
1-1-2012

Parsing the Relationship Between Baserunning and Batting Abilities Within Lineups

Ben S. Baumer

The Doctorate-Granting Institution of the City University of New York, bbaumer@smith.edu

James Piette

University of Pennsylvania

Brad Null

Stanford University

Follow this and additional works at: https://scholarworks.smith.edu/sds_facpubs



Part of the [Data Science Commons](#), [Other Computer Sciences Commons](#), and the [Statistics and Probability Commons](#)

Recommended Citation

Baumer, Ben S.; Piette, James; and Null, Brad, "Parsing the Relationship Between Baserunning and Batting Abilities Within Lineups" (2012). Statistical and Data Sciences: Faculty Publications, Smith College, Northampton, MA.

https://scholarworks.smith.edu/sds_facpubs/40

This Article has been accepted for inclusion in Statistical and Data Sciences: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu

*Journal of Quantitative Analysis in
Sports*

Manuscript 1429

Parsing the Relationship between Baserunning
and Batting Abilities within Lineups

Ben S. Baumer, *CUNY Graduate School and University
Center*

James Piette, *University of Pennsylvania*

Brad Null, *Stanford University*

Parsing the Relationship between Baserunning and Batting Abilities within Lineups

Ben S. Baumer, James Piette, and Brad Null

Abstract

A baseball team's offensive prowess is a function of two types of abilities: batting and baserunning. While each has been studied extensively in isolation, the effects of their interaction is not well understood. We model offensive output as a scalar function f of an individual player's batting and baserunning profile z . Each of these profiles is in turn estimated from Retrosheet data using hierarchical Bayesian models. We then use the SimulOutCome simulation engine as a method to generate values of $f(z)$ over a fine grid of points. Finally, for each of several methods of taking the extra base, we graphically depict the surface $f(z)$ over changes in the probability of advancing via that method. This framework allows us to draw conclusions both about optimal baserunning strategies in general, and about how particular offensive profiles affect a player's optimal baserunning strategy. We present many informative visualizations and analyze specific aspects of several well-known Major League players.

KEYWORDS: baseball, baserunning, Bayesian modeling, simulation, data visualization

Author Notes: We are grateful to several reviewers for constructive criticisms that we feel ultimately strengthened our paper.

1 Introduction

“Big power hitters swing and miss and strikeout. Or they hit home runs and walk. And at the end of the year their OBP is always going to be higher than most of the other guys on the team because they clog the bases.” – Harold Reynolds (2009)

It has long been debated among baseball fans and pundits whether it is advisable for teams to “manufacture runs” by aggressively stealing bases and attempting to advance baserunners at every opportunity or to “clog the bases,” as Harold Reynolds puts it, and wait for the three-run homer. Certainly different teams have taken different approaches (e.g. the speedy Cardinals of the mid 1980’s versus the slugging A’s of the early 2000’s), and to some extent, it is obvious that particular players are better suited to one strategy as opposed to another (e.g. Vince Coleman versus Jason Giambi).

Yet clearly, any optimal baserunning strategy depends not only on the situation (i.e. number of outs, score, etc.), but also in a meaningful way upon the composition of batting skills present in the lineup. In this paper, we develop a framework for exploring how baserunning strategies are affected by batting abilities. Our simulation-based approach enables one to estimate the magnitude of the change in run scoring caused by a change in baserunning strategy/ability, given a fixed batting profile.

For the purposes of this paper, we must define three important terms and how they relate to a player’s abilities. A *batting profile* is a vector that gives the probability of observing each member of a pre-defined set of outcomes for a plate appearance. A player’s batting profile will be used to represent their batting ability. A *baserunning profile* is made up of two vectors that are used to describe a player’s baserunning ability. The first vector indicates the probability of attempting each of a number of pre-defined baserunning events, while the second vector is the probability of successfully executing each those baserunning events. When combined together for a specific player, these two profiles represent a player’s *offensive profile*.

1.1 Previous Work

Several researchers have addressed baserunning in the past. Tango, Lichtman, and Dolphin (2007) attempted to assess circumstances under which it was optimal to steal and along with Click (2007), estimated the break even probability below which a team, in a particular situation, should never attempt to steal a base. Complementarily, Click (2005) and Fox (2005) estimated the number of runs gained or lost by

an individual baserunner by assigning a run value to each baserunning event and summing the events for each player from play-by-play data. Observing that batting outcomes have an effect on baserunning events due to factors related to the batted ball in play (i.e. flyball vs. groundball, etc.), Baumer (2009) constructed a simulator that would generate hypothetical play-by-play data based on both batting and baserunning abilities. This analysis was performed using both empirical and Bayesian approaches to estimate the values related to baserunning ability for each player (Baumer (2009), Baumer and Terlecky (2010)).

Null (2009) generated accurate predictions of batting outcomes using a nested Dirichlet model. In this paper, we incorporate this model to compute not only estimates for batting profiles, but also for baserunning profiles. There are other existing Bayesian treatments of modeling player ability, most notably Albert (2006), who chose to place priors of nested Beta distributions on each individual's batting outcomes. While this approach has its own advantages, the nested Dirichlet model used by Null involves a form of dynamic nesting, rather than a fixed nesting pattern. Given the degree of granularity within the simulator, the dynamic approach is favored.

1.2 Purpose

The goal of this paper is to explore the fundamental relationship between baserunning ability and batting skill and obtain a meaningful assessment of their interaction. We pose three important questions related to this relationship that we aim to address:

- Given an initial batting and baserunning profile, which baserunning method(s) increases run scoring the fastest/slowest (i.e. what is the magnitude and direction of the steepest change in run scoring among the different methods of taking the extra base)? Specifically, if my team consists of 9 copies of Derek Jeter, what changes in his baserunning ability or strategy will help that team the most?
- Is it the case that the answer to the previous question differs based on the initial offensive profile? For example, would I get a different answer if my team consisted of 9 copies of Russell Branyan, as opposed to Jeter?
- Can we tailor an optimal¹ baserunning strategy to a particular style of player? And how does this strategy differ among players with dramatically different batting abilities? For example, how does the optimal set of baserunning

¹Optimal here is defined as maximizing the expected number of runs scored per nine innings.

strategies for teams built like Adam Dunn differ from those built like Jose Reyes?

