Howard, J. (2018). 'Keep the line moving': residual system state potential as a baseball hitting-style indicator. *Journal of Sports Economics & Management*, 8(1), 16-36.

'KEEP THE LINE MOVING': RESIDUAL SYSTEM STATE POTENTIAL AS A BASEBALL HITTING-STYLE INDICATOR

Mantener la línea en movimiento: potencial residual del estado del sistema como indicador del estilo de bateo de béisbol

Jeffrey Howard

Northern State University, USA

ABSTRACT: The sport of baseball has a strong tradition of utilizing forward-projecting probabilistic measures, such as batting average, for performance assessment. Albeit highly informative and quite accurate, such measure of batting probability merely looks at how a player can probabilistically act upon the system in a forward manner, based on how they have acted upon that system in the past, but does not provide information about the state of the system itself once a player completes an at-bat. The current paper proposes a statistic called the Residual Run Potential Index (RRPI) as a measure of the 'residual system state' as it remains after a completed at bat. Derived from Retrosheet (2017) event data, the RRPI reflects a system-efficiency measure of player hitting style that retains maximized run-scoring potential of the system state. RRPI calculations derived from a proprietary software simulation experiment, as well as for several Major League Baseball (MLB) Hall-of-Fame players, are also presented.

Keywords: baseball, batter, performance, statistic, analytic, system

RESUMEN: *El deporte del béisbol tiene una fuerte tradición de utilizar medidas probabilísticas de proyección futura, como es el promedio de bateo para la evaluación del rendimiento. Aunque altamente informativa y absolutamente exacta, tal medida de la probabilidad del bateo examina simplemente cómo un jugador puede actuar probabilísticamente sobre el sistema de una manera adelantada basada en cómo han actuado sobre ese sistema en el pasado, pero no proporciona la información sobre el estado del propio sistema una vez que un jugador termine un at-bat. El presente trabajo propone una estadístico llamado Índice Potencial de Carrera Residual (RRPI, por sus siglas en inglés) como una medida del "estado residual del sistema" que permanece después de un bateo. Derivado de los datos de eventos de Retrosheet (2017), el RRPI refleja una medida de eficiencia del sistema de estilo de golpeo del jugador que retiene el potencial maximizado de puntuación de carrera del estado del sistema. También se presentan los cálculos de RRPI derivados de un experimento de simulación de software propietario, así como para varios jugadores del Salón de la Fama de las Grandes Ligas de Béisbol (MLB).*

Palabras Clave: baseball, batear, rendimiento, estadístico, analítico, sistema

Received/recibido: 13-03-2018 Accepted/aceptado: 30-04-2018

Contact information:

Correspondence author

Jeffrey Howard
jeffrey.howard@northern.edu
Northern State University
1200 S Jay St; Aberdeen, SD 57401

Introduction

Perhaps the most well-known statistic in baseball is the batting average (BA). The BA provides a good measure of hitter probability in a number of situations, and it's ubiquitously understood—even by laypersons. However, as a descriptive statistic, the BA merely reflects past performance, so as to give insight into future performance—which is good insight into the way a batter 'acts upon a system'—but devoid of insight into the condition/state that a player leaves the system post at-bat. Thus the BA does not provide a snapshot of the 'system', but rather a gauge of participant probability to act upon the system in a certain way for any specific set of circumstances.

Other non-BA means have been proposed for assessing hitter performance, such as that whereby runners in scoring position are compared vs bases empty, no outs situation so as to reveal if player probability of getting a hit might be parsed by analyzing such situation comparisons (Albert, 2002). However such a system again is forward-projecting so as to assess likelihood that a hitter will act upon the system state in a particular way, and not how the system state remains after it has been acted upon. In addition, there are at least 10 specific research-documented models of determining offensive performance in the game of baseball (Bennett & Flueck, 1983), again none of which are predicated upon assessing the residual system state as an after-the-event measure of hitter performance. There are even models that address neither forward measures such as BA, or residual system state, but rather non-statistical characteristics and factors such as age and experience (Ng, 2017), or of ongoing and consistent motivational performance—or lack thereof, called 'shirking'—whereby long-term contract incentives have been shown to drive player productivity down (Krautmann & Solow, 2009). Nevertheless, 'how things remain' post-performance seems to garner a dearth of attention in sabermetrics, and the means by which such 'residual inquiry' it can give insight into player performance must not be devalued by default.

The state of the system after an at-bat—the residual system state—can be a valuable commodity when it comes to assessing player 'hitting-style'. Such system state information tells the story of how the hitter 'left things' once an at-bat is completed. The current paper proposes a statistic called the Residual Run Potential Index (RRPI) that reflects not hitter probabilistic outcome from an at-bat, but rather the residual state of the system as it remains after an at-bat.

The essence of the RRPI.

Within the scope of the RRPI, the term 'system state potential' as a concept referring to a baseball diamond and the distribution of such potential using three bases—first, second, and third, is not decided upon flippantly. The decision derives from the influence that the number three has over the entire game itself. Three strikes (opportunities) given to a batter to succeed in some way; three outs to each half inning; three times through a batting order to pitch a perfect game; three runners as the maximum potential the system can contain on three bases; three outfielders being the standard defensive perimeter for any fair ball leaving the infield (Caraher, 1991).

