8.95	0.87	2.85	5.30	3.93
As Reliever (50%)	55763	0.12	1.60	8.51	0.75	3.04	6.09	3.61
Delta			-75%	-5%	-14%	+7%	+15%	-8%

We already know that relievers tend to strike out and walk more batters because of the different mindset of relieving (full out) versus starting (some pacing oneself). Note the home runs.

More detail:

LEFT-HANDED PITCHERS (96)

Median ERA+: 98

	G	SV/G	IP/G	H/9	HR/9	UBB/9	
SO/9							ERA
As Starter (52%)	19853	n/a	6.28	8.96	0.87	2.84	
5.44							3.88
As Reliever (48%)	18522	0.09	1.45	8.53	0.74	3.19	
6.22							3.64
Delta			- 77%	- 5%	- 14%	+ 12%	+
14%							- 6%
Difference from Total			- 2%	=	=	+ 5%	-
1%							+ 2%

RIGHT-HANDED PITCHERS (195)

Median ERA+: 98

	G	SV/G	IP/G	H/9	HR/9	UBB/9	
SO/9							ERA
As Starter (50%)	36803	n/a	6.32	8.94	0.87	2.86	
5.22							3.96
As Reliever (50%)	37241	0.14	1.67	8.51	0.75	2.97	
6.03							3.60
Delta			- 74%	- 5%	- 14%	+ 4%	+
16%							- 9%
Difference from Total			+ 1%	=	=	- 3%	+
1%							- 1%

Similar profile except for walks.

The following are for 30 pitchers whose number of starting and relieving appearances were within 10 percent of one another:

CLOSEST TO EQUAL NUMBER OF STARTING/RELIEF APPEARANCES (30 Pitchers)

Median ERA+: 96

13 LHP, 17 RHP (43%, 57%)

		G	SV/G	IP/G	H/9	HR/9	UBB/9	
SO/9	ERA							
As Starter (50%)		5157	n/a	6.13	9.10	0.99	3.06	
5.28	4.17							
As Reliever (50%)		5152	0.10	1.67	8.55	0.79	3.39	
5.94	3.74							
Delta				- 73%	- 6%	- 20%	+ 11%	+
13%	- 10%							
Difference from Total				+ 2%	- 1%	- 6%	+ 4%	-
2%	- 2%							

Nothing drastically different.

Here is the best and worse 10 percent:

BEST CAREER ERA+ (29 Pitchers)

Median ERA+: 115

11 LHP, 18 RHP (38%, 62%)

		G	SV/G	IP/G	H/9	HR/9	UBB/9	
SO/9	ERA							
As Starter (52%)		8034	n/a	6.75	8.34	0.78	2.64	
6.29	3.41							
As Reliever (48%)		7358	0.28	1.47	8.08	0.67	2.65	
6.95	3.11							
Delta				- 78%	- 3%	- 14%	+ 0%	+
10%	- 9%							
Difference from Total				- 3%	+ 2%	=	- 7%	-
5%	- 1%							

WORST CAREER ERA+ (29 Pitchers)

Median ERA+: 84

9 LHP, 20 RHP (31%, 69%)

		G	SV/G	IP/G	H/9	HR/9	UBB/9	
SO/9	ERA							
As Starter (49%)		3415	n/a	5.70	9.50	1.00	3.40	
4.98	4.69							
As Reliever (51%)		3583	0.05	1.66	9.13	0.84	3.61	
5.76	4.29							
Delta				- 71%	- 4%	- 16%	+ 6%	+
16%	- 9%							
Difference from Total				+ 4%	+ 1%	- 2%	- 1%	+
1%	- 1%							

10 percent earliest and latest, most and fewest walks, most and fewest strikeouts no differences; see the original for details. In contrast, most and fewest homers:

BEST HOME RUN PREVENTERS (29 Pitchers)

Median ERA+: 101

10 LHP, 19 RHP (35%, 65%)

SO/9	ERA	G	SV/G	IP/G	H/9	HR/9	UBB/9	
As Starter (46%)		5428	n/a	6.33	8.77	0.60	3.07	
5.31	3.65							
As Reliever (54%)		6370	0.15	1.70	8.61	0.54	3.01	
5.49	3.37							
Delta				- 73%	- 2%	- 10%	- 2%	+
3%	- 8%							
Difference from Total				+ 2%	+ 3%	+ 4%	- 9%	-
12%	=							

WORST HOME RUN PREVENTERS (29 Pitchers)

Median ERA+: 93

8 LHP, 21 RHP (28%, 72%)

SO/9	ERA	G	SV/G	IP/G	H/9	HR/9	UBB/9	
As Starter (49%)		4162	n/a	5.96	9.52	1.30	3.03	
5.62	4.72							
As Reliever (51%)		4372	0.06	1.48	9.03	1.05	3.13	
6.64	4.23							
Delta				- 75%	- 5%	- 19%	+ 3%	+
18%	- 10%							
Difference from Total				=	=	- 5%	- 4%	+
3%	- 2%							

Next one is interesting:

MOST SAVES PER RELIEF APPEARANCE (29 Pitchers)

Median ERA+: 104

4 LHP, 25 RHP (14%, 86%)

SO/9	ERA	G	SV/G	IP/G	H/9	HR/9	UBB/9	
As Starter (36%)		5793	n/a	6.51	8.77	0.82	2.86	
5.81	3.82							
As Reliever (64%)		10331	0.30	1.46	8.25	0.67	2.64	
6.56	3.22							
Delta				- 78%	- 6%	- 18%	- 8%	+
13%	- 16%							
Difference from Total				- 3%	- 1%	- 4%	- 15%	-
2%	- 8%							

FEWEST SAVES PER RELIEF APPEARANCE (29 Pitchers)

Median ERA+: 94

12 LHP, 17 RHP (41%, 59%)

SO/9	ERA	G	SV/G	IP/G	H/9	HR/9	UBB/9
------	-----	---	------	------	-----	------	-------

As Starter (48%)	4027	n/a	5.93	9.46	1.13	2.99		
5.43 4.56								
As Reliever (52%)	4435	0.01	1.48	9.05	0.95	3.18		
6.00 4.23								
Delta			- 75%	- 4%	- 16%	+ 6%	+	
10% - 7%								
Difference from Total			=	+ 1%	- 2%	- 1%	-	
5% + 1%								

So is this:

MOST IMPROVED ERA WHEN RELIEVING (29 Pitchers)

Median ERA+: 98

5 LHP, 24 RHP (17%, 83%)

SO/9	ERA	G	SV/G	IP/G	H/9	HR/9	UBB/9	
As Starter (45%)		5257	n/a	6.10	9.15	0.90	3.11	
5.56 4.25								
As Reliever (55%)		6297	0.17	1.49	7.91	0.62	2.91	
6.72 2.99								
Delta				- 76%	- 14%	- 31%	- 6%	+ 21
% - 30%								
Difference from Total				- 1%	- 9%	- 17%	- 13%	+
6% - 22%								

MOST DECLINED ERA WHEN RELIEVING (29 Pitchers)

Median ERA+: 98

14 LHP, 15 RHP (48%, 52%)

SO/9	ERA	G	SV/G	IP/G	H/9	HR/9	UBB/9	
As Starter (64%)		6647	n/a	6.32	8.62	0.82	2.79	
5.55 3.67								
As Reliever (36%)		3674	0.07	1.51	9.08	0.87	3.67	
6.34 4.47								
Delta				- 76%	+ 5%	+ 6%	+ 32%	+
14% + 22%								
Difference from Total		- 1%	+ 10%	+ 20%	+ 39%	- 1%	+ 30%	