1.3 Organization

The layout of this paper is as follows. In section 2, we describe the parameterization of the model, the framework needed to explain the baserunning-batting relationship and the method of estimation. Section 3 goes on to explain how this framework is implemented and tested. We present the results in section 4, and we discuss our findings and possible future research directions in section 5.

2 Model

2.1 Parameter Spaces

Three separate parameter spaces exist for each Major League Baseball player within this model, one for a player's batting profile and two for their baserunning profile.

Batting Profile Let $E = \{1B, 2B, 3B, HR, BB, SO, HBP, GO, AO\}$ be a set of k mutually exclusive outcomes of a plate appearance². A batting profile for player j is a vector \mathbf{p}_j that describes the true, unknown probability that any of j 's plate appearances will end with each of the outcomes in E . We denote the number of times that outcome i for player j was observed in n plate appearances as X_{ij} , where $i \in \{1, \dots, k\}$. Then, the vector $\mathbf{X}_j = (X_{1j}, \dots, X_{kj})$ is distributed by

$$\mathbf{X}_j \sim \text{Multinomial}(n, \mathbf{p}_j),$$

where \mathbf{p}_j is a vector of scalars (p_{1j}, \dots, p_{kj}) , p_{ij} is the true probability of observing outcome i from player j , and $\sum_{i=1}^k p_{ij} = 1$. One example of a batting profile is provided by Prince Fielder who, in 2094 plate appearances between 2007 and 2009, recorded the following: 267 singles, 100 doubles, 7 triples, 130 home runs, 284 walks, 393 strikeouts, 35 hit by pitches, 424 ground outs, and 454 air outs. Our empirical estimate of his batting profile, $\hat{\mathbf{p}}_j$, is shown in Table 1. Later, we will improve upon this estimate by incorporating it into a more sophisticated model.

Baserunning Profile It is necessary to have two parameter spaces for a baserunning profile, one for the probability of an attempt of a given baserunning outcome

²For details on these notations, please see Appendix A.

	1B	2B	3B	HR	BB	SO	HBP	GO	AO
$\bar{\mathbf{p}}_j$	0.128	0.048	0.003	0.062	0.136	0.188	0.017	0.202	0.217

Table 1: Empirical Estimate of Batting Profile for Prince Fielder (= j).

and one for the probability of a success on that attempt. The specifications for baserunning attempt and success rates are slightly different from batting outcomes in that we are dealing with independent probabilities of disjoint events. We define ℓ methods of “taking the extra base,” (e.g. stealing 2nd base, or advancing from 1st to 3rd on a single) or baserunning events, as in Baumer (2009)³. Let N_{ij} be the number of attempts of baserunning event i by player j and similarly, let Y_{ij} be the number of successes of baserunning event i by player j . Then the number of attempts by player j of baserunning event i are distributed by

$$N_{ij} | M_{ij} \sim \text{Bin}(M_{ij}, q_{ij}),$$

where M_{ij} is the number of opportunities by player j for baserunning event i and q_{ij} is the true probability that player j attempts baserunning event i . It follows that the number of successes by player j of baserunning event i are distributed similarly:

$$Y_{ij} | N_{ij} \sim \text{Bin}(N_{ij}, r_{ij}),$$

where r_{ij} is the true probability that an attempt of baserunning event i is successful by player j . Again, using Prince Fielder as an example, during the 537 plate appearances when he was on first base with no runner on second base, he attempted to steal second base just 10 times (1.9%) and was successful in only 5 of those attempts (50%). Following similar reasoning for the other methods of taking the extra base, we compute Fielder’s baserunning profile, which is shown in Table 2.

	Adv12	Adv13	Adv14	Adv24	GIDP	GrOut3	Steal2	Steal3	TagUp2	TagUp3
$\bar{\mathbf{q}}_j$	0	0.080	0.345	0.444	0.680	0	0.019	0.008	0.172	0.750
$\bar{\mathbf{r}}_j$	N/A	0.667	0.938	0.850	0.314	N/A	0.500	0.333	0.909	0.905

Table 2: Empirical Estimate of Baserunning Profile for Prince Fielder (= j). In the second cell in the first row, 0.080 is the observed probability of Fielder attempting to advance from 1st to 3rd on a single. Among those attempts, he was successful with probability 0.667 (second row).

To summarize, for any player j , \mathbf{p}_j is a point in a k -dimensional simplex \mathcal{P}^k that represents our estimate of his batting profile. Similarly, a baserunning profile is

³See Appendix B for a list and description of these methods. Note that we do not consider taking more than one base as a viable baserunning event in this model, given its rarity.

	Adv12	Adv13	Adv14	Adv24	GIDP	GrOut3	Steal2	Steal3	TagUp2	TagUp3
$\hat{\mathbf{q}}_j$	0	0.080	0.345	0.444	0.680	0	0.019	0.008	0.172	0.750
$\bar{\mathbf{q}}$	0.097	0.260	0.440	0.614	0.690	0.533	0.074	0.017	0.240	0.577
$\hat{\mathbf{q}}_j$	0.067	0.120	0.373	0.509	0.685	0.478	0.020	0.008	0.205	0.603
$\hat{\mathbf{r}}_j$	N/A	0.667	0.938	0.850	0.314	N/A	0.500	0.333	0.909	0.905
$\bar{\mathbf{r}}$	0.823	0.960	0.939	0.953	0.395	0.994	0.728	0.765	0.863	0.953
$\hat{\mathbf{r}}_j$	0.823	0.947	0.938	0.950	0.371	0.994	0.645	0.700	0.866	0.951

Table 3: Empirical estimates ($\hat{\mathbf{q}}_j, \hat{\mathbf{r}}_j$), league averages ($\bar{\mathbf{q}}, \bar{\mathbf{r}}$), and model point estimates ($\hat{\mathbf{q}}_j, \hat{\mathbf{r}}_j$) of Prince Fielder’s baserunning profile.

defined by $\mathbf{q}_j, \mathbf{r}_j \in [0, 1]^\ell$ for every player j . Thus, the batting and baserunning tendencies and abilities of player j can be summarized by the triple $z_j = (\mathbf{p}_j, \mathbf{q}_j, \mathbf{r}_j) \in \mathcal{P}^k \times [0, 1]^\ell \times [0, 1]^\ell = Z$.