When it comes to referencing ‘system state potential’ in baseball as a performance parameter, perhaps the most common term that comes to mind is men ‘left on base’(LOB)—which represents offensive ‘potential lost’ and not capitalized upon. It is the LOB that is the creative catalyst for the development of the RRPI. The LOB statistic is looked upon as a performance parameter for both pitching and offense, and demonstrates how the system state remains after an at-bat—but it does so only for the last out of an inning. Similarly, the front-half of the RRPI—the ‘Residual Run Potential’ (RRP)—looks at LOB with respect to ‘hitters’ as a function of hitting-style and how the bases remain after an at-bat. In essence, the RRP is the LOB as assessed after each at-bat with respect to hitter influence on runners left on base, rather than once at the end of each half inning to assess pitcher performance or wasted offensive capital. LOB has been shown to have a strong formulaic influence on pitching measures such as ERA (Beneventano, Berger, & Weinberg, 2012), and being that the RRP is essentially looking at the converse of the LOB for batters, it reasons that similar influence, being positive or negative, may also exist. The RRPI ‘composite’ is in essence a performance statistic representing batter influence on the overall system state potential, as represented by a hitter’s hitting-style and the average of the number of men left on base for a hitter over some period of time. The goal for a defense at any point in time is to have low RRP situation; the goal of the offense is to have high RRP situation—and the LOB statistic merely reflects the current RRP when the third out is recorded. The RRPI is all about the concept of average RRP (ARRP), and how batters contribute over some defined period of time to post at-bat RRP as a function of their hitting style.

An analogy of a ‘restaurant inspector’ provides a good basic explanation of the RRPI and what it entails. Imagine that the health department went into the kitchen of a restaurant one hour before closing, so as to evaluate the 'state of cleanliness’ of that kitchen. The evaluation would indeed tell you the state of cleanliness of that kitchen, but it would tell you nothing about any of the employees responsible for leaving the kitchen in that 'state'—such as their hand washing behaviors, their personal attire (hairnets or hats etc.), their current health (sick or healthy), or the manner in which they handle food, utensils, or machinery. One would need to get data on the employees individually (specific information about them) so as to determine the degree to which each employee contributes to the 'state of kitchen cleanliness'. It is assumed that the state of kitchen cleanliness is indeed due to these employee factors, but the inspector cleanliness rating of the kitchen is merely a global reflection of its overall state, absent any specific highresolution detail regarding the characteristics of those who contributed to that overall state.

Whereas an ‘average’ might provide a measure of say, how many dishes a certain chef prepares in any one hour time period, the RRPI would analogously make indication as to the ‘state of the kitchen’ after said time period and preparations by that chef were completed. It is easy to see that the ‘state of a kitchen’ could be equated with the probability of simple food contamination, or even food poisoning—neither probability of which could be estimated from the average number of dishes a chef prepared in any single hour at work in a kitchen. Thus, the chef analogy demonstrates that the probability of receiving a specific ‘result’ is also determined in part by the a priori ‘inherited system state’, and not always solely via forward-projecting measures of participant performance within the system.

Similarly, the probability of scoring runs is not always wholly determined by a hitter, it’s facilitated by the system state they inherit for each at-bat—a state that is determined by the previous batter—and each batter is both a beneficiary of system state inheritance, and one who bequeaths a system state to others.

Calculating the RRPI

Average Residual Run Potential (ARRP).

The RRPI (Figure 1) is comprised of two components, the first of which is called the Average Residual Run Potential (ARRP), and the second which is the total of the weighted (squared) averages of five different types of base-acquisition events. The ARRP statistic reflects the average state of the bases (as a system) that remains following a player at-bat—which can also be referred to as the ‘residual system state’. The ARRP can be calculated for a team as well as at the individual level, and can reflect an entire season of data, a custom defined time range, or an entire career.

Figure 1. The Residual Run Potential Index (RRPI) equation

$$RRPI = \frac{\sum b_i}{AB} + \left[\left(\frac{s}{H}\right)^2 + \left(\frac{d}{H}\right)^2 + \left(\frac{t}{H}\right)^2 + \left(\frac{h}{H}\right)^2 + \left(\frac{w}{w + AB}\right)^2\right]$$

b = residual base state	h = home runs
s = singles	H = hits
d = doubles	w = walks
t = triples	AB = at bats

ARRP calculation can be carried out by using Retrosheet (2017) event data database field END_BASES_CD so as to acquire data representing the base-state after a player at-bat over a specified period of time, as defined via database query. This Retrosheet (2017) event field is comprised of values ranging from 1 to 7 that represent each possible state for the bases after an at-bat (Marchi & Albert, 2009). Albeit the default coding values are helpful in their own right, it is necessary to re-code these base-state END_BASES_CD event 1 to 7 value codes (Table 1) so as to allow for creation of a coding procedure reflective of and sensitive to the potential of the system-state. Thus a coding value procedure was created to represent the potential of the system (the bases) by accounting for both [1] the number of baserunners (3), and [2] the quantity of the system-state that those men have usurped with respect to their current position on the bases.

The bases—first, second, and third—as a 'system' are partitioned into two pathways, each with a .33 value—first base to second base equaling .33, and second base to third base equaling .33—with the maximum amount that the system state can hold equaling .99 plus 3 baserunners (3|.99). For example, this 3|.99—the maximum that the system state can hold—indicates there are three men on base, and that those three men have traversed the .33 pathways a collective total of 3 times (3 x .33 = .99). The vertical separator symbol "|" has been added to separate the number of men currently on base from the summative value that the men on base have traversed within the system. Therefore in the 3|.99 example, the man on third has traveled both pathways (2 x .33 = .66), the man on second has traveled one pathway (1 x .33 = .33) and the man on first has yet to move across either pathway (0 x .33 = .00). Such a system allows for baserunner occupancy scenarios such as first and third, second and third, and first and second to be weighted with respect to the balance of the overall system, as there is an inherent difference between, for example, men at first and third and men at second and third regarding probability of scoring. The men at second and third demonstrate a 'back-half weighting' of the system (closer to scoring) more so than do the men at first and third—who represent a 'balance' of the system—or men at first and second who represent a 'front-half weighting' of the system (farther from scoring).

Table 1. Code modification for END_BASES_CD event

Code	**Bases occupancy**	**Re-code value**
0	Empty	0|.00
1	1B only	1|.00
2	2B only	1|.33
3	1B & 2B	2|.33
4	3B only	1|.66
5	1B & 3B	2|.66
6	2B & 3B	2|.99
7	Loaded	3|.99

The Sum of the Squares of Hit-Type and Walk Ratios.

