wSLG: Redefining the Relative Values of Various Hits in Baseball

Alex Cardonick[1], Mac Carso[1], Rick Cleary[2,*], Jacob Eckerle[1], Steven J. Miller[1], Ethan Prout[1], and Evan Ruschil[1]

[1] Department of Mathematics and Statistics, Williams College, Williamstown, MA, [2] Mathematics, Analytics, Science and Technology Division, Babson College, Wellesley, MA.

*Corresponding e-mails: rcleary@babson.edu and sjm1@williams.edu

Abstract
Slugging percentage (SLG) has been a popular baseball statistic for decades and important today as one of two components of the ubiquitous OPS (On-base plus Slugging) statistic. The traditional SLG used in OPS uses a linear combination of singles, doubles, triples and home runs, with the weights equal to the number of bases the batter advances on each type of hit. A study of the process of scoring runs leads us to suggest alternate weightings that more accurately reflect the value of each type of hit. We present two versions of weighted slugging percentage (wSLG) with these redefined weights. Both versions suggest that in traditional SLG, and thus in OPS, singles are undervalued relative to doubles.

Keywords: Slugging percentage, linear combinations, Markov chains, expected runs.

---

1 Introduction: A Traditional Statistic and a Rationale for an Alternative

In Major League Baseball (MLB) today, players' offensive capabilities are measured by certain key metrics. These include batting average (AVG), which has been used for generations, and some relatively new measures like on-base percentage (OBP), which gained popularity as interest in analytics grew late in the 20th century. Another common measure, slugging percentage (SLG) has been used even longer than OBP (see [Crashburnalley (2025), Schell (2016), Vanderwerken (2021), Wikipedia (2024)] for a history, some issues, and an analysis of the greatest sluggers of all time). Batting average and on-base percentage are commonly used metrics, as they show how successful a player is at getting hits or getting on base, both of which are key contributions to a team's offense. These statistics, however, do not account for a hitter's ability to produce extra-base hits. Certainly, a home run provides a larger contribution to a team's offense than a single, but both would have the same effect on a player's batting average or on-base percentage. The slugging percentage metric attempts to remedy this by calculating the number of bases per at-bat:

$$SLG \quad = \quad \frac{1\times(singles) + 2\times(doubles) + 3\times(triples) + 4\times(home\ runs)}{at-bats}.$$

This statistic has become much better known as hitters are now often judged by their "slash line", listing their batting average, on-base percentage and slugging percentage separated by slashes. As an example, Boston Red Sox outfielder Jarren Duran had a 2024 slash line of .285/.382/.492. Many observers like to combine the last two parts of the slash line into a single statistic, the on-base plus slugging (OPS) metric, so OPS = OBP + SLG, so Duran's OPS for the season was 0.834. The OPS formula gives equal weighting to the two variables.

SLG values certain types of hits more than others in order to try to provide a better representation of a player's offensive contributions. Clearly, a home run should carry the most weight, followed by a triple, then a double, then a single. The weights of four, three, two, and one certainly make the statistic very easy to calculate and follow the correct intuition for calculating a player's offensive capabilities, with the weights equivalent to the number of bases the batter advances. However, these integer weights are too simple to truly provide the best measure of offensive contribution. While slugging percentage, sometimes called "total base percentage", was cited occasionally for many years, it was in the 1950s that it became most commonly reported. At that time simple integer weighting made sense, but in the 21st century it is possible both to find better weights and then to use them. In this paper our suggested alternative weightings will produce measures that are consistently smaller than traditional SLG, and this in turn suggests that perhaps SLG is "over valued" in the traditional OPS calculation.

Small but meaningful tweaks to established statistics are standard practice in baseball analytics. A well-known example of improving a metric by replacing integer values with decimals that give better results involves a modification to the Pythagorean winning percentage (PWP) formula. This formula suggests that a team's winning percentage should be nearly equal to the ratio of their runs scored, squared, to runs scored, squared, plus runs allowed, squared:

PWP = $(\text{runs scored})^2 / ((\text{runs scored})^2 + (\text{runs allowed})^2)$

This formula is a useful predictor of winning percentage, and teams that outperform their expectation through the first half of a season are good bets to regress slightly in the second half. But an even more accurate version is observed by replacing the "square" in the formula with a value of about 1.87 [Baseball-Reference (2024)].

