Why Don't We Have FIP for Hitters? A Need and Attempt to De-luck Hitter Performance

Sam M. Young*
Department of Economics
UC San Diego
smyoung@ucsd.edu

August 22, 2014

Abstract

To date, hitting sabermetrics have lagged behind pitching metrics in the removal of luck. While most sabermetricians recognize the pervasive impacts that batting average on balls in play (BABIP) can have on a hitter's batting line, attempts to precisely adjust players' "triple slash," OPS, or weighted on-base average (wOBA) for luck remain relatively uncommon. Thus, under existing metrics, luck disproportionately distorts assessment of hitters, particularly when performance differs from expectations. Given the number of hitters each year who "break out" or "decline," the ability to recognize when these incidents stem from luck versus fundamental changes in play is an essential part of player evaluation.

In this paper, I propose a method to estimate de-lucked hitter performance by estimating rates of individual hit types using underlying peripherals (e.g. fly ball percentage and fly ball distance). These rates can be used to compute expected versions of many additional statistics, as well as expected counts for each hit type. With a slight modification for home runs, expected versions of batting average, on-base percentage, slugging, isolated power, OPS, and wOBA all outperform their realized values in predicting subsequent-year performance. This aligns with existing findings for pitchers, and hints at the potential to successfully evaluate hitters on a more luck-neutral basis.

Tags: hitter luck; expected hitting; xHitting; xWOBA; xBABIP

*This paper was originally submitted to the 2014 MIT Sloan Sports Analytics Conference, and has been expanded since. I am grateful to Dan Brooks and Chuck Korb for a poster opportunity at Saber Seminar 2014, as well as Matt Goldman, Lester Lusher, Kevin Ray, Ling Shao, Alex Smith, Matt Swartz, Igor Vaynman, and various readers on FanGraphs.com for comments and feedback. Any errors are my own.

1 Introduction

Within the "sabermetric revolution," there exists an unusual asymmetry between pitching and hitting metrics in the treatment of luck. On the pitching side, many of the most popular metrics seek (as a main goal) to remove luck by estimating ERA from underlying peripherals. These estimators, such as Fielding Independent Pitching (FIP), have enjoyed a successful track record overall, and in many baseball circles are now the standard way of discussing pitching performance.[1] Yet this emphasis on removing luck has been much less pronounced on the hitter side. Even most advanced hitting metrics still rely on realized hit outcomes, which are sensitive to luck.

It is not as though sabermetricians are blind to the notion of hitter luck. To the contrary, sabermetric discussions of hitting frequently mention batting average on balls in play (BABIP) — a main avenue of hitter luck. Hitters with unusually high BABIP during a stretch are often deemed "lucky," while those with unusually low BABIP are deemed "unlucky." Yet merely recognizing *whether* a hitter has gotten lucky is not the end goal; in many cases, a related question is what level of performance would have occured under neutral luck.

On this second point, things are complicated by the fact that BABIP tends to inflate/deflate many aspects of a hitter's batting line simultaneously. That is, high BABIP raises not only batting average, but also on-base percentage, slugging, and other more modern measures such as weighted on-base average (wOBA) or its scaled version wRC+. While many people recognize this problem, attempts to precisely adjust hitter statistics for luck remain relatively uncommon. Thus, even when it is (seemingly) apparent that a hitter has gotten lucky or unlucky, the corresponding luck-neutral performance often remains a mystery. This stands in stark contrast to pitchers, where FIP and its variants provide specific estimates of de-lucked ERA.

As a result, under existing metrics luck disproportionately hinders assessment of hitters, particularly when performance differs from expectations or when there exists uncertainty about a player's ability (e.g. for rookies or players returning from injury). Given the number of players each year who "break out" or "decline," the ability to recognize when these incidents reflect luck versus fundamental changes in play is an

[1] Numerous studies have found FIP and its variants to outperform ERA in predicting future performance. Some examples include Swartz (2011) and Wyers (2009).

essential part of player evaluation. Teams that are adept in this area enjoy a significant advantage in making smart roster decisions.

In this paper, I propose a method to estimate de-lucked hitting performance by estimating rates of individual hit types among balls in play using hitting peripherals (such as fly ball percentage and fly ball distance). These rates can be used to compute expected versions of a variety of statistics, including batting average, on-base percentage, slugging, OPS, and wOBA, as well as expected totals for individual hit types. With a slight modification for home runs, expected versions of batting average, on-base percentage, slugging, isolated power, OPS, and wOBA all outperform their realized values in predicting future performance.