The second component of the RRPI is comprised of the sums of the squares (SS) of the ratios for singles, doubles, triples, home runs, and walks. The purpose in using the sums of the squares is to give a 'weighting' to all the ratio values so as to give some data points either a heavier or lighter influence amongst their group (Deviant, 2016). Being that singles are the most common hit, and singles also provide the greatest value with respect to maintaining both 'high system potential' and maximizing of runs scored, the weightings diminish the value of doubles, triples, and home runs—types of events that are greater at 'depleting' system potential.

Upon having calculated both the ARRP and the Hit-Walk Ratio SS values, one simply adds these values together to arrive at the RRPI composite value. At this point, it is important to clarify why the addition of the SS Hit-Walk values is included, rather than just rely on the ARRP value itself. To make this clarification, one must revisit the analogy of the 'restaurant inspector' so as to be reminded that although the 'cleanliness rating' that the inspector provides is indeed explanatory regarding the 'state' of the kitchen, one is reminded that this rating value says nothing about the characteristics of those who used the kitchen. Thus the addition of the SS Hit-Walk Ratio values provides a unique value that is descriptive of 'hitting style'—an important addition to the RRPI when one wants to assess the hitting style of an individual player. In addition, the ARRP does not purely reflect the hitting style of a hitter in that there are a number of events that can occur during a player at-bat that could change the system potential for a hitter when his at-bat concludes—events that are not caused by the hitter—such as runners being picked off, runners caught stealing, runners interfering with the defense, player ejection for arguing/fighting, wild pitch/passed ball, or a base runner being thrown out on a base-advance attempt etc. Each of these aforementioned events will affect the base-state potential, which could then reflect on the hitter once the at-bat is over (e.g. one man on base, no outs, and a pickoff accounts for out #1; batter then hits a popup for out number two leaving the system state at 0 for the next batter).

Table 2 Retrosheet data counts for singles, doubles, triples, and home runs, walks 1945-2015

Retrosheet data 1945 to 2015	
Event Type	**Count**
Singles	1,617,757
Doubles	411,762
Triples	57,937
Home Runs	233,489
Walks	791,793

With respect to the historical data, there is definitive support that more singles are hit than are doubles, triples, or home runs (Table 2). Capitalizing on this via the RRPI, if a team wishes to recruit a hitter that 'leaves' the system state at a higher maximum run-scoring potential on average, then indeed they want a more 'homogenous' hitter—this is the very intent of the RRPI—to assist in highlighting who these hitters are. The converse is also true—the RRPI can help reveal which hitters are not homogenous with respect to hit type.

Thus Table 2 data shows the value of hitting singles, and validates the incidence of such over a 71-year period. Table 3 highlights a comparison of singles vs doubles with respect to their influence on ARRP—the front-end component of the RRPI. The Table 3 data demonstrate ARRP results for hitting 10 singles in a row vs 10 doubles in a row, while holding baserunner advance constant (runners always advance the same number of bases as the batter), so as to highlight the resulting ARRP values for each scenario.

Table 3. ARRP example for 10 hits in a row (singles vs doubles)

Continuous Singles		**Continuous Doubles**	
Batter #	**Men on base post-hit**	**Batter #**	**Men on base post-hit**
1	1	1	1
2	2	2	1
3	3	3	1
4	3	4	1
5	3	5	1
6	3	6	1
7	3	7	1
8	3	8	1
9	3	9	1
10	3	10	1
ARRP	27/10 = 2.7	ARRP	10/10 = 1

*Note: runners always advance same number of bases as batter

Continuously hitting 10 doubles in a row generates 9 runs when holding base-advance by runners constant; continuously hitting 10 singles in a row produces only 7 runs under this same scenario. However, the ARRP—the Average Residual Run Potential across all 10 at bats in each scenario—is quite different. For doubles the ARRP across the 10 at-bats = 9/10 or .90; for singles the ARRP = 2.7—thus the system when being advanced predominantly by singles will nearly always sit in a higher average run potential state, which is more advantageous in real life where baserunner advance is not being held constant, and the potential for the intermittent 'big hit' always looms on the horizon. In fact, mixing homogenous hit-type hitters with non-homogenous hit-type hitters across a batting lineup would suggest a highly efficient means of keeping run-scoring potential of the system state at a high average, while simultaneously providing high probability of the occasional/intermittent 'big hit' event to drain the system state potential. Perhaps most importantly, such homogeneity across players has been effectively applied as tool to hold certain statistical elements constant in the analyses of pricing athletes regarding their 'expected financial return', and to assist in variability assessment for 'return on investment' (Kedar-Levy & Bar-Eli, 2008).

Table 4. Base-state change-event type percentages and squares of percentages for Rod Carew 1982

Event Type	**Carew, R. 1982**	
	%	**% squared**
Singles	.8024	.6438
Doubles	.1497	.0224
Triples	.0299	.0009
Home Runs	.0180	.0003
Walks	.1059*	.0112
TOTALS	1.11	.6787

*Walks = BB / (AB + BB)

To illustrate how the RRPI reveals differences in hitters, a comparison of Rod Carew 1982 season data is contrasted to Willie Mays 1965 data. For Carew, the 1982 season was rather ordinary and slightly below average given his career numbers; for Mays the 1965 season was his best for home runs (52). Table 4 provides 1982 season hit and walk percentage data for Rod Carew, along with the squares and sum of the squares of the percentage values.

Table 5. Base-state change-event type percentages and squares of percentages for Willie Mays 1965.

Event Type	**Mays, W. 1965**	
	%	**% squared**
Singles	.5706	.3256
Doubles	.1186	.0141
Triples	.0169	.0003
Home Runs	.2938	.0863
Walks	.0971*	.0094
TOTALS	1.10	0.4357

*Walks = BB / (AB + BB)

Similarly, Table 5 provides these same data for Willie Mays from the 1965 baseball season. A comparison of these tables demonstrates how Carew and Mays differ with respect to the sum total of the squares of these percentages.