We derive two versions of an augmented slugging percentage metric, weighted slugging, which we will denote as wSLG. Each version provides weights that better capture the relative value of singles, doubles, triples, and home runs and, in future work, these weights can be tailored to a team's particular tendencies. We do this by considering how much each type of hit impacts the team's expected runs scored in a game. Since runs scored are the defining characteristic of whether a team wins or loses, we will use the change in expected runs to define the value of a particular hit.

Our first version is computed under the assumption that base runners advance as many bases as the batter, and then we allow for additional base runner advances, for instance a runner on first scoring on a double. Our key result is that in either version, under reasonable assumptions, a home-run is worth a bit more relative to a triple than the standard slugging suggests, while a double is not worth quite as much relative to a single. As triples are rare the consequences of the first result are minor; however, as singles and doubles occur frequently even a small improvement can have significant impacts on the run production of a team and the evaluation of a player's production.

These small improvements can be especially valuable now that all major league teams have analysts who understand the fundamentals of making good decisions with data. Organizations use increasingly sophisticated approaches as they try to gain modest but potentially important competitive advantages. Our approach is similar to the work done on weighted on-base average as seen in Slowinski (2024).

2 Mathematical tools: Expected values and Markov chains

Baseball games are readily modeled using Markov processes [Bukiet (1997), Statshacker (2023)]. We use two key tools to derive appropriate weights for our wSLG statistic. The first is a run expectancy matrix with values of the expected runs in the rest of the inning from any possible base runner configurations, and a game state probability matrix that includes probability of being in each configuration. Both matrices are based on data from the 2010-2015 MLB seasons. From this, we calculated the effect of each type of hit on a team's expected runs in a given inning in order to calculate how each type of hit should be weighted.

The run expectancy matrix (Table 1a.) measures the change in the expected number of runs scored from the beginning of a batter's at-bat to the end of it. Run Expectancy (RE) is the number of runs a typical team scores in the remainder of the inning based on the number of outs and location of baserunners. For example, at the beginning of an inning with 0 outs and no one on base, denoted by 000, we would expect the total number of runs scored to be 0.481. If a team has already scored one or more runs in an inning, we add the runs scored to the expected runs to get an updated value. For example, if a player leads off an inning with a home run then we remain in the same state, but the expected runs increases to 1.481.

As expected, these estimates generally decrease as the number of outs increase but increase as the number of baserunners increase. The game state probability matrix (Table 1b) shows the probability of the 24 different scenarios of a standard 9-inning MLB game. To give an interpretation, the probability of being in the game state "nobody out, nobody on" is 0.244. It is not surprising that this is the most likely state as it is the only one that must occur in each inning. The run expectancy and game state probability matrices will be used to calculate the relative values of a single, double, triple and home run and are fundamental to our analysis.

Tables 1a and 1b.) Table 1a (left) shows run expectancy based on base/out state; and table 1b (right) shows the probability of being in that state at any given time.

Run Expectancy

	Outs 0	Outs 1	Outs 2
000	0.481	0.254	0.098
100	0.859	0.509	0.224
020	1.100	0.664	0.319
120	1.437	0.884	0.429
003	1.350	0.950	0.353
103	1.784	1.130	0.478
023	1.964	1.376	0.580
123	2.292	1.541	0.752

Game State Probability

	Outs 0	Outs 1	Outs 2
000	0.244	0.175	0.139
100	0.059	0.070	0.071
020	0.015	0.026	0.033
120	0.014	0.025	0.031
003	0.002	0.009	0.014
103	0.005	0.011	0.016
023	0.003	0.007	0.008
123	0.004	0.009	0.011

Our initial inspiration for the development of a model that seeks to understand the true value of a single, double, triple, and home run arose from the fact that baseball can be modeled as a Markov chain. The Markov model uses the transition probabilities that represent the chance of moving from any one of 24 possible game states (eight possible base runner configurations, each with one, two or three outs). A 25th state can be considered as the "three outs, inning over" state. The Markov model has been applied in the analysis of run distributions and the predictive modeling of the number of games that a team should win. The run expectancy matrix generated from the 24 game states provides the basis to analyze the weights of slugging percentage. The simplified nature of the slugging percentage weights suggests a more detailed analysis could provide a more accurate measurement of the true value of a single, double, triple, and home run. In our review of previous literature, we have yet to find an event-based approach that explicitly seeks to analyze the true value of various hits. Thus, this paper presents our approach to modifying slugging percentage with alternative weights to more accurately represent the value of a single, double, triple, and home run.