2 Comparison to xBABIP and Projections

To date, the most common applications of luck-neutralized hitting have been in the form of "expected BABIP" or xBABIP calculators.[2] An example by Boden (2009) is provided below.

$$\begin{aligned} xBABIP = 0.3916 + 0.2877 \cdot LD\% &- 0.1520(GB\% - (GB\% \times IFH\%)) \\ &- 0.1875(FB\% - (FB\% \times HR/FB\%) - (FB\% \times IFFB\%)) \\ &- 0.8345(IFFB\% \times FB\%) + 0.4997(IFH\% \times GB\%) \end{aligned}$$

These formulas return an expected/luck-neutralized value of BABIP based on a hitter's peripherals. Although informative, this number provides only so much value by itself. Users often wish to know what level of batting average, OPS, or wOBA corresponds to this.

But while xBABIP can be converted to expected batting average and on-base percentage with only minor assumptions, the same is not true for slugging. If a hitter is deemed to have lost x hits due to ball-in-play luck, for slugging purposes the *types* of hits lost also matter. This creates a problem when removing luck from comprehensive statistics like OPS or wOBA, using existing methods, since the xBABIP value does not itself remove luck from slugging (required for OPS) or individual hit types (required for wOBA).

Although one can proceed by assuming the types of hits gained/lost, or perhaps that BABIP leads iso-

[2]There have also been occasional attempts to estimate expected versions of other statistics; for instance, Bradbury (2005) estimates an expected OPS model. But these calculators, including xBABIP, appear to be a collective work in progress and have yet to catch on en masse to the degree of FIP. While xBABIP occasionally appears in sabermetric articles, it is not readily available on player pages at either FanGraphs or Baseball Prospectus.

lated power unaffected, either of these approaches is potentially subjective.[3] The newly proposed method instead uses an empirical approach to estimate the types of hits gained/lost, thereby offering a solution to these problems.

It is also worth clarifying that the new method is not a direct substitute for projection systems like ZiPS and Steamer. Strictly, the new method actually generates *backward*-looking estimates, rather than forward. But in much the same way that actual past performance is useful in forecasting, so too may be expected past performance.

When used as a forecaster, compared to ZiPS and Steamer, the new method may be faster to incorporate changes in playing style. If a shoulder injury causes a player to hit fewer fly balls and more line drives (e.g.) than in the past, the new method readily reacts to this change, while ZiPS and Steamer continue to put much weight on the old playing style. There is some justification to this, of course, as the style itself may gradually revert; which the new method does not directly address. So for forecasting purposes the new method likely gives too much credit for recent style changes, though its results can still be interpreted in-sample as a "de-lucker."

3 New Method: A Complete Distribution of Hit Types

The main innovation of this piece is to map a hitter's peripheral performance not just to a single xBABIP value, but to a complete distribution of outcomes among balls in play. To do so, I first compute the rate of each hit type (singles, doubles, triples, and home runs) among a player's balls in play. Regressing each rate on a corresponding set of peripherals yields predicted rates of each outcome. Multiplying these rates by the number of balls in play yields expected counts for each hit type, which can be used to compute expected versions of many additional statistics (such as OPS or wOBA).

Throughout this analysis, the player-year (e.g. 2013 Mike Trout) is the micro-unit of observation. The current sample consists of all such player-years with at least 100 plate appearances between the 2010 and 2013 MLB seasons. With the exception of fly ball distance, which comes from BaseballHeatMaps.com, all other data come from FanGraphs.com.

[3]Gross (2011) discusses possible assumptions to use when converting xBABIP to slugging in more detail.

To be concrete, below I walk through the process for the case of singles. For other outcomes, the process itself is similar, but the explanatory peripherals used differ slightly.[4]

The rate of singles is computed as:

$$\textit{Singles Rate}_{it} = \frac{\mathit{1B}_{it}}{(AB - K + SF)_{it}} \tag{1}$$

where $\mathit{1B}_{it}$ denotes the number of singles hit in year t by player i, and AB, K, and SF denote at bats, strikeouts, and sacrifice flies respectively. The denominator denotes the number of balls in play, which differs here from other BABIP contexts by also including home run balls as "in play" (since home run rate will also be estimated).

The next step is to estimate the relationship between singles rate and explanatory peripherals, which in turn yields a predicted rate of singles. That is, I use the data to estimate:[5]

$$\begin{aligned}\textit{Singles Rate}_{it} = {} & \beta_0 + \beta_1 GB\%_{it} + \beta_2 FB\%_{it} + \beta_3 IFFB\%_{it} + \beta_4 FB\ Dist_{it} + \beta_5 LF\%_{it} \\ & + \beta_6 RF\%_{it} + \beta_7 (LF\%_{it} \times LHB_{it}) + \beta_8 (LF\%_{it} \times Switch_{it}) \\ & + \beta_9 (RF\%_{it} \times LHB_{it}) + \beta_{10} (RF\%_{it} \times Switch_{it}) + \beta_{11} Spd_{it} \\ & + \beta_{12} (Spd_{it} \times GB\%_{it}) + \beta_{13} Park\ Factor_{it} + \epsilon_{it}\end{aligned} \tag{2}$$

and the predicted rate of singles, denoted $\widehat{\textit{Singles Rate}}_{it}$, is simply the value of the right-hand side using the estimated values of $\beta_0, \beta_1, ..., \beta_{13}$.