BEST HOME RUN PREVENTERS (29 Pitchers)

Median ERA+: 101

10 LHP, 19 RHP (35%, 65%)

SO/9	ERA	G	SV/G	IP/G	H/9	HR/9	UBB/9	
As Starter (46%)		5428	n/a	6.33	8.77	0.60	3.07	
5.31 3.65								
As Reliever (54%)		6370	0.15	1.70	8.61	0.54	3.01	
5.49 3.37								
Delta				- 73%	- 2%	- 10%	- 2%	+
3% - 8%								

Difference from Total		+ 2%	+ 3%	+ 4%	- 9%	-
12%	=					

WORST HOME RUN PREVENTERS (29 Pitchers)

Median ERA+: 93

8 LHP, 21 RHP (28%, 72%)

		G	SV/G	IP/G	H/9	HR/9	UBB/9	
SO/9	ERA							
As Starter (49%)		4162	n/a	5.96	9.52	1.30	3.03	
5.62	4.72							
As Reliever (51%)		4372	0.06	1.48	9.03	1.05	3.13	
6.64	4.23							
Delta				- 75%	- 5%	- 19%	+ 3%	+
18%	- 10%							
Difference from Total		=	=	- 5%	- 4%	+ 3%	- 2%	

Turkenkopf, Dan (2009). Adjusting steals for win value.

<https://tft.fangraphs.com/adjusting-steals-for-win-value/>

The point here is to adjust stolen base totals by their timeliness given the game situation as measured by win value. Dan did not calculate for every attempt, but rather calculated average run value for an attempt for each year from 1954 through 2008 using Retrosheet data. The result of each attempt was calculated as

Successful Steals minus Pickoffs minus (2 X Caught Stealing)

The highest single-season wins per attempt figure was 0.03, attained by Pokey Reese in 2001; with 30 attempts, Reese's effort translated to 0.9 wins. As he ran willy-nilly and was caught and picked off quite a bit, Rickey Henderson's 1982 was only worth .004 runs per attempt and a run value of 0.7. The most impressive career was Tim Lincecum, with 0.011 wins per attempt (just .001 less than the highest figures) and 911 attempts giving him a little over 10 wins. These are actually approximations; a more accurate method would replace annual averages with the actual run value of each attempt given base/out/score/inning situation at each attempt.

Turocy, Ted (2004). A theory of theft: The strategic role of the stolen base in baseball. Unpublished manuscript.

Turocy, Ted (2014). An inspection game model of the stolen base in baseball: A theory of theft. Available at www.gambit-project.org/turocy/papers/theft-20140822.pdf

In two unpublished papers, Ted Turocy presented mathematical models on the strategic value of the stolen base attempt.

Turocy, Theodore L. (2005). Offensive performance, omitted variables, and the value of speed in baseball. *Economics Letters*, Vol. 89, pages 283-286.

Ted Turocy (2005), using Retrosheet data from 1974 to 1992, came up with an overall breakeven point for stolen base attempts of .717.

Turocy, Theodore L. (2008). In search of the “last-ups” advantage in baseball: A game-theoretic approach. *Journal of Quantitative Analysis in Sports*, Vol. 4, Issue 2, Article 5.

Is there a last up advantage? Ted Turocy (2008) used Retrosheet data from 1973 through 1992 as data for a simulation assuming two teams of equal quality, and concluded that there is a infinitesimal last-ups advantage of .001 in winning percentage, equivalent to an extra win every six years.

Uelkes, Peter (2012). Game scores. *Baseball Research Journal*, Vol. 41 No. 2, pages 30-36.

This is a detailed examination of starting pitchers using Bill James’s Game Score concept, based on more than 117,000 Retrosheet games. The most important part is the discovery that home starters have had a 14.7% advantage over road starters in strikeout/walk ratio, consistent with other research revealing pitch f/x data revealing umpire bias in ball/strike counts in favor of home teams.

Walsh, John (2005). Do batters try to hit sacrifice flies? <https://tbt.fangraphs.com/do-batters-try-to-hit-sacrifice-flies/>

Walsh, John (2005). Can batters successfully modify their batting approach? <https://tbt.fangraphs.com/can-batters-successfully-modify-their-batting-approach/>

A report on sacrifice flies, using 2003 and 2004 Retrosheet data for runner on third and fewer than two outs (sample size of more than 19,800).

The first row displays all relevant PAs (does not include IBB, HBP, or SH). The second row shows sac fly opportunities as defined above. John purposely scaled the first row so that the total number of PAs equals that for the second row. The third row is the raw difference between the two, and the fourth the percentage difference.

Fly	LD	GB	Popup	K	BB	HR	H	TB	AB
4193	2802	6555	1200	3426	1624	603	4777	7752	18183
4422	2667	6951	1215	3021	1520	507	5080	7834	18286
229	-135	396	15	-405	-104	-96	303	82	103
5.5	-4.8	6.0	1.3	-11.8	-6.4	-15.9	6.3	1.1	0.6

Note that fly balls were indeed higher in sac fly situations, but so are grounders, but for some reason not liners. Strikeouts and walks were both lower, indicating that batters seem to be trying to make contact. Homers were lower in sac fly situations, but other hit types were higher. Overall, the data suggest that batters do try to hit fly balls in sac fly

situations, but the higher grounder and lower HR rates imply that they are not consistently successful.

	AVG	OBA	SLG	ISO	RC	OUTS	RC27
All:	0.263	0.323	0.426	0.163	2505	13406	5.0
SF:	0.278	0.333	0.428	0.150	2611	13206	5.3

I added the isolated power figures to make the following point: These data suggest that, perhaps due to the increase in grounders, batting average was noticeably higher. But since (as noted above) this tendency probably limited homers, isolated power was lower and so slugging average about the same. The Runs Created and RC/27 figures show that batters were more productive in sac fly situations than overall, suggesting that the increase in ground-ball singles more than made up for the homer decrease.

Success rate on flies was 60.3%.

Below is evidence that the increase in hits was at least partly due to different defensive alignments: First is all relevant PAs, second one only SF situations:

batted-ball type	outs	1B	2B	3B	HR
F	0.734	0.055	0.078	0.012	0.120
G	0.764	0.215	0.020	0.001	0.000
L	0.264	0.515	0.178	0.015	0.028
P	0.982	0.014	0.003	0.000	0.000

batted-ball type	outs	1B	2B	3B	HR
F	0.755	0.068	0.069	0.012	0.097
G	0.749	0.227	0.023	0.001	0.000
L	0.171	0.580	0.202	0.018	0.029
P	0.977	0.021	0.002	0.000	0.000

As implied by the data above, sac fly situations did increase the likelihood that grounders became singles and doubles, suggesting that infielders were playing in. Importantly, they also increased the single rate for fly balls, but also increased the fly out rate and decreased the double rate. Liners were also more likely to become hits of all types. These data imply that outfielders were playing deeper in sac fly situations than overall.

Walsh, John (2006). Endangered species: The three-base hit.