3 Derivation of appropriate weights: Fixed baserunning

Yet our analysis begins by first comparing the relative value of a triple versus a home run. This comparison provides the simplest starting point for beginning to evaluate the relative weights of all types of hits. For both hits, every runner currently on base will score. Therefore, we know exactly how many runs will score every time a triple or home run is hit. Additionally, we also know the exact game state that will result from a triple and a home run. A triple will clear all the bases and leave the batter at third base, resulting in a "003" game state, meaning one runner on third base. A home run will clear all the bases, including the batter, which results in a "000" game state,meaning no one on base. This example means we do not have to make any assumptions about baserunning, avoiding the need for adjustments required when there may be multiple resulting game states. Finally, in all the analyses below we ignore the effects of errors so we do not worry about a triple with say a throwing error, allowing the batter to score, which is rare at the major league level.

From this basic understanding of how the triple and home run interact with the run expectancy matrix, we can then use this to provide the relative gain from a triple and a home run from each of the 24 game states. This calculation uses what we call the RE24 statistic, which is a run expectancy-based measure that uses the 24 game states. The calculation is as follows:

$$\text{RE24} := \text{RE End State} - \text{RE Beginning State} + \text{Runs Scored}$$

As described previously, the RE End State will remain fixed for a triple with the values 1.350 (0 outs), 0.950 (1 out), and 0.353 (2 outs), and similarly for a home run with the values 0.481 (0 outs), 0.254 (1 out), and 0.098 (2 outs), but with an additional run scored compared to the triple. This calculation will be done for beginning state run expectancy for each of the 24 game states in the matrix for both home runs and triples resulting in the average contribution for a triple and a home run from each game state. We then normalize the values above depending on the probability of being in each of the game states to begin with, which we obtain from the game state probability matrix. Tables 2 and 3 show the results.

Tables 2a (left) and Table 2b (right): The gain from in run expectancy for each game state for triples and home runs.

Triple Gain

	Outs 0	1	2
000	0.869	0.696	0.255
100	1.491	1.441	1.129
020	1.250	1.286	1.034
120	1.913	2.066	1.924
003	1.000	1.000	1.000
103	1.566	1.820	1.875
023	1.386	1.574	1.773
123	2.058	2.409	2.601

HR Gain

	Outs 0	1	2
000	1.000	1.000	1.000
100	1.622	1.745	1.874
020	1.381	1.590	1.779
120	2.044	2.370	2.669
003	1.131	1.304	1.745
103	1.697	2.124	2.620
023	1.517	1.878	2.518
123	2.189	2.713	3.346

Tables 3a (left) and 3b (right): The gain metrics multiplied by the probability that a game is in the corresponding state. The sum of all entries is the expected gain for each type of hit.

Triple Gain - Normalized

	Outs 0	1	2
000	0.212	0.122	0.035
100	0.088	0.101	0.080
020	0.019	0.033	0.034
120	0.027	0.052	0.060
003	0.002	0.009	0.014
103	0.008	0.020	0.030
023	0.004	0.011	0.014
123	0.008	0.022	0.029

Expected Triple Gain 1.033

HR Gain - Normalized

	Outs 0	1	2
000	0.244	0.175	0.139
100	0.096	0.122	0.133
020	0.021	0.041	0.059
120	0.029	0.059	0.083
003	0.002	0.012	0.024
103	0.008	0.023	0.042
023	0.005	0.013	0.020
123	0.009	0.024	0.037

Expected HR Gain 1.420

The result we find is a normalized "Expected Triple Gain" of 1.033 and a normalized "Expected HR Gain" of 1.420. The meaning of these values is that on average, accounting for the probabilities of the 24 game states, over time we can expect a triple to contribute 1.033 runs and a home run to contribute 1.420 to any given inning. The more important results are in how our model values the triple relative to the home run versus how SLG values this ratio below.