2.2 Fitting the Models

In order to draw insights from our analysis, we want to use accurate estimates of player ability at our disposal. Thus, to model batting ability, we leverage the nested Dirichlet model of player abilities presented in Null (2009). This model has been shown to be among the most accurate publicly available models for point estimation of player ability⁴. Like many popular models of player ability and performance, this model effectively takes observed historical performance data for individual players and regresses them to some general population mean.

To estimate the underlying baserunning attempt and success rates we use a similar approach to Null (2009). For each baserunning event i , we assume that the underlying attempt and success rates for all players are distributed according to a common beta distribution, and we estimate the parameters of these distributions using the data. Thus, given the historical success and failure data for all players (from 2007 to 2009) with respect to baserunning event i , we use the method of Minka (2003) to estimate parameters α_{r_i} and β_{r_i} for each event i . If player j has N_{ij} observed attempts for event i and Y_{ij} observed successes, then the point estimate for \hat{r}_{ij} is $\frac{\alpha_{r_i} + Y_{ij}}{\alpha_{r_i} + \beta_{r_i} + N_{ij}}$. We follow the same approach to estimate \hat{q}_{ij} for all events i and all players j . This approach, again, regresses player historical results towards some mean. The ratio and magnitude of the α and β terms dictate what this mean is and how strongly player estimates are pulled towards that mean. Table 3 presents Prince Fielder’s historical attempt and success rates for all of the baserunning events

⁴It has many other nice properties, most notably providing a valuable posterior distribution of player abilities, but for our purposes in this paper, we are only concerned with the point estimates it provides.

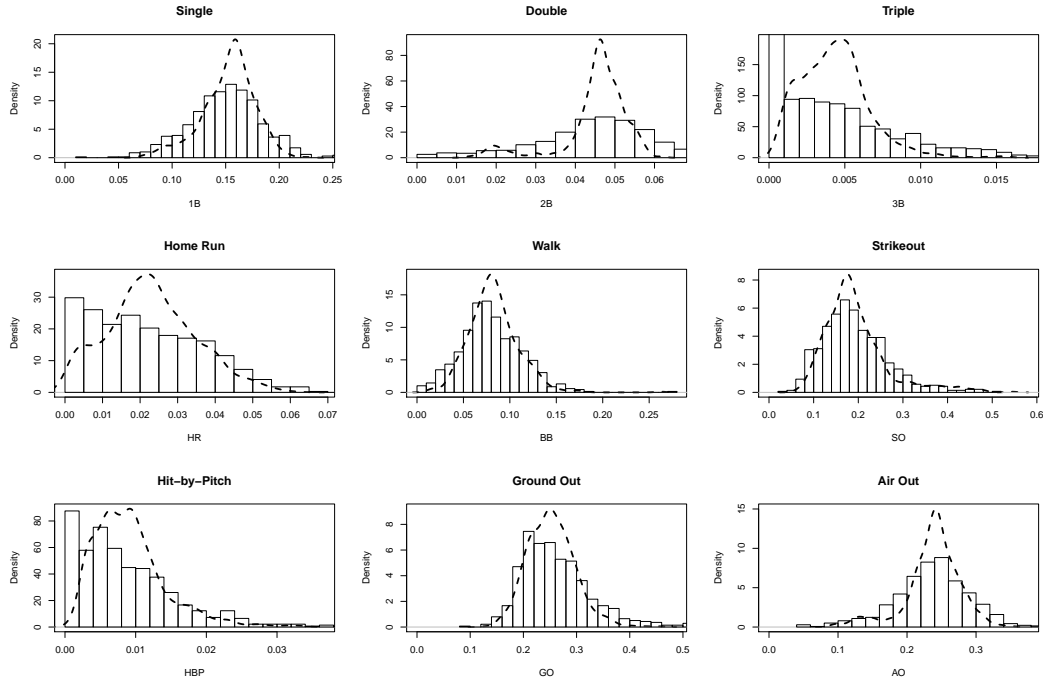


Figure 1: Density estimates of raw and smoothed batting profiles.

studied here, alongside the overall population rates over the period studied and the point estimates for Fielder’s true probability using this model.

Figure 1 shows the distribution of batting ability gleaned from historical data from the 2007 to 2009 seasons in both its raw and smoothed forms. The bimodal nature of batting ability sometimes evident in the smoothed estimates reflects our use of separate models for position players and for pitchers. Table 4 presents our previous empirical estimate of Prince Fielder’s batting ability ($\bar{\mathbf{p}}_j$), alongside the overall league average over these three seasons ($\bar{\mathbf{p}}$) and the model’s estimate of Fielder’s actual ability ($\hat{\mathbf{p}}_j$). Note that due to re-normalization of the $\hat{\mathbf{p}}_j$ vectors, not all values of $\hat{\mathbf{p}}_j$ will lie between the empirical estimate and the league average.