The predicted *number* of singles is $\widehat{\textit{Singles Rate}}_{it}$ multiplied by the number of balls in play.

$$\widehat{\mathit{1B}}_{it} = \widehat{\textit{Singles Rate}}_{it} \cdot (AB - K + SF)_{it} \tag{3}$$

Predicted values for batting average, on-base percentage, slugging, etc., are computed using their standard formulas, but with predicted numbers of singles, doubles, triples, and home runs in lieu of their actual values. As mentioned earlier, these are all technically *backward*-looking estimates of how a hitter would have performed during the same stretch under neutral luck.

But while officially backward-looking, testing whether these estimates achieve their overall purpose can-

[4]The peripherals for each hit type are allowed to vary since the intuitive determinants of each hit type also differ. There is, however, a significant amount of overlap. The exact peripherals used for each outcome can be found in Table A1 of the Appendix.

[5]Since these abbreviations may be confusing, a glossary of variables is provided in the Appendix.

not be done on a same-time-period basis, where realized performance necessarily correlates perfectly with itself. Instead, a standard exercise in these settings is to compare whether actual or predicted performance correlates better with *next season's* performance.[6]

The rationale is that, whether a player was lucky or unlucky in the past, future luck should remain neutral, and thus there is no reason for future performance to deviate systematically from true talent. So if expected wOBA (e.g.) makes a better predictor of subsequent-year wOBA, it suggests that expected wOBA may be more faithful to a player's true hitting ability, by removing luck that occurs in finite samples.

4 Results

For brevity, the text of this paper skips over the estimated relationships between rates of each hit type and their peripherals, and proceeds straight to forecasting. The estimated relationships, however, can be found in Table A1 of the Appendix.

Table 1 below displays correlations between a season's rate of singles, doubles, triples, and home runs (among balls in play) and the preceding season's actual versus predicted rates. In each case, the predicted rate exhibits the stronger correlation.

Table 1: Correlations with Subsequent Season Rates

	Singles rate	Doubles rate	Triples rate	HR rate
Actual rate	0.227	0.167	0.367	0.523
Predicted rate	0.346	0.235	0.416	0.530

Sample: 100+ plate appearances in "prior" season

This same pattern holds for batting average and on-base percentage, but somewhat surprisingly not for slugging, isolated power (ISO), OPS, and wOBA under the initial specification.

Table 2: Correlations with Subsequent Season Values

	AVG	OBP	SLG	ISO	OPS	wOBA
Actual value	0.364	0.413	0.444	0.548	0.432	0.421
Predicted value	0.399	0.433	0.416	0.516	0.421	0.412

Sample: 100+ plate appearances in "prior" season

The discrepancy between Tables 1 and 2 appears to be due to home runs. Although $\widehat{Home\ Run\ Rate}_{it}$

[6]This method has been used to assess ERA estimators, like FIP and SIERA.

appears (in Table 1) to be mildly better than unadjusted *Home Run Rate*$_{it}$ in predicting future home run hitting, this pattern is primarily true among players with not-that-many plate appearances. As the number of plate appearances increases, actual home run rate becomes the better standalone forecaster.[7] Thus, there is a reasonable argument to maintain the actual number of home runs. Doing so — but using predicted numbers of singles, doubles, and triples — expected OPS and wOBA also outperform their realized values in predicting subsequent season performance.

Yet predicted home run rate does not appear to be useless. When home run rate is regressed on lagged home run rate and its *residual*, results (in Table A3 of the Appendix) indicate regression toward the predicted rate;[8] and this is true even among players with many plate appearances. So predicted home run rate ultimately seems to convey additional information, and it may be best to use a *combination* of actual and predicted home runs when predicting future performance. This appears true in the data, as a 50/50 average of actual and predicted home runs does better than using either option alone when forecasting any of AVG, OBP, SLG, ISO, OPS, or wOBA; and this pattern becomes even more true when allowing the weight on actual home runs to vary with the number of plate appearances.[9]

Table 3 below shows an augmented version of Table 2 with the newly discussed alternatives. For each outcome, the weighted average between actual and predicted home runs (and using regular predicted values for singles, doubles, and triples) appears most predictive of future performance. And though the exact numbers in Table 3 change if one increases the number of plate appearances, the overall pattern remains the same.

In this sense, it appears possible to extend the principles of "de-lucking" to various aspects of hitting performance, with a degree of success. Yet it remains a natural question how literally the model should be taken. Almost every model is sure to omit *some* explanatory variables — in part because certain skills are difficult to measure — and as the treatment of home runs shows, there do appear be other skills involved that are not captured by the current specification. So when a hitter's wOBA (e.g.) differs from its expected

[7] Interestingly, this statement only appears true for home runs, and not for singles, doubles, and triples. See Table A2 in the Appendix for more detail.