The RRPI reveals differences between players who were clearly different types of hitters—albeit a unique type of difference in that it accounts for how the hitter actually ‘left’ the system state potential after an at-bat, rather than how he simply might predictively influence said system beforehand (batting average). In these contrasted seasons, Carew and Mays batted .319 and .317 respectively, yet the RRPI between the two reveals two distinct hitting styles—something a retrospective statistic like the RRPI can reveal that a predictive measure, such as batting average, does not. Additionally, if one desires a more composite measure, one could actually take batting average and add the RRPI value to get a value that represents both the probability that the system state will change (e. g. batting average), combined with a historical/retroactive measure of a player’s residual system state potential (RRPI) after an at-bat.

Table 6 goes a bit further by presenting the career base-state change-event data for several Hall-Of-Fame players of different hitting styles, along with their career ARRP and RRPI values.

Table 6. Historical player career RRPI, ARRP, and career totals by base-state change-event type

Player	RRPI	ARRP	BB	Single	Double	Triple	Homer
Mantle, M.	1.488	1.023	1571	1444	339	70	530
Carew, R.	1.657	1.006	874	2404	445	112	92
Mays, W.	1.385	0.949	1227	1960	509	133	641
Williams, T.	1.628	1.173	1108	939	312	30	332
Gwynn, T.	1.561	0.952	587	2378	543	85	135
Brett, G.	1.451	0.972	868	2035	665	137	317

The first two lines of data in Table 6 offer up a particularly interesting RRPI data comparison with respect to Mantle and Carew. With respect to walks vs singles, a walk either: [1] increases the system state potential by 1; or [2] keeps the system state potential constant (e.g. a BB when bases are loaded). Singles may do both of these as well, but when one factors in the probability that as many as 3 runs could score on a single, the walk and the single show great disparity. In an 18 year career, Mantle had more walks than Carew did in his 19 years in pro ball (nearly twice as many walks). But Carew had more singles, more doubles, and more triples than Mantle in a nearly equivalent career, with a nearly equivalent number of plate appearances for these Table 4 events (Mantle = 3954, Carew = 3927). It was Mantle's 'system depleting' 530 home runs that makes the difference in hitting styles so profound. Clearly, these were two different types of hitters—and albeit both left the bases in a nearly equivalent ARRP state career-wise, the addition of the sum of the squares component for hit-type and hitting style is the clear differentiator in the comparison. Simply put, the RRPI attempts to magnify and reveal these differences in hitting style.

Table 7 presents correlations of these Table 6 base-state modifying change-events with the ARRP and RRPI for each player's historical data. These correlations demonstrate the manner in which the application of the mathematical weightings provides a value to each event-type based on how that event-type contributes to maximizing the residual potential of the system state after an at-bat. The Single is the variable that most strongly contributes to the RRPI—with respect to 'hits' this is reasonable—the single (A) is more likely to leave greater run potential within the system state as compared to other hit-events; while (B) simultaneously providing the most flexible range of potential system state change in the form of runs (0 runs, 1 run, or 2 runs may score). Again, although 0 runs scoring as the result of system state change event may be seen as a negative, when it comes to system efficiency, it is a preferred result as it is more desirable to retain high/maximized system state potential.

Table 7. Historical player career total base-state change-event type by ARRP and RRPI correlations

Event-Type	RRPI	ARRP
BB	-.367	.279
Single	-.027	-.837
Double	-.443	-.744
Triple	-.549	-.829
Homer	-.791	-.038

The value of walks (BB) is highly negatively correlated with RRPI—perhaps because the BB lacks the most variability with respect to its influence on the system state. The BB is the most efficient base-state change-event when it comes to consistently changing the state to a higher potential—the BB is always predictable in that the system state either [1] receives an increase in potential by a value of 1; or [2] remains in same state it was in prior to the at-bat (a run is walked in with the bases remaining loaded). As Aforementioned, singles are less predictable in how they will leave the base-state than the BB because sometimes 2 runs will score on a Single; sometimes only 1 run will score on a single; sometimes no runs score on a Single, but for the BB the system will nearly always gain in potential by a value of 1.

Another consideration that must be made is that regarding the 'global potential' of the residual system state vs the 'change potential' of that system state. With men on 1st and

3rd or men on 2nd and 3rd, the global-potential of the system state is at a value of 2.66 or 2.99 respectively, but the change-potential of the base system state between these two scenarios is quite different. In addressing the change-potential of the residual base state system the system is assessed with respect to the degree to which it is 'balanced'—either it is leaning closer toward a possible two-run 'purge' (men on 2nd and 3rd) which would deplete all base state potential, or it is balanced (men on 1st and 3rd; bases loaded) or leaning farther away from a possible two-run 'purge' (men on 1st and 2nd). Thus style of hitting is not just important with respect to reaching base and doing so consistently, it is also important with respect to the type of outs a batter makes and the ability of those outs to leverage the residual system state in a favorable way. The RRPI from the experimental simulation incorporates the full gamut of hitting style, including the influence of the type of outs a batter is inclined to make, in that it recorded all base states after every at-bat, not just base states where a batter was able to reach base.

A Recent MLB RRPI Example

The 2015 Kansas City Royals and their 'keep the line moving' philosophy of hitting (Kerkhoff, 2015) provides a good recent MLB example of the RRPI. The strategy of keeping the batting order moving in a systematic and incrementally stable manner, if successful, should reflect higher average system state potential in the long run. In fact, the 2015 champion Kansas City Royals did indeed demonstrate a higher overall team RRPI (1.428) in 2015 than did the runner-up New York Mets (1.377). Perhaps most interesting though, was the 'style' of hitting that differentiated these two teams when looking at RRPI, as evidenced by their season ARRP values. The Mets had a lower ARRP value (.899) than the Royals (.914)—and the sums of squares for hit and walk ratios for the Royals was .513 vs. the Mets' .477—a disparity endorsing that KC was more likely to engage in 'system-efficient hitting binges' over the course of the season.