	1B	2B	3B	HR	BB	SO	HBP	GO	AO
$\bar{\mathbf{p}}_j$	0.128	0.048	0.003	0.062	0.136	0.188	0.017	0.202	0.217
$\bar{\mathbf{p}}$	0.136	0.035	0.003	0.015	0.071	0.235	0.007	0.303	0.195
$\hat{\mathbf{p}}_j$	0.133	0.051	0.003	0.052	0.118	0.183	0.017	0.211	0.233

Table 4: Empirical estimate ($\bar{\mathbf{p}}_j$), league average ($\bar{\mathbf{p}}$), and model point estimate ($\hat{\mathbf{p}}_j$) of Prince Fielder’s batting profile.

2.3 Methodology & Procedure

Each inning, an offensive team uses the batting and baserunning abilities of its members to score runs. This process can be modeled by an unknown, real-valued, continuous, scalar function $f : Z \rightarrow \mathbb{R}$. In reality of course, f is a highly complex, non-deterministic function that depends upon many external factors not captured by any of our parameters. While estimates of $f(z)$ can be obtained using aggregate formulae such as Linear Weights Furtado (1999), we want to capture more subtle interactions between batters and baserunners. Thus, we use a known event-based simulation engine, SimulOutCome, whose accuracy has been previously verified by Baumer (2009).

A full description of SimulOutCome is beyond the scope of this paper, but interested readers are encouraged to consult Baumer (2009) for additional details, or download the actual source code, which is freely available on SourceForge⁵. Briefly, the simulation engine works by loading an offensive profile z_j for each of the nine copies of player j , and then constructing complete innings through randomized event-based simulation. Event generation is governed by the probabilities in z_j , and simple but realistic baserunning logic. Additional variation in the form of randomized fielding errors is included at the appropriate league-wide frequencies. The output generated is synthetic play-by-play data.

Our procedure is to use SimulOutCome as a black box to estimate values of f , the expected number of runs per inning, for specified values of z , a player's offensive profile. Let \hat{z}_j be our estimate of player j 's offensive profile. Our primary focus in this paper is to understand the impact upon run scoring of batting and baserunning ability, so we want to find the rate of change of f in neighborhoods around \hat{z}_j . To do so, we fix \hat{z}_j , but then estimate f while allowing q_i (the probability of attempting baserunning advancement method i) and r_i (the corresponding probability of success) to vary. In this application, for each method of taking the extra base i , we vary q_i and r_i within a 20×20 grid of values in $[0, 1]^2$. That is, if we consider $\hat{z}_j(q_i, r_i)$ to be a function of two variables q_i and r_i (but otherwise fixed and equal to \hat{z}_j), we compute $f(\hat{z}_j(q_i, r_i))$ for 400 different combinations of q_i and r_i for each player j . The value of each cell in this grid is estimated as the mean number of runs scored per 100,000 simulated innings. Thus, since we have 10 methods of taking the extra base, complete analysis of each player requires simulating 400 million innings⁶.

In reality, a lineup is a heterogeneous collection of players with different offensive profiles. However, in this case our goal is to understand the effect that

⁵<http://sourceforge.net/projects/westest/>.

⁶This took an average of about 1.6 seconds per one million innings on a consumer-grade 64-bit AMD Phenom machine.

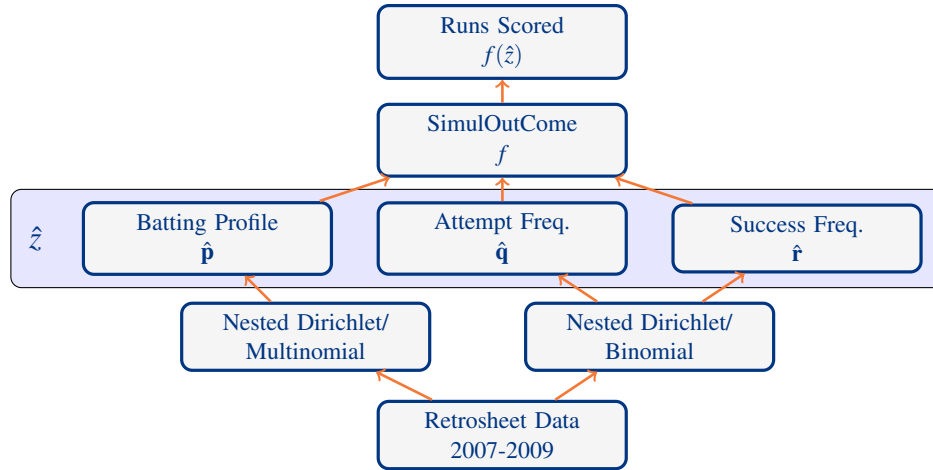


Figure 2: Schematic of our procedure. Batting data from Retrosheet is fed through a Nested Dirichlet-Multinomial model to produce a batting profile $\hat{\mathbf{p}}$, while baserunning data is fed through a Nested Dirichlet-Binomial model to produce baserunning profiles $\hat{\mathbf{q}}$ and $\hat{\mathbf{r}}$. The three comprise the offensive profile $\hat{\mathbf{z}}$, which is fed into the simulated function f to produce the estimate of runs scored $f(\hat{\mathbf{z}})$.

batting and baserunning collectively have upon run scoring. Therefore, we perform our tests on a homogeneous lineup consisting of 9 identical copies of the same offensive profile. In this manner, we hope to gain insight into the interplay between batting and baserunning abilities of a particular style of play.

A schematic of our procedure is shown in Figure 2. We begin with raw Retrosheet data from 2007 to 2009 and feed it into our Bayesian models to obtain smoothed estimates of each player’s offensive profile $\hat{\mathbf{z}}$. We then use SimulOutCome as a black box approximating f , and pass it $\hat{\mathbf{z}}$ to compute $f(\hat{\mathbf{z}})$, the estimated number of runs scored per inning by a team consisting of 9 players with offensive profile $\hat{\mathbf{z}}$.