Table 4: The difference in value of a triple and home from the probability model, compared to the standard slugging percentage formula

Model Triple/HR Value	**SLG Triple/HR Value**
0.728	0.750
Model HR/Triple Value	**SLG HR/Triple Value**
1.374	1.333
% Difference	
3.08%	

.

Clearly a triple is less valuable than a home run; the question we are addressing is how much less? Based on these results in Table 4, we find that SLG overvalues the triple relative to the home run since our model ratio is smaller than 0.750, representing a 3.08% overvaluation of the triple relative to SLG's value. A difference on the order of 3% is not particularly large and, since there are few triples, relatively unimportant in explaining run production. But what is important is the general principle of trying to find the correct relative values. When we compare the standard integer weights of a single versus double, a similar analysis yields a much larger difference. Since these are the two most common hits, having correct relative weights will help us evaluate both scoring potential and, eventually, player contributions.

In our second level of analysis, we used the run expectancy matrix and game state probabilities to examine the relative value of a single and a double. Unlike a triple and homerun, a single or double does not clear all the baserunners, therefore, we need to account for the movement of baserunners. To simplify the analysis, we make the assumption that a single moves a baserunner one base, while a double moves a baserunner two bases. Of course, this assumption does not match the realities of baseball, as a single or a double can advance a runner multiple bases. We will adjust for the dynamic running of baserunners in section four.

For this analysis, we will have a beginning game state and a resulting game state and determine the number of runs scored after a single or double is hit. The resulting game state and number of runs scored for both types of hits is shown in Table 5.

Table 5: For each game state, the runs scored and resulting game state after a single, under the assumption that all baserunners advance exactly one base.

	Runs Scored: Single	Resulting Game State: Single	Runs Scored: Double	Resulting Game State: Double
000	0	100	0	020
100	0	120	0	023
020	0	103	1	020
120	0	123	1	023
003	1	100	1	020
103	1	120	1	023
023	1	103	2	020
123	1	123	2	023

The calculation of RE24 is straightforward and can be repeated with a double. Starting with a single, we find the difference between the run expectancy in the initial game state and the resulting game state and add the number of runs scored, if any. For example, the run expectancy for an initial game state of 0 outs and no runners on, denoted as "000", is 0.481. After a single is hit, which results in a game state of "100", the run expectancy shifts to 0.859. In this situation, no runners will score, so the relative value of a single will be the difference between the two, 0.378. We then make this calculation for each of the 24 game states that exist, projected in the matrix below.

Table 6: The increase in run expectancy after a single in each game state under the assumption that all baserunners advance exactly one base.

Single Gain

	Outs		
	0	**1**	**2**
000	0.378	0.255	0.126
100	0.578	0.375	0.205
020	0.684	0.466	0.159
120	0.855	0.657	0.323
003	0.509	0.559	0.871
103	0.653	0.754	0.951
023	0.820	0.754	0.898
123	1.000	1.000	1.000

Under the assumption that baserunners can only advance one base on a single, the gain from hitting a single will never exceed one, as the maximum number of runs scored is one, which is only when the initial game state has a runner on third base. Interpreting the results, the value of a single is generally higher when there are fewer outs and when a runner advances into a more likely scoring position.

The corresponding analysis for doubles is summarized in Table 7.

Table 7: The increase in run expectancy after a double in each game state under the assumption that all baserunners advance exactly two bases.

Double Gain

	Outs		
	0	1	2
000	0.619	0.410	0.221
100	1.105	0.867	0.356
020	1.000	1.000	1.000
120	1.527	1.492	1.151
003	0.750	0.714	0.966
103	1.180	1.246	1.102
023	1.136	1.288	1.739
123	1.672	1.835	1.828

The gain from a double exceeds one in most circumstances, as a double will bring in a run every time a base runner is on second or further and leave another runner in a dangerous scoring position. After calculating the relative gains for a single and double, we then normalize these values based on the probability of being in each game state, as seen in Table 8.