<https://tbt.fangraphs.com/endangered-species-the-three-base-hit/>

Triples per team adjusted for 162 games varied between 70 and 90 from 1900 through about 1930, and then dipped to about 60 in 1940, 40 in 1960, and 30 in the 1990's through 2005. Reasons – smaller ballparks, faster outfielders.

Some trivia – left-handed hitters a bit more triple-productive (0.58% of PAs, with righties at 0.41%). Center fielders fielded the most (1239), with right fielders at 1166 and left fielders 303. This does not necessarily mean that more went to center than right, as perhaps center fielders fielded them in right field if the right fielder fell down. A very simple regression model in which triples were regressed on at bats and stolen bases accounted for 86 percent of the variance, showing the impact of speed on both SBs and 3B hits.

Walsh, John (2006). Three-and-oh. <https://tbt.fangraphs.com/three-and-oh/>

2000-2005 Retrosheet data. Batters swung on 3-0 8.3 percent of the time, and here is the batting line when swinging:

	3-0 All Counts	
BABIP	.380	.327
HR/BIP	.094	.040
HR/H	.25	.12

Pretty clearly better than overall.

Walsh, John (2006). Base stealer intangibles (part 1). <https://tbt.fangraphs.com/base-stealer-intangibles-part-1/>

Just having a runner on first results in a lot fewer grounders fielded by first base and with fewer than two outs a few fewer by second and shortstop – see Infield Positioning section of Fielding Strategy chapter for details.

Here is comparison between outcomes for batters batting behind 10 most prolific base stealers 2003-2005 for when the base stealer was and was not on first, corrected to make PAs the same:

	Fly	LD	GB	Pop	K	BB	HR	H	TB
AB									
All Opps:	753	523	1236	206	593	381	105	904	1439
3286									
S1B Opps:	798	596	1286	204	502	303	94	1007	1504
3355									
Diff:	45	73	50	-2	-91	-78	-11	103	65
69									
Diff Pct:	6.0	13.9	4.0	-0.8	-15.4	-20.5	-10.6	11.3	4.5
2.1									

“There are several interesting things to note about these numbers:

- **Hits** – As we expected from our discussion of defense above, there are more hits in the S1B situations.
- **Balls in Play** – There are significantly more balls put into play in S1B situation, as seen by the increase in the batted ball type categories and the decrease in strikeouts and walks.
- **Power** – While hits increased by 11%, total bases did not keep pace, increasing by only 4.5%. The main reason is a 10% decrease in home runs.”

	AB	H	2B	3B	HR	BB	K
All:	3286	904	180	20	105	381	593
S1B:	3355	1007	187	14	94	303	502

	AVG	OBP	SLG	RC	OUTS	RC27
All:	0.275	0.351	0.438	504	2381	5.72
S1B:	0.300	0.358	0.448	539	2348	6.19

Impact of runner on first on percentage of grounders fielded.

Percentage of ground balls fielded

Situation	GB	1B	2B	3B	SS	OF
All GB	143302	0.12	0.22	0.17	0.24	0.17
Runner on 1B	38218	0.09	0.21	0.17	0.23	0.21

Note that runner on first means first baseman fields considerably fewer due to covering first, and second and short a tiny bit less due to shading toward second for double play. The next one

Percentage of ground balls fielded, two outs

Situation	GB	outs	1B	2B	3B	SS	OF
All GB	43441	2	0.12	0.22	0.18	0.24	0.16
Runner on 1B	13911	2	0.09	0.23	0.18	0.24	0.18

With two outs no double play positioning, so percentage fielded for 2B and SS not fewer with runner on first. 1B still less, still covering first.

Percentage of ground balls fielded, left/right splits

Situation	GB	Bats	1B	2B	3B	SS	OF
All GB	41034	L	0.22	0.31	0.06	0.15	0.17
Runner on 1B	10577	L	0.17	0.31	0.06	0.15	0.22
All GB	54460	R	0.04	0.14	0.26	0.30	0.17
Runner on 1B	14907	R	0.03	0.14	0.26	0.29	0.20

Note handedness impact for 1B.

Walsh, John (2006). Base stealer intangibles (part 2). <https://tbt.fangraphs.com/base-stealer-intangibles-part-2/>

The following are the overall lines for the pitchers who faced a top-10 base stealer; note that overall they are slightly better than those who did not.

	AVG	OBP	SLG	OPS		RC27
All pitchers:	.264	.335	.423	.758		5.16
S1B pitchers:	.264	.330	.420	.749		5.04

As they are 0.12 better overall yet 0.05 worse above uncorrected for, they are actually affected 0.17 RC27

Second way of doing all this: The following simplifies the whole thing by comparing what happens with top-10 stealer on first vs any runner on first, second empty. First is stealer on first, repeating a chart from part 1.

	AVG	OBP	SLG	RC	OUTS	RC27
All:	0.275	0.351	0.438	504	2381	5.72
S1B:	0.300	0.358	0.448	539	2348	6.19

Here is any runner (should be any runner minus top-10 stealer).

	AVG	OBP	SLG	RC	OUTS	RC27
All:	0.265	0.337	0.429	14908	75777	5.31
R1B:	0.282	0.342	0.443	15824	74888	5.71

Difference is 0.07 more in RC27 for top 10 basestealers, as compared with 0.17 in above analysis. Don't know which is more accurate.

In addition, about 20% fewer walks with top-10 base stealers (you can see this at beginning table I part 1), about half due to inherently better pitching and about half due to strategic differences (he says pitchers not wanting to put more baserunners on and batters wanting to hit the ball). Slightly more HBP (42 for top 10 basestealers vs. 35 expected otherwise when adjusted for sample size difference), but a lot more balks (5.5 per 1000 batters vs. 2.7 per 1000 batters for any runner on first second open).

Anyway he concludes that given the 0.17, which really isn't much, top-10 basestealers add maybe 2 runs per season at most due to disruption.

Walsh, John (2007). Going the other way. <https://tbt.fangraphs.com/going-the-other-way/>

This webpost reports the proportion (from my reading off of charts) of batted balls hit to the opposite field, from 1957 to 2006 Retrosheet data when hit location data was

available. Batted balls were considered to go to the opposite field if hit by a lefty(righty) batter and fielded by left(right) fielder.

The Proportion of homers for lefty and righty batters about the same, 10 percent of batted balls until about 1990, over 15 percent in the 1990s, down a bit in the 2000s.

Doubles – more for lefty batters, generally between 30 and 40 percent until a drop to about 30 in the 2000s. For righties, between 20 and 30 percent throughout, with perhaps a slight decline in the 2000s. The difference was probably due to right fielders usually having stronger arms, so righty batters were less likely to try for two.

Triples – righty batters between 40 and 50 percent. Lefties 10 to 20 percent, perhaps dipping toward 10 percent over time. The difference was likely due to the longer throw from right field to third base.

Outfield singles – Lefty and righty batters about the same, with a general increase from about 20 to about 25 percent during the interim.

Fly and line outs – also rose, from between 30 and 35 percent to the high 30s.

Walsh, John (2007). The advantage of batting left-handed.

