Player Valuation

By Todd Zola with Rob Leibowitz

Introduction

When it comes to fantasy baseball, ***value*** is an extremely schizophrenic word, finding itself at the heart of a myriad of discussions. The relative ***value*** of different players is constantly debated. The ***value*** of a stolen base versus a home run can be discussed for hours. There is earned ***value***, projected ***value***, bid ***value*** and inflated ***value***. What is ***valuable*** to one is of no ***value*** to another.

The on-line Merriam-Webster Dictionary offers a whopping twenty-five definitions for value. In almost every instance, a corresponding application to fantasy baseball can be insinuated. The focus of the next several pages will be this specific definition: *"a numerical quantity that is assigned or is determined by calculation or measurement."*

What will follow is detailed presentation of the theory and mechanics of calculating one's own dollar values. The method will be applicable to leagues of all types, shapes and sizes. It is customizable to meet any category or scoring requirements. It is more logical and mathematically sound than any method published.

The method was revised for 2003,and was initially published in the inaugural 2001 Masters Guide to Fantasy Baseball. What ensues is a streamlined version that addresses many of the concerns and questions our readers have expressed the past couple of years. All of the original concepts are detailed; we have simply taken the liberty of altering the presentation.

Before the procedure is formally presented, a few words of warning must be issued. As important as value is to this hobby, the owner of the team whose players earn the most value during the season, does not necessarily win. The owner whose team earns the most ***POINTS*** wins. There is a distinct and discernable difference. Try not to get hung up on value. The important thing is what you do with the value. That is what separates the winners from the losers. Remember, *second place is first loser*. While it is obviously helpful to have the most accurate values possible, it is more necessary to know what to do with them. Understanding and using the Replacement Level System to calculate your values is only the first step. While you may enjoy an advantage at your draft, it is by no means enough to insure being the recipient of the Yoohoo shower come next fall. Lucky for you, The Masters are here to provide you with all the additional information you need.

Let us go back to the definition of value, *"a numerical quantity that is assigned or is determined by calculation or measurement."* That perfectly describes the manner in which you should consider these values. In essence, the system is a little black box. Statistics are inputted, yielding an output of values. Simply put, statistics are converted to a dollar value.

General Concepts

Baseball statistics are the fuel for player valuation and come in two main forms: actual or projected. Actual statistics are the basis for retro studies, where values are calculated at season's end and are used to aid in strategy studies, a subject discussed in great detail in later chapters. Chances are, the primary use of this chapter will be to calculate customized bid values, so a little extra time will be spent discussing projections.

Projections are the basis for bid values. Note the specific use of the word basis. Projections are not the sole determinant of bid values, but should more properly serve as a guide. If the little black box says John Doe projects to a value of $24, you are not required to say "$24" if the bidding halts at $23. There is also nothing preventing you from saying "$25" if someone beats you to "$24". There are a plethora of reasons why this is so. Leading the pack is the mere fact that there is error associated with projections, so the dollar value the black box renders is truly represented more accurately as a range. The intrinsic dynamics of every auction or draft lead to players being draft above or below their projected values. Inflation, especially in keeper leagues is a prime example. Finally, remember that value is not what wins. Sometimes it may be necessary to bid what it takes to get a player to provide your squad with a better balance of statistics or even to let a bargain go if he provides redundant statistics.

Projection theory is a subject worthy of its own book. For the purposes of this discussion, it is recommended to adhere to a few simple guidelines. In short, make sure your projections have the chance to be correct. The very definition of the word projections connotes uncertainty; else they would be called 'definitisms' or something like that. What is meant by using projections that have the chance to be correct is incorporating a few common sense principles that are really nothing more than logical checks and balances. The cumulative statistics projected for each team should be consistent with previous season's team totals, while of course at the same time accounting for off-season roster movement Cumulative wins should be somewhat reflective of the team's likely placement in the standings and can be approximated by using Bill James Pythagorean theorem, by using each team's runs for and against to generate a wining percentage. Overall leagues statistics do not vary all that greatly from year to year, and the cumulative totals of your projected player pool should fall within these anticipated boundaries. Note the relative consistency in the overall standard rotisserie statistics the past five years:

