

APPENDIX 11: RUNS SCORED PER INNING FOR ALL GAMES; AND RUNS SCORED PER INNING FOR THOSE GAMES LASTING 20 OR MORE INNINGS

Runs scored per inning for all games, tabulated by inning for innings 1-45, based on Smith's major league data for innings 1-19 and Lowry's 20 innings and more data for innings 20-45:

INNING	INNING DATAPOINTS	RUNS PER INNING
SMITH		
1	245,812	0.561
2	245,812	0.431
3	245,812	0.491
4	245,812	0.497
5	245,812	0.491
6	245,446	0.507
7	244,856	0.493
8	244,088	0.486
9	243,394	0.433
10	23,026	0.399
11	12,906	0.398
12	7,220	0.385
13	4,044	0.396
14	2,282	0.397
15	1,240	0.392
16	704	0.380
17	378	0.384
18	202	0.302
19	116	0.440
LOWRY		
20	412	0.279
21	254	0.319
22	156	0.314
23	96	0.250
24	62	0.145
25	48	0.167
26	32	0.156
27	24	0.250
28	18	0.111
29	14	0.143
30	10	0.300
31	8	0.250
32	8	0.250
33	6	0.167
34	4	0.000
35	4	0.500
36	4	0.000
37	2	0.000
38	2	0.000
39	2	0.000
40	2	0.000
41	2	0.000
42	2	0.000
43	2	0.000

44	2	0.000
45	2	0.500

Runs scored per inning for just those games lasting 20 or more innings, tabulated by inning for innings 10-45 based on Lowry data for all games which have gone 20 or more innings and have a box score:

INNING	INNING DATAPOINTS	RUNS PER INNING
LOWRY		
10	412	0.019
11	412	0.039
12	412	0.024
13	412	0.019
14	412	0.010
15	412	0.029
16	412	0.005
17	412	0.015
18	412	0.034
19	412	0.063
20-45	see above	see above

Parts of this database are less valid than others simply because of a lower number of existing datapoints. Smith's database has a total of 2,259,116 datapoints. It starts with 245,812 inning datapoints (which represents 122,906 games) in the first inning and remains there through the fifth inning. As games start to get called due to weather and darkness, the number of datapoints falls gradually to 243,394 in the ninth inning. That number then falls off dramatically to 23,026 in the tenth inning, 12,906 in the eleventh, etc. By the time it gets to the 19th inning, it has fallen to only 116 datapoints.

Similarly, my database on all games that have lasted 20 or more innings has a total of 5,298 datapoints. It starts with 412 inning datapoints (which represent all 206 games lasting 20 or more innings for which box scores exist) in the tenth inning and remains there through the 20th inning. Then it falls off to 254 datapoints in the 21st inning, and by the 28th it is down to only 18 datapoints, meaning that we only have box scores for 9 games, or 18 inning datapoints, which have gone 28 or more innings.

As the Smith data for runs per inning approaches inning 20, and as the Lowry data approaches inning 30, we begin to see significantly higher variability in runs per inning, simply due to a smaller number of datapoints.

The Smith data shows 0.384 in the 17th inning, then plummets down an unrealistic 27% to 0.302 in the 18th, and then zooms up an unrealistic 46% to 0.440 in the 19th. Similarly, the Lowry data shows 0.156 in the 26th inning, then zooms up an unrealistic 60% to 0.250 in the 27th, and then plummets down an unrealistic 56% to 0.111 in the 28th.

However, it must be remembered that this is the very best database available. For innings 20-45 it is 100% complete data, based on all box scores in existence for games lasting 20 or more innings.

The data shows that run production is 0.000 for innings 34 and 36-44 because no team has ever scored in those innings. Only four teams have ever played a

36th inning; only two teams have ever played innings 37 through 45.