Tables 8a (left) and 8b (right): The gain metrics multiplied by the probability that a game is in the corresponding state. The sum of all entries is the expected gain for each type of hit.

Single Gain - Normalized

	Outs		
	0	1	2
000	0.092	0.045	0.018
100	0.034	0.026	0.015
020	0.010	0.012	0.005
120	0.012	0.016	0.010
003	0.001	0.005	0.012
103	0.003	0.008	0.015
023	0.002	0.005	0.007
123	0.004	0.009	0.011
Expected Single Gain			0.379

Double Gain - Normalized

	Outs		
	0	1	2
000	0.151	0.072	0.031
100	0.065	0.061	0.025
020	0.015	0.026	0.033
120	0.021	0.037	0.036
003	0.002	0.006	0.014
103	0.006	0.014	0.018
023	0.003	0.009	0.014
123	0.007	0.017	0.020
Expected Double Gain			0.701

As seen in Table 8, we find that the expected gain from a single and a double are 0.379 and 0.701, respectively. We again compare how these values relate to the SLG statistic. Within the slugging percentage formula used by the MLB, the relative value of a single over a double is 0.5. However, our analysis shows that relative value should be 0.541. Therefore, the SLG statistic is undervaluing the single over the double by 7.53%. These values, in the context of slugging percentage, are shown in Table 9.

Table 9: The difference in value of between a single and a double from the probability model, compared to the standard slugging percentage formula.

Model Single/Double Value	SLG Single/Double Value
0.541	0.500

Model Double/Single Value	SLG Double/Single Value
1.849	2.000

% Difference
-7.53%

This is not to say that a single is more valuable than a double, but that it may be important for managers to think about which statistics to value when assessing the quality of a hitter.

4 Derivation of appropriate weights: Variable baserunning

In this section we include the possibility that runners on base could progress more bases than the batter. It is important to note that the main simplifying assumption here is that each person on base will still progress the same number of bases. For example, we have not yet modeled the case where after a single, one runner goes from first to second base while another runner progresses from second base to home. Likewise we do not allow for the possibility that the runners on base advance fewer bases than the batter, an infrequent but not unheard of play. For example, a slow runner on second base and a speedy hitter could change our game state from 020 to 023 after a double if the runner on second has to delay leaving to see if the ball is caught.

To do this analysis, we found probabilities of runners advancing different numbers of bases on singles or doubles. These probabilities indicate that on a single, there is a 59% chance that each runner advances one base, a 40% chance that each runner on base will advance two bases and a 1% chance that each runner on base advances three bases. For doubles, the probabilities are about 75% for two bases and 25% for three bases. Note that in each case, if a runner "advances" more bases than it would take for them to get home, they will simply make it to home plate and earn their team a run. So, for example, there would be no difference in the outcome if each runner "advanced" two or three bases if the at-bat started with only a runner on second.

Using the probabilities in the paragraph above, we calculate the resulting game state for each type of hit depending on how many bases the runners advanced, as well as how many runs would be scored in each case. Table 10 lists all of the possible results.

Table 10a (top, for singles) and 10b (bottom, for doubles): The game states resulting from each type of baserunner advance.

Resulting Game State - Single

	1 Base		2 Bases		3 Bases	
	Result	Runs Scored	**Result**	Runs Scored	**Result**	Runs Scored
000	100	0	100	0	100	0
100	120	0	103	0	100	1
020	103	0	100	1	100	1
120	123	0	103	1	100	2
003	100	1	100	1	100	1
103	120	1	103	1	100	2
023	103	1	100	2	100	2
123	123	1	103	2	100	3

Resulting Game State - Double

	2 Bases		3 Bases	
	Result	Runs Scored	**Result**	Runs Scored
000	020	0	020	0
100	023	0	020	1
020	020	1	020	1
120	023	1	020	2
003	020	1	020	1
103	023	1	020	2
023	020	2	020	2
123	023	2	020	3

For example, if a batter comes to the plate with a runner on first and second, the batter hits a single, and the runners each advance two bases, the resulting game state would be "103" (runners on first and third), and one run would be scored. If the batter were to hit a double, with the runners advancing two bases, the resulting game state would be "023", and one run would be scored. From this, we took the run expectancy in the resulting game state, in addition to the number of runs scored, for each possibility given a starting game state for both singles and doubles. Then, we took a weighted average of these resulting game states based on the probabilities described above and subtracted the original run expectancy to determine the average contribution of a single and double from each game state, and we present these in Table 11:

Table 11a (left, singles) and 11b (right, doubles): Expected runs gained on each type of hit for singles and doubles for each possible game state.