<https://tbt.fangraphs.com/the-advantage-of-batting-left-handed/>

John Walsh (2007) computed the following proportions of grounders fielded by infielders that batters beat out, categorized by batter handedness, from 2003-2006 Retrosheet data. Note that lefty batters did not get a higher proportion of infield hits despite the fact that they are closer to first after their swing, likely because they tend to hit the ball to the right side, and there are fewer infield hits there than the left side.

bats	GB Fielded	IF Hits	IF Hit Fraction
Left	63188	4679	0.074
Right	111969	8790	0.079

Next, John showed plate appearances by handedness and position, along with batting averages, from 2000 to 2006 Retrosheet data:

Pos	Left	Right	Both	AVG
C	0.114	0.736	0.150	0.259
1B	0.563	0.363	0.073	0.278
2B	0.154	0.588	0.258	0.273
3B	0.189	0.681	0.130	0.268
SS	0.054	0.660	0.287	0.269
LF	0.465	0.437	0.098	0.278
CF	0.433	0.401	0.165	0.271
RF	0.424	0.497	0.079	0.276

Taking this into consideration, division into positions with only righty throwers and positions in which either hand will do.

Pos	AB_L	AB_R	AVG_L	AVG_R	L_minus_R
1B-OF	243784	223599	0.276	0.275	0.001
C-3B-SS-2B	65579	343551	0.269	0.266	0.003

The latter group hits about equally better left or right. But compare it to the following, which shows the overall platoon effect.

bats	pitc	AVG	OBP	SLG	OPS
L	R	0.275	0.356	0.452	0.808
L	L	0.253	0.328	0.396	0.724
R	L	0.271	0.346	0.443	0.788
R	R	0.260	0.323	0.414	0.737

As it is about twice as big as the differential for position, It looks like the positional difference may account for about half of the overall lefty differential.

Walsh, John (2008). Taking control. <https://tbt.fangraphs.com/taking-control/>

John Walsh discovered the following from Retrosheet data, years not specified. Walks with 3-0 count occurred about 35 percent of the time whether or not there was a runner on first. The average walk rate for pitchers more likely to go 3-0 was not surprisingly higher (8.4%) than those less likely to go 3-0 (7.9%).

Walsh, John (2008). The origin of the platoon advantage. In Dave Studenmund (Producer), *The Hardball Times Baseball Annual* (pages 165-171). Skokie, IL: Acta Sports.

John Walsh (2008) used the Neyer/James Guide to Pitchers as a source for pitcher repertoire and Retrosheet performance data from 1957-2006 to compute platoon differentials for pitchers, and compared those for one group of pitchers who had a relatively large platoon differential (he did not say how much) and a second group with a reverse differential; here is what he found concerning pitch usage (table cut and pasted from Jared Cross, 2015). John gave 5 points if the pitch was listed as the most used in Neyer/James, 3 points if the second most used, and 1 point if the third:

John Walsh's Pitch Usage Points				
		Pitch Usage Points		
Extremum	# of Pitchers	Slider	Curveball	Changeup

High-Split	22	53	21	10
Low-Split	29	22	62	42

Note that sliders were associated with high splits and curves and changeups with reverse splits. It is unfortunate that he had not included fastballs in the analysis, as subsequent research has revealed that fastballs along with sliders are the pitches associated with significant pitcher platoon differentials. On his list of pitchers included in the samples, fastball was listed first for all but one of the 22 of the high-split set whereas for the reverse aka low split group 12 of the 29 did not.

Walsh, John (2008). Hit 'em where they ain't—if you can. <https://tbt.fangraphs.com/hit-em-where-they-aintmdashif-you-can/>

2000-2007 Retrosheet data, runner on second only. With no out, 41 percent of grounders to the right side of the infield; with one of two outs, 36 percent. Very indirect evidence of attempt with no out to move the runner to third. The difference between zero and one/two out was closer to 7 percent in the early 1950's and 6 percent in the 1970's and 1980's, so the strategy has decreased use over time.

Walsh, John (2009). Beyond OPS: Filling in the gaps. <https://tbt.fangraphs.com/beyond-ops-filling-in-the-gaps/>

John Walsh (2009) proposed the metric. Double Plays Avoided (DPA), based only on times grounded into double plays; no DPs on liners or flies.

1 – Find the average rate of grounded into double plays per opportunity (he didn't define the latter, but it should be all base-out situations including a runner on first for either no or one out). This figure averaged about 11.5 percent “in recent years” (and 6 percent in 1911, perhaps due to less skilled infielders or to bunting/stealing/hit-and-run strategies to stay out of one).

2 – Find the number of GIDPs and DP opportunities for given batter.

3 – Multiply opportunities by .114, giving you expected GIDPs for that batter.

4 – Subtract from Step 3 the actual number of GIDPs. The results will be negative(positive) if there were more(fewer) than expected. The general range (based on 2006 to 2008 Retrosheet data) was about +7 (very good) to –7 (very bad) a year. The best ones were for players who are either fast or from high Three True Outcomes types who hit a lot of flies relative to grounders and walked and struck out a lot.

Now, a DP was worth –.85 runs, a lot, as it both erased a base runner and added two outs. As a generic out in a DP situation was worth –.3 runs, the DP was worth –.47 more. So +/-7 means about +/- 3½ runs in a year.

Walsh, John (2010). Philosophy of batting leadoff. <https://tbt.fangraphs.com/philosophy-of-batting-leadoff/>

Changes in time concerning what type of batter led off. All off diagrams. In the early 1950's, the leadoff batter was high OBA, about .365 whereas MLB average was about .355. By the mid 1960's, MLB average had dropped to about .335, but lead off collapsed to as low as .320 as teams looked for base stealers no matter the OBA. By the mid 1970's both were at about .340. Both went up and down in next decades, but leadoff hitters did so faster. As a consequence, leadoff hitters were then consistently close to 10 points higher through about 1995, when it reversed to 10 points higher for leadoff in the early 2000's. By 2010, leadoff OBAs again ahead, closing in on .360 with average about .350.

Weinstein, Max (2013). Who deserves credit for throwing out base-runners?

<https://www.beyondtheboxscore.com/2013/7/18/4522508/who-deserves-credit-for-throwing-out-base-runners>

Between 2002-2012 with at least 100 innings either pitched or caught, the year-to-year correlation of caught stealing was 0.39 for catchers and 0.25 for pitchers. For at least 200 innings as battery, catcher caught stealing rate correlated only 0.39 with battery rate, whereas pitcher rate correlated 0.79. The implication is that pitchers are more responsible for caught stealing rates than catchers. Overall, catcher and pitcher rates correlated 0.19.

Weinstein, Max (2013). Exploring the battery effect.

<https://blogs.fangraphs.com/exploring-the-battery-effect/#more-134161>

A WithOut an With You study encompassing 2002 to (I think) mid-August 2013 and including all battery combinations with at least 200 innings together and 150 separately Max computed a correlation of 0.52 between pitchers and wild pitches alone and in batteries, and 0.66 for passed balls. For catchers, correlations were 0.20 for wild pitches and 0.16 for passed balls. Combining WP and PB as independent variables led to pitchers alone correlating with battery PB at 0.83 and WP at 0.76; for catchers alone, 0.60 with PB and 0.51 with WP. Given that modern analysts do not distinguish between wild pitches and passed balls, PB and WP should have been combined throughout.

Weinstein, Max (2013b). 2013's top batteries at preventing the running game.