2.4 Players Tested

Given the computational time required, we focused on 21 prominent major league players listed in Table 5. These players were chosen somewhat arbitrarily to fulfill a variety of different offensive archetypes. Specifically, we targeted both power hitters and speedy baserunners, while also considering how often each player strikes out. For example, Albert Pujols and Ryan Howard are two of the most feared sluggers in the game, but Howard strikes out frequently, while Pujols does not. Both Luis Castillo and Curtis Granderson rely on speed, but Granderson hits for power

Player	Player	Player
Abreu, Bobby	Dunn, Adam	Markakis, Nick
Brayan, Russell	Fielder, Prince	Pedroia, Dustin
Cabrera, Miguel	Gonzalez, Adrian	Pujols, Albert
Cabrera, Orlando	Granderson, Curtis	Ramirez, Hanley
Castillo, Luis	Holliday, Matt	Reyes, Jose
Crawford, Carl	Howard, Ryan	Reynolds, Mark
Cust, Jack	Jeter, Derek	Suzuki, Ichiro

Table 5: Players Tested

while Castillo does not. By studying players of vastly different archetypes, we hope to answer questions about how differences in offensive profiles affect players' baserunning strategies.

3 Results

For our analysis, we used all Major League Baseball data from 2007 to 2009 and applied the fitting procedure discussed in Section 2.3 to estimate (1) a league average \bar{z} vector, as well as (2) \hat{z}_j vectors for the 21 individual players listed above. We summarize our results using surface plots that show the change in the expected number of runs scored per 9 innings as a function of attempt and success rate. We do this for each of the ℓ methods of taking the extra base separately. The magnitude and direction of steepest increase from the target point is of primary interest.

3.1 Surface Plots

The surfaces plotted are defined as $g_{ij}(q_i, r_i) = 9 \cdot (f(\hat{z}_j(q_i, r_i)) - f(\hat{z}_j(\hat{q}_{ij}, \hat{r}_{ij})))$. This scaling provides a natural interpretation of values for g_{ij} : it is the expected number of runs per game gained or lost by a lineup consisting of 9 copies of player j , due to changes in the frequency of attempts and successes of the i^{th} baserunning event. A mild 2-dimensional kernel smoother is applied to improve legibility, and a white crosshair is placed on each plot to indicate the location of the point which describes the player's estimated profile. Note that the value of the surface at this point ($\hat{z}_j(\hat{q}_{ij}, \hat{r}_{ij})$) is 0 by definition, but the crosshair may not intersect the 0 contour line exactly due to the application of the kernel smoother.

These surface plots can be interpreted in a variety of ways. Their primary function is to give a sense of how a lineup consisting solely of one player with certain batting abilities would operate given all combinations of their baserunning

profile (i.e. attempt and success rates) for a particular baserunning scenario. That is, if a lineup has the attributes of a particular player, we can analyze the effect on offensive production if baserunning behavior/skills were to change. This information enables a manager to make informed decisions about whether to encourage (or discourage) baserunning attempts of a certain type. Moreover, armed with a quantitative understanding of the marginal benefit of specific baserunning acts for *his* team, he could focus his attention on developing those baserunning skills that are particularly valuable.

Outside of this general purpose, these surface plots also serve to inform baseball researchers on conventional wisdom held by the baseball community, such as Harold Reynold’s opinion of slow, power hitters similar to Adam Dunn. In our forthcoming analysis, we address several of these opinions.

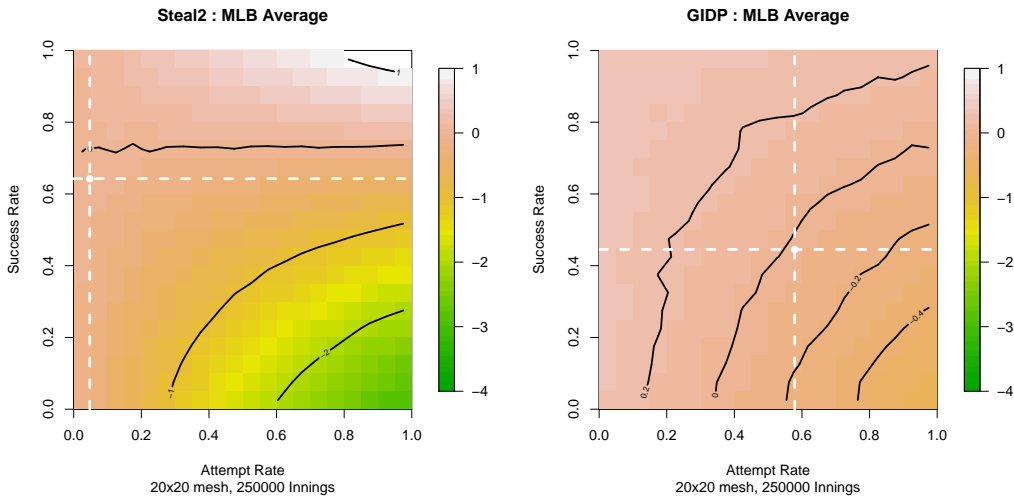


Figure 3: Surface Maps for MLB Average for Steal2 and GIDP. At left (right), the surface represents the expected change in runs scored per 9 innings for a lineup of the average MLB player as the attempt rate and the success rate for stealing 2nd base (grounding into a double play) vary.

3.2 General Observations

In Figure 3, two such plots are shown for a lineup composed of the Major League average player at that position in the batting order⁷. For each method of taking the extra base, the lower right hand corner of the plot (always attempt, never succeed)

⁷The effects of Adv12, TagUp2, TagUp3, and GrOut3 upon run-scoring were relatively small and uninteresting, and accordingly have been omitted for space considerations.

is the worst-case scenario for the offense (i.e. the lowest run-scoring region in the plot). For most methods of taking the extra base the upper right hand corner of the plot (always attempt, always succeed) represents the best-case scenario for the offense (i.e. the highest run-scoring region of the plot). The exception is GIDP, where we recall that a GIDP attempt corresponds to a potential force out at second base⁸. Thus, two distinct surface patterns emerge in the gradient visualizations across all lineups, where the optimal region is either the upper right or the upper left hand corner.