A game specific example can be seen in a game between the Kansas City Royals and the Los Angeles Angels on April 26, 2016 (MLB.com). In the bottom of the 5th inning Carlos Perez of the Angels hit a bases-loaded 'bloop' single that resulted in two runs scoring; as a follow up, the next batter, Johnny Giavotella, hit a 3-run home run—which highlighted a 5 RBI sequence between these two at-bats. However, had this scenario been reversed and Perez hit a grand slam—perhaps the most valued hit in baseball—and Giavotella followed with a single, only 4 RBI would have been tallied via the two at-bats. Ultimately, the manner in which the sequence unfolded in real-life was more efficient in that it left the system with a higher residual base-state potential.

A Baseball Simulation Experiment and RRPI

It was hypothesized that a team with a higher computer-simulation RRPI (high efficiency hitting) would show superiority in the long run over a team that demonstrated a lower computer-simulation RRPI (high power hitting). Thus, the most telling support for the RRPI might manifest in an actual 162-game season simulation environment that holds elements such as pitching and batting constant, while pitting two teams with disparate RRPI. Such an experiment was conducted using the 1982 World Series matchup of the St. Louis Cardinals vs the Milwaukee Brewers. This matchup was ideal for experimental inquiry in that the 1982 Brewers hit 216 home runs (an MLB team best that season), as compared to a mere 67 homers by the Cardinals—a lower home run total than Cardinal Mark McGwire put up on his own in 1998 (Posnanski, 2010), and the lowest MLB home run team total for 1982 (Baseball-Reference.com, 2018). In addition, home runs have this same lack of variability in that they always 'empty' the system state and thus find themselves most negatively correlated with high residual system state potential.

Method

In order to hold pitching constant across a 162 game season, both teams faced the same pitcher—Bert Blyleven from 1986—the year Blyleven gave up 50 home runs, the most in MLB history for a single season. Using the record-breaking Blyleven home run season data would seemingly 'weight' the outcome toward the Brewers, the intent of which was to demonstrate that 'hitting efficiency' could overcome 'hitting power' even when the team of greater 'efficiency' was at a distinct disadvantage.

To hold offensive and defensive performance constant, both teams used batting orders and defensive lineups that were comprised of the players who played the most games at each of the non-pitching defensive positions for the year (Baseball Almanac.com). Batting order was set by using the batting lineups for 1982 World Series Game 3 (World Series History, 2007). This lineup utilized a designated hitter (DH) which would maximize the amount of hitting data generated over the entirety of the simulation. There were no substitutions on offense or defense, no pitching changes, and no base stealing, bunting, hit-and-run, or other supplementary strategies applied during the simulation sequence.

Each team had 5 right-hand batters and 4 left-hand batters—batters that were switch-hitters were defaulted to right-hand batting in the team data files (1 Milwaukee batter; 3 St. Louis batters) in order to achieve this 5 to 4 balance. Thus, the Blyleven .268 higher average against right-hand batters would work not work in favor of either team based upon the number of right vs left-hand batters. Milwaukee had a .274 average for all right-hand batters vs right-hand pitching, and St. Louis had a .282 average for all right-hand batters vs right-hand pitching. Milwaukee had a .283 average for all left-hand batters vs right-hand pitchers, and St. Louis had a .286 average for all left-hand batters vs right-hand pitchers.

The simulation was performed using a proprietary software program called Stat-A-Rama Baseball (SARBB) (Howard, 2017) as coded by the author in Visual Basic 6. A special capability was coded into the program that captured and saved the 'base state' values (0 to 3|.99 as per Table 1) after each at-bat, along with individual and aggregate batting, pitching, and win-loss statistics. With respect to pitching fatigue, the SARBB program applied a pitching fatigue algorithm as a 3.2% average reduction in pitching effectiveness per inning for each team, and thus there would be a near 30% reduction in pitching effectiveness for both teams as the game entered the ninth inning.

Simulation team data was contained within Microsoft Excel .csv files. These team data files were compiled from the Lahman database (Lahman, 2015), were user-editable, and were designed to contain player name, team name and year, pitching data, and batting data. A game was 'set up' manually by selecting a ballpark graphic, loading the .csv files for the visitor and home teams respectively, and then assigning the pitcher for each team. After both pitchers and lineups were loaded, the 'options' form on the main interface was utilized to set the 'auto-play' function for the game. The auto-play function allowed the simulation code to continuously loop so as to automatically advance batting orders, record all data results to a Microsoft Access file, and advance innings until game completion.

The actual software simulation algorithm consisted of using pitcher averages against left-hand batters and right-hand batters, as well as batter averages vs right-hand and left-hand pitchers. For an at-bat, the simulator would select a random number from 1 to 6 (e.g. die roll), and assign pitcher statistics to be utilized in the outcome calculations for values 1, 2, 3, and batter statistics for values 4, 5, 6. At this point a value from .1 to .999 was randomly generated and compared to the relevant batting average for the batter, or pitcher-vs-batter average—depending on the value of the 1 to 6 random number—to determine if a hit or walk (e.g. a 'base-acquisition event') had occurred. For example, if the random number from .1 to .999 was equal to or less than the batter's average against a right-hand pitcher, a base-acquisition event had occurred. The next step was determining the probability from the data as to what type of base acquisition event had occurred—walk, single, double, triple, or home run. The final step was to change the 'base-state' by assessing the current state against one of the 8 possible base-states, and then modifying the current state to reflect the new base-state as a result of the base-acquisition event. At-bats that did not result in base-acquisition events were first compared to the strikeout probability of the pitcher/batter, and if a strikeout did not occur, they were deemed 'generic outs' and a similar system (based on historical data—e.g. fly out, groundout, pop out etc.) for selecting the 'type' of out that would occur, was activated in the code.

Procedure

Each game was manually setup and then played out automatically by the software. A single game would complete in about 15-20 seconds, after which the statistics were saved to an MS Access database. Games were played in blocks of 10 until all 162 games were completed and all data had been saved. Base-state data were saved to a comma delimited .csv file, and game statistics were saved to the MS Access database.