<https://blogs.fangraphs.com/2013s-top-batteries-at-preventing-the-running-game/>

Here, Max proposed a metric called baserunning Battery Runs Saved (bBRS). It includes three components; runs saved through number stolen base attempts, caught stealing, and stolen bases allowed, each relative to average and then summed. It is computed for each pitcher-catcher battery combination (not each pitcher and each catcher individually). As of mid-August 2013, the highest was Adam Wainwright-Yadier Molina (not a surprise) with 4.10, almost one run ahead of the second ranked battery. Molina also showed up in another battery in the top 25. The worst was John Lackey-Jarrod Saltalamacchia (part of the reason why Saltalamacchia disappeared so quickly,

along with his terrible pitch framing) at –6.25, a full 2 ¼ runs worse than the second worst. Each individually also show up in other batteries in the bottom 25 list.

Weinstein, Max (2013c). The overrated value of catchers' throwing arms.

<https://blogs.fangraphs.com/the-overrated-value-of-catchers-throwing-arms/>

The following tables are based on 2011 to 2012 data encompassing 7757 attempted steals with 5641 successful:

Catcher Pop Time (s)	CS%	SBA
1.6/1.7	43.56%	25%
1.8/1.9	51.00%	58%
2/2.1/2.2	53.00%	17%

The correlation between pop time and caught stealing percentage was a moderate –0.30.

Pitcher Release Time (s)	CS%	SBA
0.9/1.2/1.3	67.07%	15%
1.4/1.5	59.00%	37%
1.6/1.7	41.00%	39%
1.8/1.9/2	22.00%	9%

The correlation between pitcher release time and caught stealing percentage was a very high –0.88.

Combined Battery Time (s)	CS%	SBA
2.8/3/3.1	85.71%	7
3.2/3.3	52.78%	36
3.4/3.5	48.72%	39
3.6/3.7/3.8	33.33%	18

The correlation between pop time and release point summed was –0.81, which is very high but actually less extreme than for release point alone. The implication is that pitchers are largely responsible for caught stealing rates.

Weinstein, Max (2014e). How do we assign credit for catching base-stealers?

<https://tbt.fangraphs.com/how-do-we-assign-credit-for-catching-base-stealers/>

Using Retrosheet data, the author described the basis for a method for estimating a caught stealing percentage figure for pitchers and catchers separately while adjusting for game situation (inning/bases occupied/outs/score margin/batter and pitcher handedness). Consistently with Weinstein's earlier research, pitchers continued to have a much larger impact on attempts and on caught stealing than catchers.

Wigley, Jay (2021). Did batters of long ago learn during a game? *Baseball Research Journal*, Vol. 50 No. 1, pages 55-59.

Jay Wigley (2021), using Retrosheet data going back to 1916, uncovered the fact that the TTOP effect appeared for the first time in 1921, the beginning of the “modern” slugging era. In the five years previous to then, a second time dip was followed by a third time return to the level of the first time through.

Wolfersberger, Jesse and Matthew Yaspan (2015). Trying to quantify recency effect. In Dave Studenmund and Paul Swydan (Prods.), *The 2015 Hardball Times Baseball Annual* (pages 360--367). FanGraphs.

Among the many routes to exploring the issue of whether streakiness is a real phenomenon, one of the more useful ones is to see if more recent plate appearances (PA) are better predictors of a given PA's outcome than more distant-in-the-past PAs. 2013 Retrosheet data implies it does not, with the exception of the result of the immediately preceding PA, which authors Wolfersberger and Yaspan attribute to the increased tendency for both current and previous PA to be against the same pitcher.

Wolfson, Julian, Joseph S. Koopmeiners and Andrew DiLernia (2018). Who's good this year? Comparing the information content of games in the four major US sports. *Journal of Sports Analytics*, Vol. 4 No. 2, pages 153-163.

Woodrum, Bradley (2013). The changing caught-stealing calculus.
<https://blogs.fangraphs.com/the-changing-caught-stealing-calculus-2/>

Base stealing break-evens are predictably negatively related with home run rates, as a lot of homers make the steal attempt superfluous. Bradley Woodrum (2021) calculated the following regression equation for break-evens based on 1950 to 2012 (I am guessing Retrosheet) data at the level of seasons:

$$0.590 + (3.33 \times \text{HR/PA})$$

The correlation between the two was a very substantial -0.828 . This analysis averages over different strategic approaches across teams and as such is probably an overestimate of the relationship. It would be better done at the team-season level; I am betting the correlation would be less extreme but still substantial.

The authors used 2010-2015 data from Retrosheet to estimate how informative different proportions of games beginning at the start of the season (first 1/8 of the game, 2/8 of the games, etc.) are for predicting team matchups for the rest of the season. Even with 7/8 of the season finished (140 games), accuracy was never higher than 58 percent for the rest, which the authors note is not a lot higher than the 54 percent home field advantage, which they used as their comparison model.

Woolner, Keith (1999). Field general or backstop?: Evaluating the catcher's influence on pitcher performance. In Clay Davenport (Ed.), *Baseball Prospectus 1999* (pages 466-474). Washington, D.C.: Brassey's Sports. Available at <https://www.baseballprospectus.com/news/article/432/field-general-or-backstop-evaluating-the-catchers-influence-on-pitcher-performance/>

Woolner, Keith (2000). Catching up with the general: A postscript: A second look at catcher defense. <https://www.baseballprospectus.com/news/article/436/catching-up-with-the-general-a-postscript-a-second-look-at-catcher-defense/>

Several researchers, including Craig Wright anecdotally in his book with Tom House (*The Diamond Appraised*) and Tom Hanrahan in three articles in *By The Numbers*, uncovered evidence based on ERA that catcher performance improves with experience. The only nay-sayer that I am familiar with was Keith Woolner (1999). Using data from Retrosheet and Total Sports from 1984 through 1997, Keith performed WOWY analyses with every pitcher with each catcher with whom he faced 100 or more batters (sample size = 6347 pitcher/catcher combination). He then calculated the overall run value for the results of those plate appearances for each of the combinations. The distribution of these run values approximated the normal distribution fairly closely, implying that performance differences among catchers either do not exist or do exist but occur randomly. Further, the year-to-year correlation for catchers was a non-existent .02, meaning that performance changes randomly from year to year. Keith re-analyzed these latter data in several ways to see if a subtle effect hidden in the overall trend would appear; the correlations remained very close to zero. After reader criticism stating that his WOWY analysis was invalid because the comparisons were often with different catchers from year-to-year, Keith (2000) restricted it to teams with the same two catchers working with the same pitchers in consecutive years. The resulting correlation, 0.01, supported the original conclusion. I find it difficult to substantively reconcile Tom and Keith's very distinct conclusions.

Woolner, Keith (2001). Temperature and OPS. <https://www.baseballprospectus.com/news/article/1058/aim-for-the-head-temperature-and-ops/>

One thing that is almost certainly not a "skill" is a tendency to hit better in warm versus cold weather. Keith Woolner (2001), using data from Retrosheet and The Baseball Workshop, developed a database of 224 players with a least 100 PA in both cold and

warm weather for both 1999 and 2000, with $72\frac{1}{2}$ degrees at gametime as the cut-off, computed a ratio of cold/warm OPS for each player and then correlated these ratios for the two seasons. The correlation was actually slightly negative, -0.15 , implying a small tendency for batters to reverse tendencies from season-to-season. This was a quick-and-dirty study in which Keith did not control for players changing teams, and given what could be relevant characteristics specific to individual ballparks, such changes could well be responsible for the reason the figure was negative rather than close to zero.