Single Gain

	Outs		
	0	**1**	**2**
000	0.378	0.255	0.126
100	0.721	0.480	0.233
020	0.715	0.621	0.465
120	1.057	0.902	0.628
003	0.509	0.559	0.871
103	0.796	0.859	0.979
023	0.851	0.909	1.204
123	1.202	1.245	1.305

Double Gain

	Outs		
	0	**1**	**2**
000	0.619	0.410	0.221
100	1.139	0.939	0.541
020	1.000	1.000	1.000
120	1.561	1.564	1.336
003	0.750	0.714	0.966
103	1.214	1.318	1.287
023	1.136	1.288	1.739
123	1.706	1.907	2.013

As in the case for triples and home runs, we normalize these values based on the probability of being in each game state. Table 12 shows the expected gain in runs for each hit type:

Table 12a (left, singles) and 12b (right, doubles): Overall expected runs gained on each type of hit for singles and doubles, normalizing across probabilities of the game states.

Single Gain - Normalized

	Outs		
	0	**1**	**2**
000	0.092	0.045	0.018
100	0.043	0.034	0.017
020	0.011	0.016	0.015
120	0.015	0.023	0.019
003	0.001	0.005	0.012
103	0.004	0.009	0.016
023	0.003	0.006	0.010
123	0.005	0.011	0.014

Expected Single Gain **0.442**

Double Gain - Normalized

	Outs		
	0	**1**	**2**
000	0.151	0.072	0.031
100	0.067	0.066	0.038
020	0.015	0.026	0.033
120	0.022	0.039	0.041
003	0.002	0.006	0.014
103	0.006	0.014	0.021
023	0.003	0.009	0.014
123	0.007	0.017	0.022

Expected Double Gain **0.736**

From Table 12, we see that a single adds 0.442 expected runs to an inning, and a double adds 0.736 expected runs to an inning.

4 Conclusion

Based on our analysis of the four types of hits, we created two possible equations for wSLG; one that consolidates our findings for the analysis with fixed base movements for singles and doubles, and another that consolidates our findings for the analysis incorporating probabilities of advancing. The equations and corresponding weights are based on the relative values of the expected gains of each hit type:

Fixed baserunning: $wSLG(\mathrm{f}) = 1B + 1.849 \times 2B + 2.725 \times 3B + 3.745 \times HR$.

Variable Baserunning: $wSLG(\mathrm{v}) = 1B + 1.665 \times 2B + 2.336 \times 3B + 3.211 \times HR$.

Any team with some data about the tendencies of their baserunners could weight these two values to produce a team specific value of wSLG for their own purposes. But the key point is reinforced in either case: Doubles are not worth twice as much as singles, and extra base hits are in general not quite as much as the traditional slugging percentage gives them credit for. Because our findings indicate that the traditional SLG metric undervalues singles relative to other types of hits, using wSLG as an alternative metric will provide a more accurate measure of how many runs a batter can be expected to produce over a season. A full comparison of how our models' relative weights compare to the traditional SLG model's weights is illustrated in Table 13.

Table 13: Relative weights of hits in traditional and proposed weighted slugging averages.

Fixed Base Running

	Model	SLG	% Diff
Single/Double	0.541	0.500	*8.15%*
Single/Triple	0.367	0.333	*10.10%*
Single/HR	0.267	0.250	*6.81%*
Double/Triple	0.679	0.667	*1.80%*
Double/HR	0.494	0.500	*1.24%*
Triple/HR	0.728	0.750	*2.99%*

	Model	SLG	% Diff
Double/Single	1.849	2.000	*7.53%*
Triple/Single	2.725	3.000	*9.17%*
HR/Single	3.745	4.000	*6.38%*
Double/Triple	1.473	1.500	*1.77%*
Double/HR	2.025	2.000	*1.25%*
Triple/HR	1.374	1.333	*3.08%*