Results

The simulation results had 1982 St. Louis with a win/loss record of 85 and 77, and 1982 Milwaukee with the opposite at 77 and 85. ARRP and RRPI values were calculated from the simulation data for both teams. Table 8 provides actual 1982 season ARRP and RRPI team average values for the 18 players used in the simulation (9 per team lineup), as compared to the average ARRP and RRPI values generated by the season replay simulation for these same 18 players.

Table 8. Computer simulation ARRP and RRPI values vs 1982 season values for St. Louis and Milwaukee

Team	1982 ARRP	SIM ARRP	1982 RRPI	SIM RRPI
1982 St. Louis	.990	.924	1.600	1.511
1982 Milwaukee	.932	.881	1.455	1.488

Table 9 provides the individual 1982 ARRP and RRPI values for the Milwaukee Brewers lineups, along with the ARRP and RRPI values as generated from the simulation data. Similarly, Table 10 provides individual 1982 season and experimental simulation ARRP and RRPI values comparisons for the St. Louis Cardinals static lineup.

Table 9. Simulation VS actual ARRP and RRPI for 1982 Milwaukee Brewers

Player	ARRP		RRPI	
	Simulation	1982	Simulation	1982
Molitor, Paul	.843	.885	1.495	1.463
Yount, Robin	.954	.958	1.452	1.377
Cooper, Cecil	.947	.979	1.613	1.454
Oglivie, Ben	.857	.926	1.470	1.384
Thomas, Gorman	.909	.942	1.523	1.333
Howell, Roy	.851	.889	1.484	1.527
Moore, Charlie	.834	.966	1.465	1.534
Gantner, Jim	.828	.922	1.431	1.624
Simmons, Ted	.904	.923	1.459	1.402

Table 10. Simulation VS actual ARRP and RRPI for 1982 St. Louis Cardinals

Player	**ARRP**		**RRPI**	
	Simulation	**1982**	**Simulation**	**1982**
Herr, Tom	.799	.867	1.400	1.578
Oberkfell, Ken	.922	1.00	1.528	1.658
Hernandez, Keith	1.03	.989	1.635	1.582
Hendrick, George	.973	1.01	1.593	1.535
Porter, Darrell	.914	1.05	1.514	1.492
Smith, Lonnie	.979	.930	1.515	1.498
Iorg, Dane	.949	1.05	1.533	1.720
McGee, Willie	.884	.985	1.492	1.652
Smith, Ozzie	.870	1.03	1.389	1.685

Independent samples t-tests were conducted to determine if actual 1982 ARRP player values for the two teams differed significantly from one another, and if simulation ARRP and RRPI for the teams differed significantly. Table 11 presents the results of these independent sample t-test comparisons.

Table 11. Independent samples t-test results of simulation VS actual 1982 ARRP and RRPI values.

Comparison	**St. Louis**	**Milwaukee**	***t***	***p***
Actual 1982 RRPI	1.600	1.455	-3.49	p = .003
Simulation RRPI	1.511	1.488	-0.71	p = .482
Actual 1982 ARRP	.990	.932	-2.57	p = .020
Simulation ARRP	.924	.880	-1.55	p = .138

*α level = .05, two-tailed tests

In addition, paired sample t-tests were conducted to make ARRP and RRPI comparisons of each team's actual 1982 data with its data generated during the simulation (Table 12).

Table 12. Paired samples t-test results of simulation VS actual 1982 ARRP and RRPI values for each team.

Comparison	**1982**	**Simulation**	***t***	***p***
St. Louis RRPI	1.600	1.511	-2.07	.071
St. Louis ARRP	.990	.924	-2.72	.026
Milwaukee RRPI	1.455	1.488	.824	.433
Milwaukee ARRP	.932	.881	-3.84	.004

*α level = .05, two-tailed tests

In order to test for significance in the difference for the simulation of wins, a bias-corrected and accelerated (BCA) 'bootstrapping' (Banjanovic & Osborne, 2016) analysis (alpha = .05) of 2000 iterations was employed in Microsoft Excel on the win/loss data, so as to arrive at 95% confidence intervals which might allude to significant differences. The data were ordered in two columns through 162 rows of data, and using binary coding of a 1 or 0 to represent a win or loss respectively, the two columns were compared. The comparison consisted of creating a column of difference scores by subtracting the first column (Milwaukee) value from the second column value (St. Louis). A mean of the difference scores was then calculated from the difference column to use as the test statistic for the resampling analysis. The results of the resampling can be seen in Table 13.

Table 13. BCA bootstrapping analysis using difference scores from win/loss data

			95% Confidence Interval	
Sample Mean	**Re-Sampling**	**Re-Sampling**	**Lower**	**Upper**
-0.0493	-0.0496	0.0558	-0.1543	0.0617

*Alpha = .05

The 95% confidence interval of the mean difference indicates a value of '0' is indeed a member of the confidence interval. This suggests 'zero difference between groups' as a plausible outcome when comparing these two teams on the win/loss data compiled by the simulation. Thus, the simulation outcome could have just as easily been in favor of Milwaukee. A 95% confidence interval simulation of the individual wins for each team (Table 14.) demonstrates a bit more intuitive outcome. This outcome also supports that Milwaukee could have been a plausible winner of the simulation; here one can easily discern the overlap of the respective confidence intervals.