Zero Contour Line One point of interest is the location of the contour line corresponding to no change in run-scoring. This line separates the plot into two distinct regions: one where run-scoring increases and the other where run-scoring decreases. While this line in the Steal2: MLB Average plot (see Figure 3) is almost exactly horizontal, it is not obvious that this should always be the case, and in most cases, it is not. The zero contour line traces the combination of attempt and success frequencies that result in no change in run scoring, and there is no reason to believe, in general, that the success rate required to maintain run scoring would be constant with respect to attempt frequency. For example, consider the baserunning event of advancing from first base to third base on a single. For light-hitting lineups, it is conceivable that the number of runs produced by that lineup could increase if advances were attempted more often, even if they occurred at a lower success rate. Moreover, that this line often has shape (in two dimensions) in our analysis represents an improvement over less sophisticated methods (e.g. Linear Weights), which cannot capture this type of interaction.

However, in the Steal2: MLB Average plot, the zero contour line is horizontal, suggesting a break even stolen base success rate. Baseball researchers have long debated the existence of a such a threshold, over which run scoring is aided, but under which run scoring will suffer. Estimates of this cut-off range from 67% to 80% [Click (2007), Tango et al. (2007)]. The surface plot confirms the existence of such a threshold *for the average team* and estimates its value at approximately 76%. Note that as we move to the right away from the white crosshair (\bar{z}), run scoring decreases. This suggests that on average, players steal second base with a suboptimal success rate (64%). We will show examples of several players for whom no such threshold appears to exist.

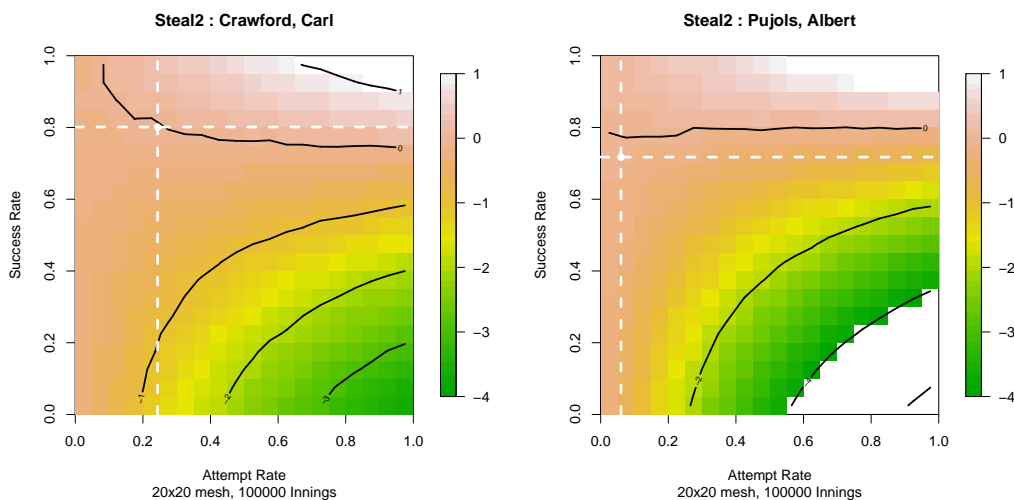


Figure 4: Changes in expected runs scored for lineups of Carl Crawford (left) and Albert Pujols (right) as the attempt and success rates for stealing 2nd base vary. Crawford should steal *more* often, since his 0 contour line is curved. For Pujols there is no benefit to stealing more often.

3.3 Offensive Profiles Affect Baserunning Strategies

In Figure 4, we compare the Steal2 plots for Carl Crawford and Albert Pujols. Crawford's zero contour line is curved, suggesting that if he could steal second base more often while maintaining his current success rate, run scoring for a team of players like him would increase. This is not true for all players, and in particular for Pujols, who has a largely horizontal zero contour line. This effect is a combination of two facts: 1) Pujols's success rate is lower than Crawford's; 2) the value of a stolen base is lower for a team comprised of Pujols clones, because of his extraordinary batting ability, particularly with respect to power. That is, not only is it the case that Pujols will not aid run scoring by stealing more often at his current success rate, but in fact there is little for a team of Pujols clones to gain from stealing more often unless they are nearly always successful. Our interpretation is that Pujols's batting ability is so great, that the loss of an out on the basepaths far outweighs the advantage of a single base gained via a steal⁹. Note (not coincidentally) how

⁸See Baumer (2009) for a more detailed explanation.

⁹This deduction is evidently lost on Pujols himself, who, with the score tied, supposedly called a hit-and-run play while at bat in the 7th inning of Game 5 of the 2011 World Series, only to see teammate Allen Craig thrown out at second base. Pujols was then intentionally walked, and the Cardinals did not score.

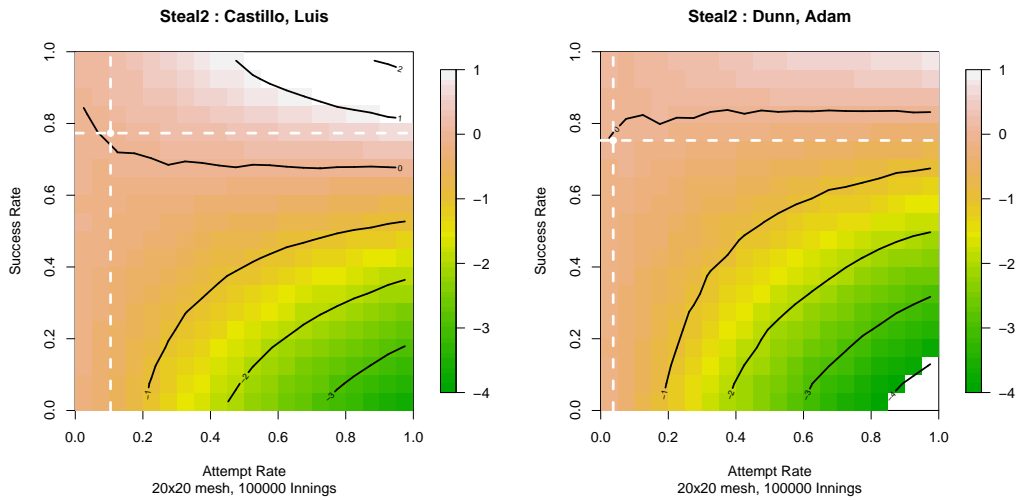


Figure 5: Changes in expected runs scored for lineups of Luis Castillo (left) and Adam Dunn (right) as the attempt and success rates for stealing 2nd base vary. Castillo has much more to gain from stealing bases than Dunn, and requires a lower rate to do so.