Variable Base Running

	Model	SLG	% Diff
Single/Double	0.601	0.500	*20.15%*
Single/Triple	0.428	0.333	*28.40%*
Single/HR	0.311	0.250	*24.57%*
Double/Triple	0.712	0.667	*6.87%*
Double/HR	0.518	0.500	*3.68%*
Triple/HR	0.728	0.750	*2.99%*

	Model	SLG	% Diff
Double/Single	1.665	2.000	*16.77%*
Triple/Single	2.336	3.000	*22.12%*
HR/Single	3.211	4.000	*19.72%*
Double/Triple	1.404	1.500	*6.43%*
Double/HR	1.929	2.000	*3.55%*
Triple/HR	1.374	1.333	*3.08%*

To examine how these alternative measures look when applied to actual players, we looked at the statistics for the top twenty leaders in slugging percentage for the 2023 MLB season. In Table 14 we report the traditional slugging percentage (SLG) and the reweighted slugging percentages for each play under both fixed and variable baserunning. The reweighted versions have lower values, and the version with variable base running has a larger difference than the fixed base running approach. This makes sense, as the reweighted versions are essentially transferring some credit from the batter to the baserunners. Note also that the difference (SLG - wSLG(v)) is smaller for Ronald Acuña Jr. than it is for teammate Matt Olson. This makes sense as Acuña Jr has many more singles and doubles than Olson, and those 'keep the line moving' hits are worth relatively more in this reformulation. Also observe that the relative ordering of players by SLG, wSLG(f) and wSLG(v) are almost identical. There are only three instances where the order would change, all for wSLG(v); we have highlighted those players in the table. Thus our new statistics align with existing metrics as to the relative value of the hitters, but differ slightly in estimating their worth.

Table 14: Traditional and proposed weighted slugging averages for 2023 MLB leaders in traditional SLG. Reversals of position occur for players marked **.

Name	Team	SLG	wSLG(f)	wSLG(v)
Matt Olson	ATL	0.604	0.573	0.515
**Ronald Acuña Jr.	ATL	0.596	0.569	0.523
Mookie Betts	LAD	0.579	0.551	0.502
Freddie Freeman	LAD	0.567	0.540	0.498

Juan Soto	SDP	0.519	0.495	0.451
Austin Riley	ATL	0.516	0.492	0.450
Bobby Witt Jr.	KCR	0.495	0.471	0.432
Julio Rodríguez	SEA	0.485	0.463	0.425
Marcus Semien	TEX	0.478	0.456	0.419
José Ramírez	CLE	0.475	0.453	0.418
**Kyle Schwarber	PHI	0.474	0.448	0.398
Francisco Lindor	NYM	0.470	0.448	0.409
Trea Turner	PHI	0.459	0.438	0.403
Paul Goldschmidt	STL	0.447	0.428	0.396
Alex Bregman	HOU	0.441	0.422	0.389
Ian Happ#	CHC	0.431	0.411	0.378
Nathaniel Lowe	TEX	0.414	0.397	0.369
**Eugenio Suárez	SEA	0.391	0.375	0.346
Nico Hoerner	CHC	0.383	0.371	0.353