Table 14. BCA bootstrapping analysis 95% confidence intervals using the win/loss values

MIL1982		St. Louis	
Lower	**Upper**	**Lower**	**Upper**
62	88	70	97

*Alpha = .05

Discussion

The hypothesis that the higher RRPI team would be superior to the lower RRPI team within the simulation was supported. However, St. Louis won the simulation by 8 games, and did so despite the fact that results of the RRPI independent sample t-tests comparing the teams on the simulation data failed to reject the null hypothesis. Such a result seemingly endorses the sensitivity of the RRPI and its influence over an extended period of time—an influence that perhaps manifests despite these RRPI values from both teams being from a common mathematical distribution. However, a plausible alternative explanation for this outcome may be that unlike a real MLB season, each team faced one of the best offensive/defensive teams in their opponent league for an entire 162 game season—there were no 'poor' teams in the simulation that would allow a team to 'ramp up' their ARRP with a greater number of run-bloated victories against a diversity of lower-quality or fatigued pitchers. Also, each team faced a pitcher with a reasonably good ERA and consistent statistics—poor pitchers having difficulties were not a part of the replay simulation. Thus, no variability in pitching rotations during the SIM, and no variability in batting orders during the SIM by holding them constant across the simulation, is validated as a control measure. The resampling analysis however does not appear to support significant differences in the simulation of wins; the indication by confidence intervals from the bootstrapping procedure indicates that there is not a significant difference between the teams as a result of the simulation. It must be pointed out though that when it comes to baseball—and in particular, 7 games of a World Series—one team does not need not be significantly better than the other, they only need to be one game better.

A serendipitous result within the simulation was that the 1982 Cardinals took full advantage of facing the homer-plagued 1986 Blyleven, by hitting more homers than the 1982 Brewers. Thus, it appears that the Cardinals, with their meager home run hitting ability, were underestimated when it came to facing a pitcher with high propensity in giving up home runs. Albeit the intent was to lean the advantage toward the power-laden Brewers, no such advantage was realized despite the power disparity between both teams.

With respect to men on base in real-life scenarios, it is also possible that order effects come into play whereby batters preceded by faster base runners find themselves more likely to face a depleted systems-state—and vice versa—something that could factor in to the earlier-mentioned 'change potential' scenarios.

Another weakness that was not addressed was the cumulative data on hit-by-pitches (HBP). Albeit a small data representation within the big picture, the HBP over a 162 game simulation could have a slight effect on the RRPI calculations as well. The SARBB program was not programmed to account for HBP data and thus it was not collected.

An ironic element outside of ARRP and the subsequent RRPI calculation process is that of the 'strikeout'—and the good vs bad influence it can have on the residual system-state. For example, a strikeout with no men on base is seen as a 'negative' for a batter—he leaves the bases at the same residual system-state potential, yet he's accomplished nothing and even contributed an 'out'. But in some circumstances a batter striking-out is a 'positive'— especially if men are on base and there are less than 2 outs. Although a batter might strikeout with men on base and less than 2 outs, he prevents the double-play and at the same time, guarantees preservation of the residual system-state potential for another batter. Again, this manifests as a weakness for the current study in that the scenarios where such negative vs positive influence strikeouts occurred, was not recorded. The idea that a strikeout could be a positive outcome is rather counterintuitive; for instance, it is an illusion of sorts that with two or fewer outs, a strikeout with the bases loaded should be 'devalued' beyond that of the walk with one man on 1st—the strikeout with the bases loaded leaves the system state with a higher residual run potential for the next batter. [1] With a strikeout and the bases loaded (and two or fewer outs), the batter has contributed another 'opportunity' for his team to drive in runs—thus the post at-bat residual system state potential for the event = 3. [2] With a walk and one man on 1st (with two or fewer outs) the post at-bat residual system state potential for the event = 2. [3] If the next event in each scenario is a home run: a home run with the bases loaded is worth 4—and a home run with 2 men on is only worth 3. [4] The difference between the run-generating-potential of both scenarios is '1' (4 – 3 = 1) which is the same difference one finds between the values of the post at-bat RRP for the system state within each scenario (3 – 2 = 1).

Also of concern is the idea that players might benefit long-term on their RRPI due to simply batting behind players who more frequently get on base. However, at least a portion, if not possibly all, of such situations could be 'cancelled out' by the fact that every batter who leads off an inning for a team, finds themselves in the exact opposite situation—100% of those hitters are 'following nobody', and always arrive in the batter's box with no runners on base (Schell, 2009). In addition, they have a guaranteed low-RRPI-influencing event that results if they make an out. For a visiting team, there would be a minimum of 9 batters in any 9-inning game who would find themselves in such a situation by leading off an inning; for home teams, it could be as few as 8. Thus, the long-term question becomes: Is the number of batters finding themselves in the batting order behind a hitter who speculatively gets on base more frequently, equivalent to the minimum number of possible leadoff hitters in a 9-inning game (9 for visitors; 8 for the home team)?

Ideally, no more than 8 batters could find themselves batting behind a man with a higher on-base-percentage (OBP). E. g. if the leadoff man has a .345 OBP, and one subtracts .015 for each subsequent batter, the last batter OBP is .225—and the leadoff man is following the .225 OBP batter as the only batter following someone with a lower OBP. A true batting order like this may occur on occasion however, the 'cancellation factor' as earlier mentioned occurs with 100% certainty—each team will have a minimum of 8 men coming to bat where they will follow 'nobody' with an OBP of .000. It is also important to note that batting orders keep rotating—thus the scenario of batters following higher OBP batters will magnify by the number of times the order rotates. Four times through the order where 4 men find themselves following a higher OBP batter leads to 16 instances—but one must attend to the fact that an instance only matters if the preceding batter in the scenario reaches base, as failure to do so renders the scenario innocuous.

With respect to the RRPI as a tool of retrospective insight, it does indeed provide a revealing picture of the past. The RRPI across seasons provides an interesting picture of how RRPI has changed across time, and is capable of revealing how the RRPI conjoins with historical events or trends that have occurred. Figure 2 displays an RRPI trend over 15-year intervals, and offers just such aforementioned retrospective insight by highlighting the 2000 Major League Baseball (MLB) season in contrast with four other seasons.