the range of g_{ij} values in Pujols's plot is much greater than in Crawford's. In fact, Pujols could lose more than twice as many runs per game due to poor baserunning than Crawford.

Of course, there are game-theoretic concerns that make advocating an increase in baserunning attempt rate somewhat problematic. That is, it may not be possible for Crawford to steal more often at his current success rate. It may be the case that, for example, he always steals when he knows he can make it, but rarely otherwise. Nevertheless, part of the value of our analysis is that it allows a decision-maker to quantify the impact of such trade-offs.

A second example of this discrepancy is evident in Figure 5 when comparing Luis Castillo's Steal2 plot to Adam Dunn's. Although both players steal at roughly the same rate, the light-hitting Castillo has much more to gain from stealing bases than the lumbering, slugging Dunn. This is clear from the upper right hand corner of the plots, in which Castillo can gain up to 2 runs per game, while Dunn can gain less than 1. Moreover, the bar to increase run scoring is set much lower for Castillo than it is for Dunn. Following their zero contour lines, Castillo can increase run scoring with success rates under 70%, while Dunn requires at least 80% for most attempt frequencies. Again, our interpretation is that this is a reflection of their dramatically different batting profiles, since a team of Castillo clones will get very few extra base hits, and thus has to work harder on the bases to move runners around. Conversely, a

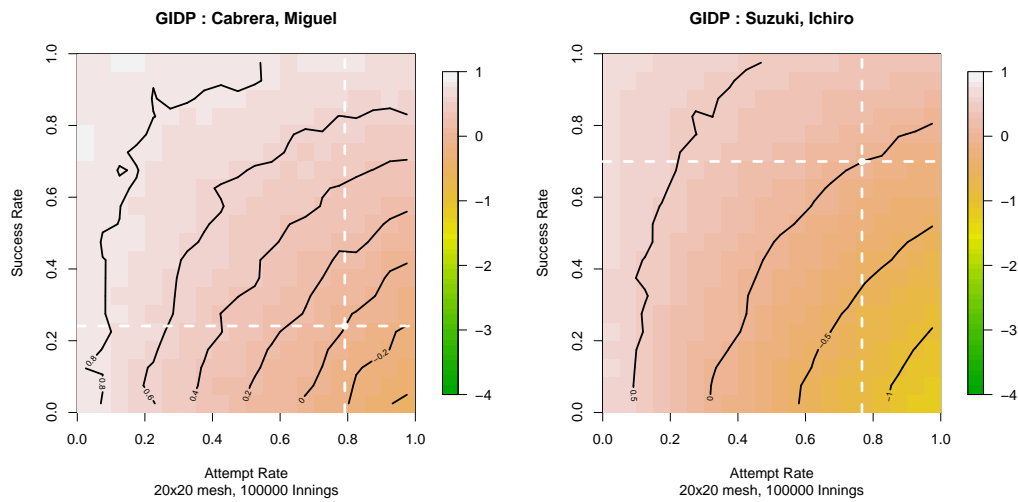


Figure 6: Changes in expected runs scored for lineups of Miguel Cabrera (left) and Ichiro Suzuki (right) as the attempt and success rates for grounding into a double play vary. Ichiro rarely gets doubled up. If Cabrera ran as well as Ichiro, he would save his team about 0.6 runs per game

team of Dunn clones is better served by staying put and waiting for an extra base hit. This comparison most directly addresses Harold Reynolds's introductory comment, as our findings suggest that a team of Adam Dunns are *well served* by clogging the bases.

3.4 Getting doubled up hurts

In Figure 6, we compare the GIDP plots for Miguel Cabrera and Ichiro Suzuki. It is perhaps surprising to note that for both players, the defense is able to force the runner at 2nd base nearly 80% of the time that they ground out in a double play situation. However, while the double play is completed against Cabrera routinely (about 75% of the time), it is a much rarer feat against Ichiro (about 30% of the time). But the extra outs lost are more valuable to Cabrera than they are to Ichiro, since Cabrera is a more productive and powerful hitter. It is apparent from the plot that if Cabrera beat out potential double play balls as often as Ichiro, a team of Cabrera clones would gain about 0.6 runs per game, a significant margin.

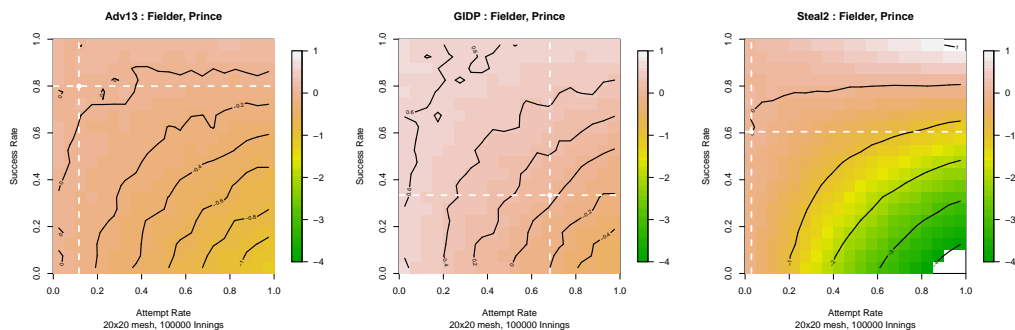


Figure 7: Changes in expected runs scored for lineups of Prince Fielder as the attempt and success rates for advancing from 1st to 3rd on a single (left), grounding into a double play (center), and stealing 2nd base (right) vary. Prince Fielder is frequently doubled up, as well as being thrown out often when trying to advance or steal. Furthermore, his offensive profile is strong enough to make any potential gains from such attempts practically insignificant.