[8] This statement also applies for singles, doubles, and triples. The residual is defined as the actual rate minus the predicted rate. Negative coefficients on residual terms indicate regression.

[9] See the Appendix for how the seemingly optimal weight is determined.

Table 3: Correlations with Subsequent Season Values

	AVG	OBP	SLG	ISO	OPS	wOBA
Actual value	0.364	0.413	0.444	0.548	0.432	0.421
Predicted value, raw	0.399	0.433	0.416	0.516	0.421	0.412
Predicted value, no HR replacement	0.403	0.441	0.440	0.539	0.440	0.431
Predicted value, 50/50 HR weighting	0.409	0.442	0.451	0.554	0.448	0.437
Predicted value, PA-based HR weighting (linear)	0.411	0.443	0.458	0.561	0.453	0.442

Sample: 100+ plate appearances in "prior" season

version, how much of this may actually be skill?

To provide a partial answer, here I estimate the persistency of a player's wOBA residual from one year to the next. On one extreme, if these deviations stem exclusively from luck, there should be zero correlation in the residuals across years, while on the other extreme, if the deviations are entirely skill, there should be a high correlation from year to year.

Table 4: Persistency of wOBA Residual

	PA in prior season			
	All (100+)	100–299	300–499	500+
Lagged wOBA residual	0.082**	0.041	0.132**	0.111**
	(0.032)	(0.050)	(0.061)	(0.050)
Constant	-0.001	-0.002	-0.001	-0.001
	(0.001)	(0.002)	(0.001)	(0.001)
Observations	1,218	343	348	527
R-squared	0.007	0.002	0.016	0.008

Notes: The outcome variable is wOBA residual (actual−expected wOBA), using the weighted average for home runs and regular predicted values for singles, doubles, and triples. Robust standard errors, clustered by player, are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$

As can be seen in Table 4, there does appear to be some skill involved, as one season's wOBA residual is a significant predictor of the following season's. Yet the magnitude of persistence is fairly small. No matter the range of plate appearances considered, at most 10–15% appears to carry over systematically. So even when a player outperforms his expected wOBA by 0.030 ($\approx$ 0.075 OPS) over a relatively full season, on average this only predicts outperformance of 0.004 the next year. Similar patterns exist for other outcomes, such as OBP and SLG.

So even with a degree of skill involved, in any *given* season quite a large share of the difference between actual and expected performance appears due to randomness that players are unable to control (or at least sustain) from year to year. Though this does not rule out systematic over- and underachievers, it means that even a known "overachiever" may do so by different amounts from year to year, and even unerachieve in some years.[10] These patterns are similar to what has been observed for FIP, suggesting broadly similar strengths as well as weaknesses.

Finally, I highlight some individual players' performance from the 2013 season, covering two main uses of the model: (1) seeing which players luck seems to have affected most; and (2) checking whether players' (apparent) breakouts or declines seem justified, based on peripheral performance.

Table 5 shows the ten leading over- and underachievers of the 2013 MLB season (by wOBA, using the weighted average for home runs). Several of these players are still in their early years in the Major Leagues, with consensus on true hitting abilities yet to fully solidify. For these players especially, it will be interesting to see whether future seasons reveal their 2013 performance to be somewhat illusory.

Table 5: 2013 Over- and Underachievers

Name	wOBA	xWOBA	Difference	Name	wOBA	xWOBA	Difference
Jose Iglesias	0.327	0.259	0.068	Kevin Frandsen	0.286	0.336	-0.050
Yasiel Puig	0.398	0.341	0.057	Alcides Escobar	0.247	0.294	-0.047
Ryan Braun	0.370	0.318	0.052	Todd Helton	0.322	0.367	-0.045
Ryan Raburn	0.389	0.338	0.051	Edwin Encarnacion	0.388	0.433	-0.045
Colby Rasmus	0.365	0.317	0.048	Ryan Hanigan	0.252	0.297	-0.045
Matt Adams	0.365	0.319	0.046	Darwin Barney	0.281	0.318	-0.037
Junior Lake	0.335	0.290	0.045	Josh Rutledge	0.281	0.318	-0.037
Justin Maxwell	0.336	0.292	0.044	Yuniesky Betancourt	0.257	0.294	-0.037
Chris Johnson	0.354	0.313	0.041	Brian Roberts	0.309	0.346	-0.037
Yan Gomes	0.359	0.319	0.040	Jordan Pacheco	0.262	0.299	-0.037

Notes: Minimum 250 plate appearances in 2013. xWOBA denotes "expected wOBA," and is computed here using the weighted average for home runs, and regular predicted values for singles, doubles, and triples. 0.020 wOBA $\approx$ 0.050 OPS, on the margin.