Woolner, Keith (2001). Reaching on errors.

<https://www.baseballprospectus.com/news/article/1145/aim-for-the-head-reaching-on-errors/>

Woolner, Keith (2001). More reaching base on errors.

<https://www.baseballprospectus.com/news/article/1167/aim-for-the-head-more-reaching-on-errors/>

Using play-by-play data from 1978 through 2000 (I shall assume Retrosheet) for 1704 players with at least 500 PA, Keith Woolner (2001f) uncovered a year-to-year correlation for reaching base on error (ROE) of only 0.21, but a more robust 0.41 for odd versus even years across careers. In follow-up work (2001g), Keith noted right-handed hitters (1.23% of PAs) to ROE more than left-handed (0.95%), with switch-hitters intermediate (1.12%); a correlation of 0.262 between ROE and groundball/flyball ratio and ROE but only 0.04 between ROE and grounding into double plays (incidentally, the ratio and GDP correlated at a surprisingly low 0.148).

Woolner, Keith (2001). Response rates.

<https://www.baseballprospectus.com/news/article/1077/aim-for-the-head-response-rates/>

Teams that score a lot tend to score in a lot of innings, and teams that score in a lot of innings tend to not be victims of shutdown innings. Between 1978 and 2000, the correlation between the latter two was .89. There was only a little evidence for a discernible team skill involved in answering opposition runs with one's own; the differences between the proportion of innings scored in and the proportion of such innings following opposition scoring was +0.17.

Woolner, Keith (2001). Walk rate spikes.

<https://www.baseballprospectus.com/news/article/1107/aim-for-the-head-walk-rate-spikes/>

Keith Woolner (2001) examined whether sudden increases or decreases in offensive production were signals of actual skill change rather than one-year flukes. To do so, he looked all but two of players from 1954 through 2000 (I am willing to bet that he used Retrosheet data) who amassed at least 1000 PA in a three-year span, a fourth season of at least 400 PA spike, and years five through seven (again minimum 1000 PA); the

two players, Ozzie Smith (spike in 1982) and Frank Tavares (spike in 1977) had zero homers their first three seasons, which would make the result moot. There were 3220 relevant player-spans His method was:

Step 1 – compute rate per PA for a given metric across the first three years.

Step 2 – compute difference in this rate between year four and result for Step 1.

Step 3 – compute rate per PA of the metric for years five through seven.

Step 4 – compute difference in this rate between years five through seven and result of Step 2

Step 5 – compute the difference between the results of Step 2 and Step 4.

It turned out that the results of Step 5 correlated at 0.42 for hits, 0.47 for homers, 0.45 for total bases, 0.51 for strikeouts (with decreases interpreted as improvements), 0.45 for on-base average, and 0.48 for walks. This means that there was some actual skill change, approximately equivalent across the six indices, that had some degree of consistency across players. Keith then replaced the original Step 5 with a different final step:

Step 5* – divide the result of Step 4 by the result of Step 2. This gives you the proportion of the year four spike that was retained in the next three years, a more accurate measure of actual skill improvement than the original subtraction because it is not affected by the size of the metric; for example, that there are so many more hits than homers. This gives you the proportion of the spike that was retained in the following three seasons; Keith called this the “retention percentage.”

He then looked at the “retention percentage” for the 300 players with the biggest gains and 300 with the worst losses, with the following outcome:

Skill	Players who increased in skill	Players who decreased in skill
H	21.9%	45.8%
HR	41.5%	47.5%
TB	29.3%	51.4%
SO	59.7%	43.1%
OBP	42.6%	37.4%
WALK	51.7%	42.1%

It appears that for the 300 biggest improvements, the retention percentage was greatest for those most closely associated with the Three True Outcomes (HR, SO, OBP, BB) and somewhat less so for hits and total bases. That for the 300 largest decreases were more equivalent across the board. Keith admitted that there is some bias in these findings as they did not include ballpark or league effects, which could be considerable in a few cases; i.e. a player moved to Colorado in the spike year and stayed there the following three.

Woolner, Keith (2002). Set lineups.

<https://www.baseballprospectus.com/news/article/1339/aim-for-the-head-set-lineups/>

Between 1978 and 2000, Keith Woolner (2002, probably using Retrosheet) data uncovered instances of teams using as few as 42 (Red Sox, 1984) and as many as 155 (Angels, 1985) different lineups across a season. The correlation between number of lineups and team wins was a credible -0.39 , as one would expect that worse teams would try different combinations looking for a winner. That Red Sox team had one specific lineup that appeared in 66 games; there was one unnamed teams that only used the same exact players in the same batting order twice. Only one team winning less than 70 games used a specific lineup more than fifteen times; in contrast, teams winning more than 90 were more variable in this matter. In this case, the correlation was 0.33 .

Woolner, Keith (2002). Quality Starts.

<https://www.baseballprospectus.com/news/article/1623/aim-for-the-head-quality-starts/>

Keith Woolner (2002) presented some interesting relevant information about Quality Starts for the 1978 to 2000 period (likely using Retrosheet as the source). From year to year, the proportion of starts that met the definition was usually in the mid to upper 40s and occasionally lower 50s during those years. When a QS occurred, team winning average was in the upper 60s and low 70s, with that for the pitcher getting credit for the win in the mid and upper 50s and occasionally lower 60s. While none of the above trended a lot during that period of time, the odds of the pitch getting saddled with the loss went up; in the low 20s during the early 1980s, it began approximating 25 percent by the late 1980s and was well over that by the mid 1990s. That decrease was compensated for by an increase in no-decisions for the starter, rising from about 21 percent to about 26 percent in the interim.

Woolner, Keith (2003). A big change for OBP.

<https://www.baseballprospectus.com/news/article/1759/aim-for-the-head-a-big-change-for-obp/>

Click, James (2004). Another look at OBP: Do speedy runners force more errors?

<https://www.baseballprospectus.com/news/article/2981/another-look-at-obp-do-speedy-batters-force-more-errors/>

Keith Woolner (2003) proposed adding reaching first on fielder's choices to OBA because, as they are the runner's fault rather than batter's, the batter should not be penalized for the out. The reason they are not included because they are only supposed to be assigned when the official scorer thinks that the batter would have been out otherwise. And they occur more often than you might think: for players with at least 300 PA between 1978 and 2000, they occurred 28.48 percent more often than HBP and 16.62 percent more often than SF. James Click (2004e) examined the issue of whether

getting on base due to opposition error should also be included, under the assumption that if faster runners do so more often then it reflects an offensive skill. However, he uncovered no evidence that they do so (but see Woolner's "Reaching on errors" above) and thus no good rationale for the inclusion.

Wyers, Colin (2008f). A run saved is a run earned? <https://tth.fangraphs.com/a-run-saved-is-a-run-scored/>

Using Retrosheet game logs from 1953 through 2008, Colin Wyers (2008) starting with all teams with at least a 50-run differential between runs scored and a runs allowed in a season, matched teams playing the same number of games with the same exact differential at a given point during the season starting at 100 games played, one with an offensive and one with a defensive surplus; an example would be a team with 51 more offensive runs scored and a team with 51 more defensive runs saved having both played 101 games. Colin then figured out the average W-L based on what these teams did in game 102. Offensive surplus teams averaged .5024 W-L, defensive surplus teams averaged .5039. The average Pythagorean error was .013 for offensive surplus teams and .022 for defensive teams. Perhaps this is evidence that defensive teams can take advantage of their superiority in that aspect of the game a bit better than can offensive teams.