5 Future Work

In creating a baseline model to examine the relative values of a single, double, triple, and home run, we instituted several simplifying assumptions. Natural extension of this work would result from relaxing these assumptions and doing the more sophisticated modeling required to understand these. By using a standardized run expectancy matrix, we assumed that each batter was a league-average hitter. However, some hitters can be more or less likely to move their team from one game state to another depending on their power, speed, and handedness. For example, a left-handed hitter who pulls the ball will be more likely to move a runner from first base to third base than a similarly-skilled right-handed hitter. In further considerations, it would also be prudent to examine the distribution of game states for hitters in the 3rd or 4th spot. Batters in the 3rd or 4th spot in the lineup have stronger hitters in front of them, and they are more likely to step to the plate with runners on base. In the model designed to account for the varying probabilities of how many bases a runner might advance on a single or double, our operating assumption was that all runners would advance the same number of bases. For example, on a single the probability of going from first to third is the same as going from second to home. Further studies could give a more accurate prediction of these probabilities of how many bases the runners advance given more data such as where the ball is hit. Surely there is a greater chance of going second to home on a single to the left fielder than there is of going first to third. Furthermore, the current model uses probabilities that are independent of the number of outs. With 2 outs and a full count the runners are off with the pitch; however, if there are 0 or 1 out(s), and a line drive single through the infield is hit, the runners must freeze to see the ball down. This will halt the runner's momentum and diminish their probability of advancing multiple bases. Further efforts could be made to study the implications of the fact that our analysis deems singles to be undervalued. Our analysis doesn't account for the speed of the baserunners, but an initial observation would be that the faster your team the more you are undervaluing the single. This offers the question of whether teams should then consider sacrificing some offensive production in the traditional sense for more speed on the basepaths. This is increasingly relevant in the current state of the MLB with the rule changes that went into effect for the 2023 season. The bases are larger, effectively shortening the distance between bases, and pitchers are now only allowed two disengagements from the mound per batter. Results from the 2023 season show increases in the number of stolen bases and the success rate, further indicating that the running game may be inching its way back into a more prevalent role (or a role that can be capitalized upon).

This analysis can be generalized by replacing our assumptions with variables. For a single or double, let p1 be the probability that all runners advance one base, p2 is the probability that all runners advance two bases, and p3 is the probability that all runners advance three bases. Note that for singles, p1 + p2 + p3 = 1, and for doubles, p2 + p3 = 1 and p1 = 0. The corresponding coefficients in the generalized analysis are computed below:

Table 15: Coefficients

Generalized Coefficients - Single

p1	p2	p3	Constant
0.853	**1.004**	**1.150**	**-0.474**

Generalized Coefficients - Double

p2	p3	Constant
1.175	**1.315**	**-0.474**

Using these formulas, one can compute the expected gain of a single or double taking into account the speed of their baserunners. This offers much more flexibility than the basic model, as it allows for teams to take their runners' speed into account and calculate their own wSLG. For example, teams with faster players will see relatively higher wSLG weights for singles and doubles compared to triples and home runs, which accurately reflects the fact that fast teams benefit more from singles and doubles than slow teams. Also note that for players who get most of their at-bats with slower runners on base, wSLG (or even traditional unweighted SLG) are more fair than RBIs since it is hard to drive in a slow runner from second with a single.

Acknowledgements

The authors thank Williams College and Babson College for their support of this work, and we are grateful to the editors and reviewers for thoughtful comments.

References

Baseball-Reference.com. "War Explained, Converting Runs to Wins." Accessed April 28, 2024. https://www.baseball-reference.com/about/war_explained_runs_to_wins.shtml#:~:text=Team%20PythagenPat%20W%2DL%25%20%3D%20(RS,2.0%2C%20but%20a%20little%20better.

Bukiet, Bruce, Elliotte Rusty Harold, and José Luis Palacios. "A Markov Chain Approach to Baseball." *Operations Research* 45, no. 1 (1997): 14–23. http://www.jstor.org/stable/171922.

Crashburnalley. "What's a Good Slugging Percentage and How is it Calculated?" Accessed February 22, 2025. https://crashburnalley.com/whats-a-good-slugging-percentage/.

Slowinski, Piper. "Weighted on Base Average (wOBA)." Accessed November 3, 2024. https://library.fangraphs.com/offense/woba/.

Schell, Michael J. *Baseball's All-Time Best Sluggers: Adjusted Batting Performance from Strikeouts to Home Runs*. Princeton: Princeton University Press, 2016.

Statshacker. "State-Based Event Probabilities in Baseball." Accessed May 11, 2023. https://statshacker.com/blog/2018/07/18/state-based-event-probabilities-in-baseball/.

VanDerwerken, Douglas. "Slugging Percentage Is Not a Percentage—And Why That Matters." *The American Statistician* 75, no. 2 (2021): 124–27. https://doi.org/10.1080/00031305.2018.1564698.

Wikipedia. "Slugging Percentage." Accessed May 18, 2024. https://en.wikipedia.org/wiki/Slugging_percentage#:~:text=A%20predecessor%20metric%20was%20developed,a%20predecessor%20to%20slugging%20percentage.