Table 6 meanwhile shows players whose performance in 2013 differed most from baseline recent performance in 2010–2012. Some players' breakouts appear very well-supported by peripherals (e.g. Gomez

[10] Readers curious about systematic over- and underachievers can find a partial list in the Appendix.

and Lind), as do some players' declines (e.g. Upton and Bernadina). In other cases, it appears luck may have played some role.[11] As with the previous list, this should also become more clear with future seasons.

Table 6: 2013 Gainers and Decliners

Name	wOBA	xWOBA	2010-12 wOBA	Name	wOBA	xWOBA	2010–12 wOBA
Hanley Ramirez	0.442	0.414	0.341	Paul Konerko	0.298	0.326	0.391
Chris Davis	0.421	0.410	0.327	Michael Morse	0.286	0.302	0.371
Ryan Raburn	0.389	0.338	0.310	Josh Hamilton	0.319	0.311	0.400
Michael Cuddyer	0.396	0.363	0.341	B.J. Upton	0.252	0.253	0.328
Adam Lind	0.368	0.359	0.313	Jeff Francoeur	0.235	0.257	0.311
Carlos Gomez	0.363	0.355	0.309	Ryan Hanigan	0.252	0.297	0.327
Marlon Byrd	0.364	0.334	0.318	Jason Kubel	0.267	0.286	0.338
Juan Uribe	0.334	0.311	0.289	Roger Bernadina	0.245	0.247	0.311
Jayson Werth	0.403	0.389	0.359	Lance Berkman	0.313	0.290	0.378
Freddie Freeman	0.387	0.370	0.343	Carlos Ruiz	0.303	0.287	0.365

Notes: Minimum 250 plate appearances in 2013, and 750 plate appearances in 2010–12. xWOBA denotes "expected wOBA," and is computed here using the weighted average for home runs, and regular predicted values for singles, doubles, and triples. 0.020 wOBA $\approx$ 0.050 OPS, on the margin.

5 Discussion

This paper proposes a new method to estimate luck-neutralized hitter performance, by using hitting peripherals to estimates rates of individual hit types among balls in play. These rates can be used to compute expected versions of a wide variety of statistics, as well as expected totals for individual hit types. A major extension over existing xBABIP-based approaches is that this method offers an empirical basis to recover expected slugging, isolated power, and wOBA.

For sure, the exact specification is far from perfect, and can certainly be improved. Yet even in its current state, with a slight modification for home runs, expected versions of batting average, on-base percentage, slugging, isolated power, OPS, and wOBA all outperform their realized values in predicting subsequent-year performance. This aligns with findings from pitcher metrics, and hints at the potential to similarly evaluate hitters on a more luck-neutral basis. As is, under the status quo luck disproportionately hinders assessment of hitters; but it seems this need not be the case.

Extensions in the future may wish to study partial-season performance, and whether projection sys-

[11] These statements also are not mutually exclusive. For instance, Raburn seems to have improved even by expected wOBA, but may have also enjoyed some good luck. The opposite is true of Hanigan.

tems (like ZiPS and Steamer) can be made more accurate by incorporating both actual and expected past performance.

References

1. Boden, Robert. 2009. "A new xBABIP calculator." *Beyond the Box Score*. Accessed December 20, 2013, <http://www.beyondtheboxscore.com/2009/11/10/1124792/a-new-xbabip-calculator>.

2. Bradbury, JC. 2005. "Introducing PrOPS." *The Hardball Times*. Accessed December 20, 2013, <http://www.hardballtimes.com/main/article/introducing-props/>.

3. Gross, Jeffrey. 2011. "Looking ahead: 2011 xBABIP-adjusted batting lines." *The Hardball Times*. Accessed December 20, 2013, <http://www.hardballtimes.com/main/fantasy/article/looking-ahead-2011-xbabip-adjusted-batting-lines/>.

4. Swartz, Matt. 2011. "New SIERA, Part Four (of Five): Testing." *FanGraphs*. Accessed December 20, 2013, <http://www.fangraphs.com/blogs/new-siera-part-four-of-five-testing/>.

5. Wyers, Colin. 2009. "How well can we predict ERA?" *The Hardball Times*. Accessed December 20, 2013, <http://www.hardballtimes.com/main/article/how-well-can-we-predict-era/>.