Wyers, Colin (2008). A quick look at baserunning.
<https://statspeakmvn.wordpress.com/2008/11/page/2/>

Using 1953-2007 Retrosheet data, Colin Wyers (2008) evaluated baserunning by summing run expectancies for every possible outcome given a relevant play. His example was a runner on first, no outs, single to left, in which the base runner could end up on second, on third, or be thrown out on the bases, each with a corresponding impact on run expectancy. The annual scale range was +/-10.

Wyers, Colin (2008g). A 10th man? <https://tth.fangraphs.com/a-tenth-man/>

Colin Wyers (2008) used Retrosheet data (years not stated) and Bill James's log5 method to compute expected team winning averages of home teams in specific games, then compared actual team W-L across games with at least 1000 more, 1000 fewer, or closer than those in attendance than team average. The actual winning averages were .555 for 1000 above, .544 closer to average, and .524 for 1000 below attending. Making the differential 15,000 extra or fewer than average, resulted in .557, .541, and .504 winning averages. In all cases the expected W-L was within 3 percentage points of .500, as it should be. This is strong evidence in support of fan support as one of the causal factors in home field advantage.

Wyer, Colin (2008). A new framework for offensive evaluation: Total Production.
<https://statspeakmvn.wordpress.com/2008/10/>

Colin Wyers (2008) offered a bottom-up regression-based method that he called Total Production using 1994-2007 Retrosheet data. It is pretty similar to the many others then available, and, to be honest, at this point in time was superfluous.

Wyers, Colin (2008). Run expectancy by count.

<https://statspeakmvn.wordpress.com/2008/11/page/3/>

Here is a run expectancy chart for the end of plate appearances at each count, from Retrosheet 1994-2007 data.

BALLS	STRIKES	RUNS
0	0	0.595
0	1	0.556
0	2	0.390
1	0	0.592
1	1	0.554
1	2	0.397
2	0	0.614
2	1	0.560
2	2	0.405
3	0	0.842
3	1	0.713
3	2	0.560

Wyers, Colin (2009). The best run estimator. In Dave Studenmund (Producer), *The Hardball Times Baseball Annual* (pages 209-215). Skokie, IL: Acta Sports.

It is customary to compare specific methods for evaluating offense, but most of them are of little value because they are limited to a given period of seasons and thus biased towards those methods that were designed in the context of those seasons. A better idea is to evaluate classes of methods to see which class works better. Wyers (2009) offered a thoughtful such attempt, showing not only that but why a method such as base runs will be more accurate than runs created or extrapolated runs using a data set big enough (all Retrosheet data from 1956 through 2007) to counter the problem of formulas designed for a specific sample of years.

Wyers, Colin (2009). How well can we predict ERA? <https://tht.fangraphs.com/how-well-can-we-predict-era/>

A study of the reliability of four metrics (ERA, FIP, xFIP, and tRA) based on 2003 to 2008 Retrosheet data, comparing performance on even versus odd calendar days, in which the root mean square error is used rather than a split-halves correlation. xFIP

was the winner (1.78), with tRA (1.87) and FIP (1.91) not far behind, and ERA taking up the rear (2.32).

Wyers, Colin (2009). What's a batted ball to do. <https://tth.fangraphs.com/whats-a-batted-ball-to-do/>

This is a study of the year-to-year correlations for different batted ball types across pitchers, in order to see if pitchers have any control over them, based on Retrosheet data. Between 1989 and 1999, the percentage of grounders hit in infielder territory (3,4,5,6) rather than between them (34,4M,56,6M) from year to year 0.07, in other words nothing, implying an absence of control. evidence for DIPS. Between 2004 and 2008, the year-to-year correlation for batting average on balls in play was only 0.096; for BABIP on grounders only just 0.104, with similar implications. However, the BABIP on fly balls was 0.457, implying differing skill across pitchers involving whether flies become hits. Combining flies with infield pops hit into "no-man's land" correlated at 0.2, so perhaps a tiny amount of pitching skill is involved there.

Wyers, Colin (2009). Putting the scissor to defense (part 1). <https://tth.fangraphs.com/fleeter-than-birds-part-1/>

Here is Colin's method for estimating infielder plays made per opportunity for seasons before batted ball location was available:

first base: $(.85 \times \text{assists}) + (.08 \times \text{pops})$

second base and shortstop: $(.85 \times \text{assists})$

third base $(.90 \times \text{assists}) + (.06 \times \text{pops})$

He also proposed a method for seasons with batted ball location: see the webpost if interested. He used Retrosheet data to produce best and worst lists: Brooks Robinson came in first (+382.8); Eddie Yost last (-171.1) followed by Derek Jeter (-157.8).

Wyers, Colin (2009). When is a fly ball a line drive? <https://tth.fangraphs.com/when-is-a-fly-ball-a-line-drive/>

It has become clear that the height of press boxes has caused variation in coder judgments concerning whether batted balls to the outfield were flies or liners. Colin Wyers (2009b), using Retrosheet data from 2005 to 2009 for visiting teams excluding pitcher at bats, noted a correlation of .16 between press box height and line drive rate (I think as a proportion of liners plus flies), and jumps to .38 with the exclusion of the five most extreme parks in either direction, in which coders acted as if they were compensating for the problem. As the difference in overall odds of making plays on each are so great, the resulting ratings for outfielders have probably been significantly affected.

Wyers, Colin (2009). What's a batted ball to do. <https://tth.fangraphs.com/whats-a-batted-ball-to-do/>

Colin Wyers's (2009e) table of outs rate for grounders at different locations supports the supposition that batters should hit 'em where they ain't, aka the infield holes...

ZONE	OUTS
3	0.8697
34	0.6212
4	0.9177
4M	0.5376
5	0.7535
56	0.5894
6	0.8773
6M	0.5851

...and, given these out rates for different outfield distances, in “no-man's land”:

DEPTH	OUTS
Infield grass	0.9581
Infield dirt	0.8866
Behind the infield	0.5343
Shallow outfield	0.4467
“Normal” outfielder positioning	0.7369
Deep outfield	0.8340
Very deep outfield	0.7925

Wyers, Colin (2011). Lost in the SIERA madre.

<https://www.baseballprospectus.com/news/article/14603/manufactured-runs-lost-in-the-siera-madre/>

The following are root mean square errors at increments of 100 innings for 2010 based on 2003-2009 aggregated Retrosheet data; note that differences were dependent on innings in estimate:

<u>IP</u>	<u>ERA</u>	<u>FIP</u>	xFIP	<u>SIERA</u>
100	1.52	1.45	1.40	1.33
200	1.20	1.19	1.18	1.14
300	1.02	1.07	1.05	1.03
400	.98	1.02	1.02	1.02

500	.77	.88	.90	.91
600	.75	.88	.93	.96
700	.81	.91	.93	.99
800	.77	.88	.91	.96
900	.77	.90	.91	.98
1000	.81	.94	.97	1.04

SIERA is the best for perhaps a season's worth of starting pitching ERA but not more; all about even at 400 IP, and note how ERA eventually does the best!!

Wyers, Colin (2011). When LaRussa should pinch-hit.