3.5 Prince Fielder is not a good baserunner

Clearly, running into outs on the basepaths is detrimental to a team’s ability to score runs. We have illustrated that stronger hitters have less to gain from stealing bases, and must do so at higher success rates than less powerful hitters in order to make it worthwhile. We’ve also shown that failing to beat out double play balls can have a significant negative impact upon run scoring. Each of these observations apply to the case of Prince Fielder. Fielder is an outstanding hitter, with exceptional home run power and the ability to get on base often, but a poor baserunner. Figure 7 reveals that Fielder is frequently doubled up, and while he does not attempt to steal or advance often, he is frequently thrown out when doing so. But perhaps the most important insight that our analysis provides, is that Fielder is such a good hitter that he has little to gain from aggressive baserunning even if he were successful. Thus, his shortcomings are not only of execution, but also of strategy.

4 Conclusion

Our approach and analysis provided graphical intuition into the relative value of the various baserunning abilities (e.g. the potential impact of increasing stolen base success rates is much greater than that of increasing success rates advancing from first to third on a single). Moreover, our framework enables us to estimate the quantitative impact of particular baserunning decisions (e.g. if players attempted

to advance from second to home on a single $x\%$ more often and maintained their current success rate, they could increase average scoring by y runs). This naturally leads to evaluations of the optimality of current Major League baserunning strategies. The surface plots generated in this paper help to identify dramatic changes in run scoring based on increases/decreases to attempt/success rates, especially when stealing second base. Perhaps most importantly, we showed that this type of analysis can vary dramatically for different types of lineups. We have seen that a player's initial offensive profile can have a great impact on the magnitude of effects for all baserunning scenarios, even those as basic as grounding into a double play. Notably, we have observed that certain teams (such as a lineup of all-or-nothing hitters like Adam Dunn or Prince Fielder) may have good reason to "clog the bases," because it is optimal for them to do so. Thus, in this paper we not only corroborate existing conventional wisdom about stealing bases, but also support new theories about more stationary hitters.

Nonetheless, there is much more research on this topic that could potentially be done. To truly generate actionable insights out of this research, future work will extend this analysis by looking at the interactions of actual lineups; that is, how should Derek Jeter run the bases when Alex Rodriguez and Mark Texiera are coming up to bat behind him? We will examine how optimal strategies and the impact of baserunning changes for particular teams against particular opponents. Finally, it is important to note that we have ignored, or assumed to be constant, many variables (e.g. the score, the number of outs, the opposing pitcher/defense, etc.) that most likely affect offensive strategies in the real world. Since this analysis focuses on aggregate measures of run scoring over a large number of innings, it did not seem prudent to add the complexity that would have been required to incorporate these effects into the model. However, future work may investigate how these variables affect the findings we have made in this paper.

A Batting Outcomes

- $1B$ = single
- $2B$ = double
- $3B$ = triple
- HR = home run
- BB = walk
- SO = strikeout
- HBP = hit-by-pitch
- GO = groundout
- AO = air out

B Baserunning Events

- Adv12 = Advancing to second base after an RBI single
- Adv13 = Advancing from first base to third base on a single
- Adv14 = Scoring from first base on a double
- Adv24 = Scoring from second base on a single
- GIDP = Beating out a double play opportunity
- GrOut3 = Scoring from third base on a groundout
- Steal2 = Stealing second base when second base is unoccupied
- Steal3 = Stealing third base when third base is unoccupied
- TagUp2 = Tagging up from second base on an air out
- TagUp3 = Tagging up from third base on an air out

References

- Albert, J. (2006): “Pitching statistics, talent and luck, and the best strikeout seasons of all-time,” *Journal of Quantitative Analysis in Sports*, 2, 1–30.
- Baumer, B. (2009): “Using Simulation to Estimate the Impact of Baserunning Ability in Baseball,” *Journal of Quantitative Analysis in Sports*, 5, 1–16.
- Baumer, B. and P. Terlecky (2010): “Improved Estimates for the Impact of Baserunning in Baseball,” in *JSM Proceedings, Statistics in Sports, ASA*.
- Click, J. (2005): “Station to Station: the Expensive Art of Baserunning,” in *Baseball Propectus 2005*, New York: Workman Publishing, 511–519.
- Click, J. (2007): “What if Rickey Henderson Had Pete Incaviglia’s Legs?” in *Baseball Between the Numbers: Why Everything You Know About the Game Is Wrong*, Basic Books, chapter 4-1, 112–126.
- Fox, D. (2005): “Circle the Wagons: Running the Bases Part III,” <http://www.hardballtimes.com/main/article/circle-the-wagons-running-the-bases-part-iii/>.
- Furtado, J. (1999): “Introducing XR (Extrapolated Runs),” in *The 1999 Big Bad Baseball Annual: The Book Baseball Deserves*, Masters Press, 479–484.
- Minka, T. (2003): “Estimating a dirichlet distribution,” *Annals of Physics*, 2000, 1–13.
- Null, B. (2009): “Modeling Baseball Player Ability with a Nested Dirichlet Distribution,” *Journal of Quantitative Analysis in Sports*, 5, 1–37.
- Reynolds, H. (2009): “Enjoy it for what its worth,” <http://haroldreynolds.mlblogs.com/2009/06/18/enjoy-it-for-what-its-worth/>.
- Tango, T., M. Lichtman, and A. Dolphin (2007): *The Book: Playing the Percentages in Baseball*, Potomac Books Inc.