Glossary of Regression Variables

$FB\ Dist$ = average fly ball distance in feet
$FB\%$ = fly ball percentage among balls in play (0.5 = 50 percent)
$GB\%$ = ground ball percentage among balls in play
$IFFB\%$ = infield fly ball percentage among balls in play
$LD\%$ = line drive percentage among balls in play
$LF\%$ = percentage of balls in play hit to left field
LHB = indicator for left-handed batter
$Opp\%$ = percentage of balls in play hit to the "opposite field" direction
$Park\ Factor$ = outcome-specific park factor (100 = league average)
$Pull\%$ = percentage of balls in play hit to the "pull" direction
$RF\%$ = percentage of balls in play hit to right field
Spd = speed score
$Switch$ = indicator for switch hitter

Table A1: Estimated Relationships between Peripherals and Rates of Individual Hit Types

	(1) Singles rate	(2) Doubles rate	(3) Triples rate		(4) HR rate
GB%	-0.318***	-0.120***	0.015**	GB%	-0.034***
	(0.030)	(0.014)	(0.007)		(0.010)
FB%	-0.419***	-0.094***	0.021***	FB%	-0.531***
	(0.024)	(0.018)	(0.007)		(0.098)
IFFB%	-0.064***	-0.033***	-0.005	IFFB%	-0.027***
	(0.016)	(0.010)	(0.003)		(0.008)
FB dist/100	0.000	0.046***	0.002**	FB dist/100	-0.404***
	(0.005)	(0.003)	(0.001)		(0.063)
LF%	-0.081***	0.050***	-0.008**	Pull%	-0.286**
	(0.019)	(0.013)	(0.003)		(0.119)
RF%	-0.056**	0.057***	-0.001	Opp%	0.006
	(0.025)	(0.017)	(0.004)		(0.012)
LF%×LHB	0.060***	0.032**	0.000	(FB dist/100)×FB%	0.176***
	(0.019)	(0.014)	(0.004)		(0.037)
LF%×Switch	0.011	-0.008	-0.001	(FB dist/100)×Pull%	0.091**
	(0.028)	(0.017)	(0.006)		(0.045)
RF%×LHB	-0.067***	-0.024**	0.001	Pull%×FB%	0.284***
	(0.018)	(0.012)	(0.003)		(0.080)
RF%×Switch	-0.030	0.006	0.001	(FB dist/100)^2	0.072***
	(0.027)	(0.017)	(0.006)		(0.013)
Spd/10	-0.057***	-0.027**	-0.000		
	(0.021)	(0.011)	(0.005)		
(Spd/10)×GB%	0.161***				
	(0.046)				
(Spd/10)×FB%		0.077**			
		(0.031)			
(Spd/10)×LD%			0.066***		
			(0.023)		
(Spd/10)^2			0.021***		
			(0.003)		
Park factor/100	0.070*	0.037***	0.005***	Park factor/100	0.027***
	(0.037)	(0.011)	(0.001)		(0.005)
Constant	0.484***	-0.046**	-0.023***	Constant	0.578***
	(0.045)	(0.020)	(0.007)		(0.087)
Observations	1,806	1,806	1,806	Observations	1,806
R-squared	0.396	0.254	0.543	R-squared	0.671

Notes: The estimating sample includes all player-year observations between the 2010 and 2013 MLB seasons with at least 100 plate appearances. All rates are coded so that 50% equals 0.50. 'Park factor' refers to the outcome-specific park factor as listed on FanGraphs.com. Line drive percentage and center field percentage are omitted to avoid collinearity. Robust standard errors, clustered by player, are in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table A2: Correlations with Subsequent Season Rates (Various PA Ranges)

100+ PA (All)	Singles rate	Doubles rate	Triples rate	HR rate
Actual rate	0.227	0.167	0.367	0.523
Predicted rate	0.346	0.235	0.416	0.530

100–299 PA	Singles rate	Doubles rate	Triples rate	HR rate
Actual rate	0.109	0.114	0.275	0.332
Predicted rate	0.234	0.130	0.300	0.418

300–499 PA	Singles rate	Doubles rate	Triples rate	HR rate
Actual rate	0.236	0.149	0.320	0.577
Predicted rate	0.380	0.264	0.433	0.546

500+ PA	Singles rate	Doubles rate	Triples rate	HR rate
Actual rate	0.432	0.312	0.562	0.752
Predicted rate	0.522	0.413	0.563	0.678

Notes: Number of plate appearances is for the “prior” season.

Table A3: Regression of Rates by Hit Type on Previous Season’s Rate and Residual

	(1) Singles rate	(2) Doubles rate	(3) Triples rate	(4) HR rate
Lagged rate	0.721***	0.667***	0.543***	0.765***
	(0.068)	(0.066)	(0.041)	(0.034)
Lagged rate residual	-0.706***	-0.572***	-0.417***	-0.444***
	(0.089)	(0.082)	(0.061)	(0.110)
Constant	0.057***	0.019***	0.002***	0.007***
	(0.014)	(0.004)	(0.000)	(0.001)
Observations	1,196	1,196	1,196	1,196
R-squared	0.120	0.058	0.181	0.306

Notes: The estimating sample consists of all player-year observations between the 2011 and 2013 MLB seasons with at least 100 plate appearances in the preceding season. Qualitative patterns remain similar even when increasing the number of plate appearances. The residual is defined as the actual rate minus the predicted rate. Robust standard errors, clustered by player, are in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Weighting Actual versus Predicted Home Runs

The final specification in the text uses a weighted average between actual and predicted home run rate. As can be seen in Table A3, predicted home run rate appears to have its greatest usefulness (for forecasting) when the number of plate appearances is low, before gradually declining as the number of plate appearances increases. In this section, I describe an empirical approach to find the "best" weight to put on actual versus predicted home run rate, for any given number of plate appearances, for forecasting purposes.