<https://www.baseballprospectus.com/news/article/15257/manufactured-runs-when-la-russa-should-pinch-hit/>

Tango, Tom aka Tangotiger (2011).

https://www.insidethebook.com/ee/index.php/site/run_expectancy_pitchers_and_pinch_hitters

The following are differences in run expectancy difference between pinch-hitters in the ninth position and the pitchers usually occupying slot:

Baserunners	0 Out	1 Out	2 Out
None	0.01	0.04	0.04
First	0.05	0.12	0.07
Second	0.11	0.08	0.09
Third	0.06	0.18	0.14
First and second	-0.06	0.18	0.19
First and third	0.17	0.24	0.18
Second and third	0.19	0.24	0.29
Loaded	0.28	0.32	0.3

The difference generally becomes noticeable with more runners on base and at least one out. There is one situation with pitchers more productive, I am assuming due to the sacrifice bunt having some value here. Tom Tango aka Tangotiger (2011), which linked to Colin's post, included very similar figures for (he thought) 1956 to 2007.

Wyner, Adi (2021). Is the 3rd time through the order effect real? Correcting for lineup order and pitcher quality selection bias. SABR Analytics Conference

Using 2010-2019 Retrosheet data, Adi Wyner (2021) reported that there was no third time through the order effect as such, as after a “settling in” process over the first few PAs, performance decrements on average begin with the first time through the bottom third of the order and continue steadily thereafter.

Zardkoohi, Asghar, Michael W. Putsay, Albert Cannella, and Michael Holmes (n.d.). Cognitive biases in decision making: Risk taking in major league baseball, 1985-1992. College Station: Texas: Department of Management, Mays Business School, Texas A & M University.

Base stealing is a one-run strategy, and as such the attempted steal should be used late in games and, in general, in high leverage situations. However, Zardkoohi, Putsay, Cannella, and Holmes (n.d.) analyzed Retrosheet data from 1985 through 1992 of more than 200,000 situations with a runner on first only and concluded that steal attempts were actually more numerous earlier in games rather than later and increased as run differentials increase from three runs behind through tied scores to three runs ahead. The authors relate this to psychological tendencies to be risky about positive things and cautious about negative things (see work on prospect theory by psychologists Amos Tversky and Daniel Kahneman, the latter a Nobel prize winner in Economics as a result of this and analogous work), such that managers are more likely to feel comfortable risking a steal when ahead than behind and when there are plenty of innings left to overcome a caught stealing then when innings are running out. Zardkoohi et al. also noted more attempts against righthanded pitchers and when there had been previous success against the same pitcher or catcher, none of which are surprising.

Zhang, Xing, Tat Y. Chan, and William P. Bottom (2022). Relational aspects of vicarious retribution: Evidence from professional baseball. *Journal of Applied Psychology*, Vol. 107 No. 6, pages 917-931

There have been several psychological studies of strategic (rather than accidental) hit by pitches as indicators of aggression, and a few specifically directed toward studying pairs of initiating HBPs followed by retaliatory HBPs as a cycle of provocation and score-settling retribution. This one is a sophisticated example of the latter set. The authors' work was based on a sample of about 20,000 HBPs from 1991-2010 games posted on Retrosheet. The data generally supported this perspective, as the occurrence of an initialing HBP increased the odds of a retaliatory one, on average an inning later, after which the odds of additional ones decreased (two retaliatory HBPs in response to the initial one were relatively rare. The authors also looked for the impact of possible measures of similarity affecting the cycle. Retaliation was more likely if the hit batter and teammate-pitcher were both from outside the U.S., and less likely if the

teammate-pitcher and initiating team batter were both from outside the U.S., had been teammates in the past. In addition, retaliation was less likely if both had not attended college whereas college attendance for both actually increased retaliation; these two findings make no sense to me. Some results for control variables are also of interest, as retaliation more likely the greater the score difference, the better the retaliatory team's winning average, the retaliatory team being at home, being on more diverse teams, and in the American League. Perhaps unknown to these authors, some of these latter findings replicate earlier work; in particular, the A.L. one appears to be due to the presence of the DH leading to pitchers knowing that they will not bat and so won't be direct victims of retaliation.

Zimmerman, Jeff (2011). BABIP and home field advantage.

<https://blogs.fangraphs.com/babip-and-home-field-advantage/>

Jeff Zimmerman (2011) noted BABIP to be higher at home (.303) than away (.296) between 2006 and 2010. The following is from Retrosheet for 2006 through 2009:

Team	Batted Ball Type	BABIP	% of Hits	% of Batted Balls
Home	LD	0.726	47.5%	19.9%
	FB	0.156	15.3%	29.6%
	GB	0.262	36.6%	42.4%
	PU	0.022	0.6%	8.1%
	Total	0.304	100.0%	100.0%
Away	LD	0.721	47.3%	19.5%
	FB	0.156	15.6%	29.6%
	GB	0.253	36.5%	42.8%
	PU	0.023	0.6%	8.1%
	Total	0.297	100.0%	100.0%
Differences	LD	0.005	0.2%	0.4%
Home – Away	FB	0.000	-0.3%	0.0%
	GB	0.009	0.1%	-0.4%
	PU	-0.001	0.0%	-0.1%
	Total	0.007	0.0%	0.0%

Note that the home team hit a few more liners and fewer grounders, the former meaning more hits.

Zimmerman, Jeff (2009). What factors have an effect on runs scored at MLB parks?

Part 2. <https://www.beyondtheboxscore.com/2009/1/7/713479/what-factors-have-an-effect-on-runs-scored-at-mlb-parks/>

2006-2008 data. In part 1, Jeff used runs scored per game to represent park factors, but after a lot of comments switched to a method previously used by Brandon Heipp aka Patriot. This was the dependent measure in a multiple regression where the

independent variables are possible factors affecting it. Rather than showing the (confusing) equation, here is his interpretation of what its results mean:

<u>Factor</u>	<u>Change in Park Factor</u>	<u>Change Runs Scored per game (9.54 runs per game)</u>
10 degree F increase	0.0077	0.073
Increase in RH by 10%	0.0120	-0.115
10,000 sq ft increase in foul area	-0.0061	-0.058
Surface is Turf	0.0090	0.085
1000 ft increase in elevation	0.0206	0.196
1 Errors for Away Team	0.0160	0.150
10 ft increase in LF	-0.0100	-0.095
10 ft increase in LC	-0.0063	-0.060
10 ft increase in CF	-0.0101	-0.096
10 ft increase in RC	-0.0020	-0.019
10 ft increase in RF	0.0106	0.101

This accounted for 69.2 percent of variance. There were two additional variables that he removed, as the impact of other variables led these two to have counter-intuitive impacts; wind (a negative regression component implies more wind blowing out means LESS scoring) and wall height (a positive component implies more scoring for higher walls, whereas this would replace some homers with less productive doubles).

Zimmerman, Jeff (2013b). First inning home field advantage.

<https://blogs.fangraphs.com/first-inning-home-field-advantage/>

Beginning in 1950 and up to late 2013, first inning strikeout rate minus walk rate was 6.8 for home team pitchers and 3.6 for away team pitchers. In later innings, the difference between the two shrank, and was down to 0.5 in the ninth inning. First inning fastball velocity was 0.15 mph higher for home team pitchers. Jeff believed that these home

field advantages occurred because the home team pitcher has a shorter delay between before-game warmup and actual pitching.