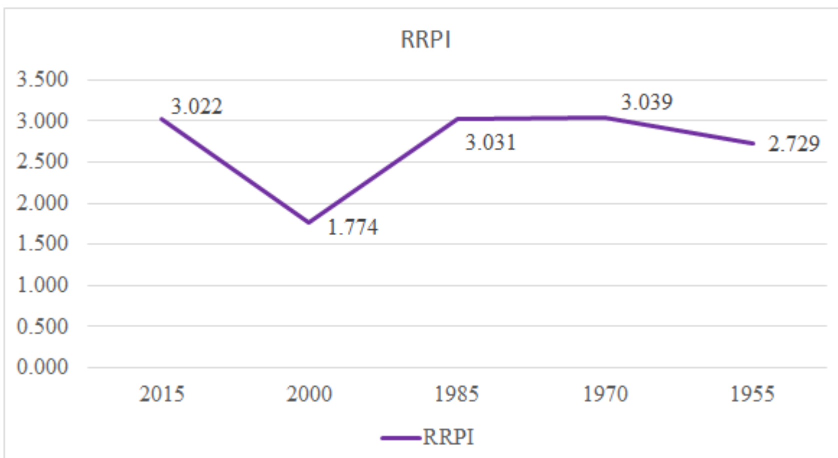

Figure 2. Historical RRPI trend for 5 seasons 1955 - 2015

The 2000 MLB season has the distinction of having the highest home run total for any season in history at 5,693. This is important in the context of the RRPI in that it is the 'home run' which has maximum effect in 'depleting the system' of run-scoring potential resulting in a reset of the system state RRP; home runs always leave the RRP of the bases at a value of zero. With post at-bat RRP and the number of home runs a player hits both contributing to player ARRP across a season, in a season with a monumental number of home runs one should see the exact trend as revealed here in Figure 2—drastically lower league RRPI.

Conclusion

Perhaps the main strength of the RRPI is its application in assessing the residual system state ubiquitously across a number of scenarios. The earlier presented double vs a single comparison provides a good working-model of this diversity of application. At minimum, the RRPI may be able to facilitate insight into the game of baseball whereby those relying on data begin to look at data from a totally different perspective—something that often leads to innovation. And many would agree that it is innovation that quite frequently leads to new found success.

A question that remains is whether or not such insight into the residual system state as applied on a player-by-player basis, could allow one to assemble a more efficient long-term lineup that would maximize residual system-state efficiency. How would one combine performance measures such as batting average, and on base percentage, with RRPI to arrive at a more solidified reflection of player ability? These are questions that can only be answered at the level of the sport itself and those in charge of assembling teams and lineups. Statistical insight such as that provided by the RRPI, is not intended to be an isolated predictor of hitting-style, but rather a supplementary piece of data that assists the dynamic process of offensive player assessment. As an assessment tool, the RRPI has marginal if any value as a forward-projecting predictive statistic that can be used 'in-the-moment' to elevate immediate probability of success in a game scenario. Its true value could be the unique perspective that it provides in assembling a specific 'type' of offensive team, the acquisition of a specific player, or defining the composition of one's own team, or that of an opponent. Ultimately, although it has its weaknesses, the RRPI seeks to provide a novel look at hitting-style from a unique angle, which may be beneficial to those seeking cutting-edge methods for gaining insight into individual player and team-lineup performance.

References

Albert, J. (2002). Hitting with runners in scoring position. *Chance*, 15(4), 8-16.

Banjanovic, E. S., & Osborne, J. W. (2016). Confidence intervals for effect sizes: Applying bootstrap resampling. *Practical Assessment, Research & Evaluation, 21*(5). Available at: http://pareonline.net/getvn.asp?v=21&n=5

Baseball Almanac: 1982 St. Louis Cardinals. Retrieved from: http://www.baseball-almanac.com/teamstats/roster.php?y=1982&t=SLN

Baseball-Reference.com (2010). 1982 MLB Team Statistics. Retrieved January 8, 2018 from: https://www.baseball-reference.com/leagues/MLB/1982.shtml.

Beneventano, P., Berger, P. D., & Weinberg, B. D. (2012). Predicting run production and run prevention in baseball: the impact of Sabermetrics. *International Journal of Business, Humanities, and Technology, 2*(4), 67-75.

Bennett, J. M. and Flueck, J. A. (1983). An evaluation of Major League Baseball offensive performance models. *The American Statistician, 37* (1), 76-82.

Caraher, B. G. (1991). The poetics of baseball: An American domestication of the mathematically sublime. *American Studies, 32*(1), 85-100.

Deviant, Stephanie (2016). Statistics how-to. *Weighting-Factor: Definition, Uses*. Retrieved March 1, 2017 from: http://www.statisticshowto.com/weighting-factor/

Howard, J.N. (2017). Stat-a-rama baseball simulation game. Available at: http;//www.statarama.com

Kedar-Levy, H., & Bar-Eli, M. (2008). The valuation of athletes as risky investments: A theoretical model. *Journal of Sport Management*, *22*(1), 50-81.

Kerkhoff, B. (2015). *For Royals, keeping the line moving starts with Alcides Escobar and Ben Zobrist*. Kansas City Star. Retrieved September 14, 2016 from: http://www.kansascity.com/sports/mlb/kansas-cityroyals/article42119892.html

Krautmann, A. C., & Solow, J. L. (2009). The dynamics of performance over the duration of Major League Baseball long-term contracts. *Journal of Sports Economics*, *10*(1), 6-22.

Lahman, S. (2015). Baseball database. *See http://www.seanlahman.com/baseball-archive/statistics*.

Marchi, M. & Albert, J. (2009). Analyzing baseball data with R. CRC Press. [p.305]

Ng, K. (2017). Analyzing Major League Baseball player's performance based on age and experience. *Journal of Sports Economics & Management, 7*(2), 78-100.

Posnanski, J. (2010). Weakest World Series Winners (since 1946). Retrieved June 3, 2017 from: http://joeposnanski.com/weakest-world-series-winners-since-1946/

Retrosheet (2017). Retrieved March 3, 2017 from: http://www.retrosheet.org/game.htm

Schell, R. (2009). A model for estimating run creation. *By The Numbers: The Newsletter of the SABR Statistical Analysis Committee (19)*, 1.

World Series History (2007). Accessed April 2, 2017: http://www.baseball-cards-and-collectibles.com/1982-World-Series.php