To find the appropriate weights, I first divide the data in many subsamples based on the number of plate appearances in the preceding season: 100–149 PA, 150–199, 200–249, etc. For each range, I consider 11 possible weightings between actual and predicted home run rate, and find which has the lowest root mean squared error (RMSE) in an AR(1) regression of home run rate.[12] An example, for 100–149 plate appearances, is shown below.

Table A4: Home Run Weighting

Weight on actual HR rate	Weight on predicted HR rate	RMSE in AR(1) regression of HR rate
1	0	0.03040
0.9	0.1	0.03028
0.8	0.2	0.03015
0.7	0.3	0.03003
0.6	0.4	0.02991
0.5	0.5	0.02981
0.4	0.6	0.02973
0.3	0.7	0.02967
0.2	0.8	0.02966
0.1	0.9	0.02968
0	1	0.02974

In this case, the lowest RMSE occurs when 20% weight is put on actual home run rate, and 80% weight on predicted home run rate. Repeating this exercise for all plate appearances ranges, I obtain the following scatterplot:

[12] That is, I regress subsequent-season home run rate on each weighting combination.

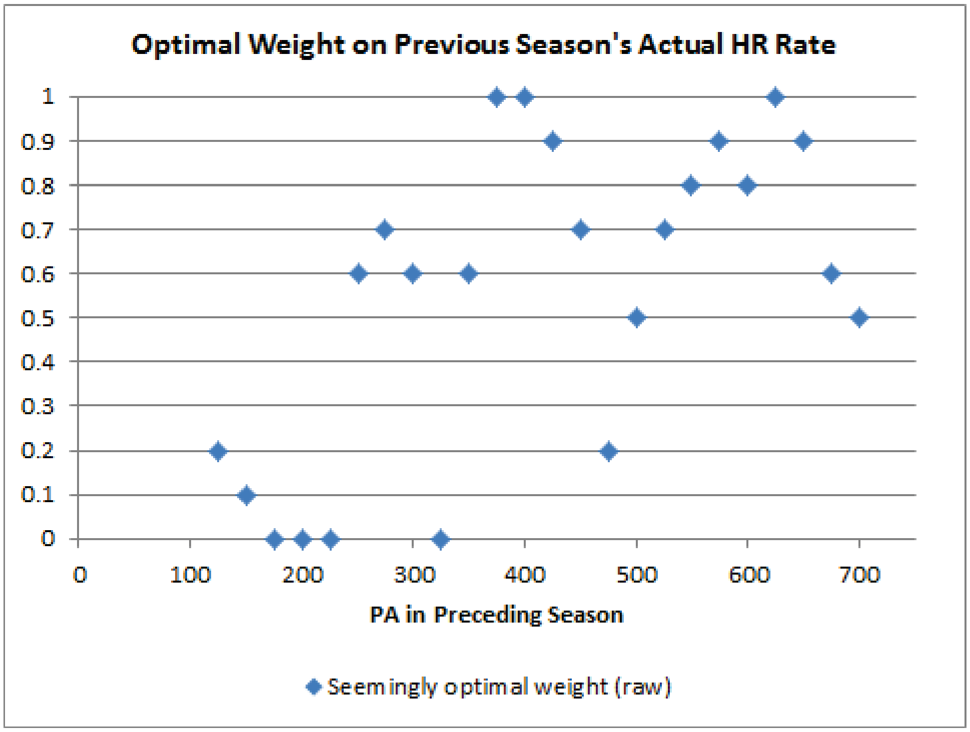


Figure 1

As expected, the optimal weight on actual home run rate generally increases with the number of plate appearances. Fitting a line through these raw data points, the optimal weight on actual home run rate is approximately:

$$Optimal\ Weight = 0.0527102 + 0.0012157 * PA \tag{4}$$

This is about 17% at 100 plate appearances, and 78% at 600 plate appearances, with the 50% mark coming at about 368 plate appearances.

Distribution of wOBA Residuals, Including Player Fixed Effects

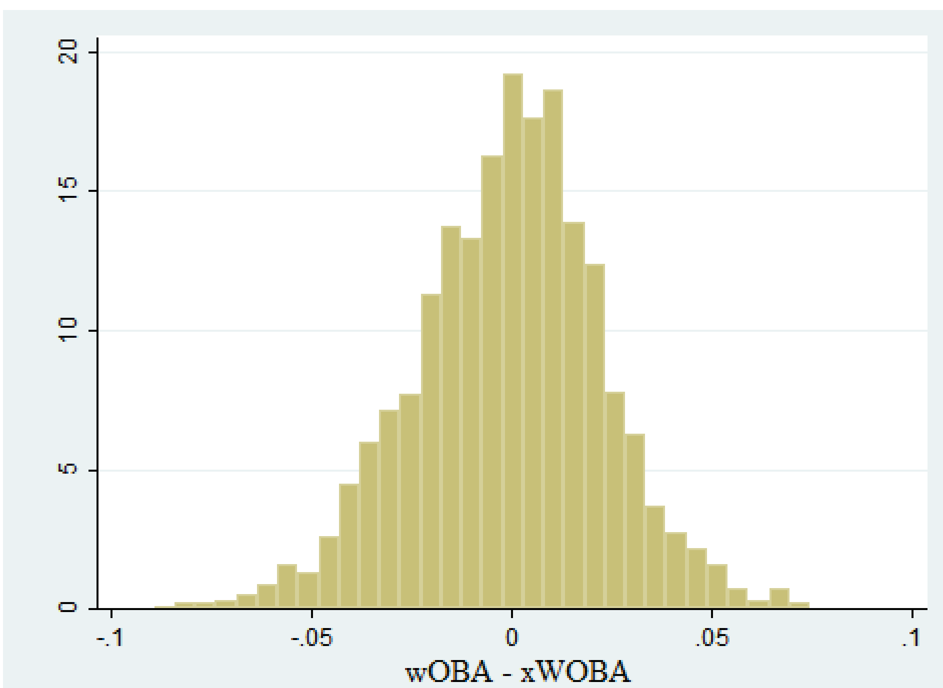
On the player-season level, the distribution of wOBA residuals (actual− expected wOBA) appears to be roughly normally distributed with mean 0. The standard deviation, meanwhile, depends on the number of plate appearances, and typically decreases with additional plate appearances.

Table A5 below shows basic summary statistics for these residuals, while Figure 2 shows a histogram on the player-season level.

Table A5: Summary Statistics for wOBA Residuals (Various PA Ranges)

	N	Mean	Std. Dev.	Share with $\lvert$Resid$\rvert \geq 0.040$
All (100+ PA)	1806	-0.001	0.024	0.099
100–299 PA	730	-0.006	0.029	0.199
300–499 PA	490	0.002	0.021	0.045
500+ PA	586	0.003	0.016	0.019

Notes: The full sample consists of player-seasons between 2010 and 2013 with at least 100 plate appearances. Number of plate appearances refers to an individual season; thus N counts player-seasons. 0.040 wOBA $\approx$ 0.100 OPS, on the margin.



Sample: 100+ plate appearances in given season

Figure 2

Table A6, meanwhile, shows the players with the largest difference between their actual and expected wOBA across the entire 2010–2013 time period. Thus there exists some evidence that these players are systematic over- and underachievers of the model. To the extent this is still a somewhat finite sample (in some cases just over 1000 plate appearances), one may wish to take these "player fixed effect" estimates with a grain of salt. But they should give a starting sense of which players have shown some ability to over/underperform the model, and by approximately how much.

Table A6: 2010–2013 Over- and Underachievers

Name	wOBA	xWOBA	Difference	Name	wOBA	xWOBA	Difference
Matt Carpenter*	0.373	0.344	0.029	Todd Helton	0.339	0.365	-0.026
Brandon Belt	0.347	0.322	0.025	Russell Martin	0.316	0.340	-0.024
Chris Johnson	0.334	0.310	0.024	Darwin Barney*	0.278	0.301	-0.023
Carlos Ruiz	0.352	0.328	0.024	Coco Crisp	0.326	0.346	-0.020
Melky Cabrera	0.336	0.312	0.024	Alcides Escobar	0.279	0.298	-0.019
Allen Craig	0.367	0.345	0.022	Jimmy Rollins	0.313	0.332	-0.019
Mike Trout	0.405	0.384	0.021	Edwin Encarnacion	0.371	0.390	-0.019
Brett Gardner*	0.334	0.314	0.020	Anthony Rizzo	0.323	0.341	-0.018
Matt Holliday	0.388	0.368	0.020	Yuniesky Betancourt	0.282	0.300	-0.018
Giancarlo Stanton	0.379	0.359	0.020	Eric Young	0.300	0.316	-0.016

Notes: Minimum 250 plate appearances in 2013 and 1000 total from 2010–2013. xWOBA denotes "expected wOBA," and is computed here using the weighted average for home runs, and regular predicted values for singles, doubles, and triples. Unless marked with (*), wOBA and xWOBA figures are for the entire 2010–2013 time period. For players marked with (*), these figures omit seasons with fewer than 100 plate appearances, where xWOBA has not been computed. 0.020 wOBA $\approx$ 0.050 OPS, on the margin.