AL HITTING	HR	RUN	RBI	SB	AVG
2002	2464	10892	10371	1234	0.264
2001	2505	11012	10511	1645	0.267
2000	2687	11995	11416	1297	0.276
1999	2635	11722	11138	1462	0.275
1998	2499	11366	10767	1675	0.271

NL HITTING	HR	RUN	RBI	SB	AVG
2002	2595	11516	10959	1512	0.259
2001	2505	11012	10511	1645	0.267
2000	3005	12976	12317	1626	0.266
1999	2635	11722	11138	1462	0.275
1998	2565	11932	11313	1607	0.262

AL PITCHING	WIN	K	ERA	S	WHIP
2002	1129	14021	4.47	557	1.38
2001	1138	14474	4.48	589	1.39
2000	1143	14032	4.92	551	1.49
1999	1122	13906	4.87	570	1.49
1998	1135	14342	4.66	590	1.43

NL PITCHING	WIN	K	ERA	S	WHIP
2002	1296	17371	4.12	665	1.37
2001	1290	17929	4.36	621	1.37
2000	1285	17323	4.64	627	1.45
1999	1305	17206	4.57	647	1.44
1998	1295	17552	4.24	675	1.38

From one year to the next, the maximum variance is around 10%. It must be noted these are the overall team totals, thus include hitting stats from pitchers and vice versa. One check and balance is to compare the hitting and pitching trend of each league to the previous year, although the advent of inter-league play skews this. For instance, if the projections indicate a rise in offense, this should be reflected in the ERA and WHIP of the pitchers. Truth be told, if the teams at bats and innings pitched are realistic, and the statistics totals fall within the ranges outlined in the tables, then that will suffice. Overall, the total statistics generated by batters and given up by pitchers should match as closely as possible. But, do not quibble over a few missing earned runs, for as will be shown, the values are determined by one's relative position in the player pool, so this slight inconsistency will effectively be flushed out anyway.

Another conceptual element regarding valuation processes is that assumptions within the mechanics of the system itself should be minimized. As much as possible should be rooted in logic and able to be represented numerically. The projections themselves introduce a significant element of uncertainty. There is no reason to compound this uncertainty with unnecessary assumptions within the system.

Finally, to be most useful, a valuation system must be flexible enough to handle any potential set of parameters, but still retain its integrity, no matter what it is asked to do. As this hobby continues to evolve, and unique methods scoring and even playing the game are further developed, the system must be able to evolve with it.

The Six Rules of Valuation

1. Value is always relative.

While the player value will ultimately be presented as a static integer, it is more correctly thought of as a range. The limits of said range are defined by the uncertainty of the projections feeding the system.

2. The amount of money in the fantasy economy is finite.

One of the elegant features of rotisserie style fantasy baseball is that it is a zero sum game. Someone gains a dollar in value, someone else loses a dollar. Someone gains a point in the standings, someone else loses a point.

The amount of money in the economy is dictated by the league's set up. Each team has a budget, so the money in the economy is simply the number of teams multiplied by the team budget.

At all times, there will be a positively valued player pool, that when summed, exactly equals this amount of money in the fantasy economy.

There is nothing that says each roster must be solely composed of players in this positively valued pool. In all likelihood, several rosters will be dotted with players earning a negative value, which means there will be players earning plus value available for pickup. However, by convention, the positively valued player pool must contain sufficient players at each position so that each team could indeed field a legal team of positively valued players.

3. A player's value is based on statistics, not his perceived value or what was paid for him.

This goes back to the schizophrenic nature of the word ***value***. Do not confuse ***value*** with ***valuable***. Keep in mind the working definition: *"a numerical quantity that is assigned or is determined by calculation or measurement."* In these terms, a player of lesser **value** can be more **valuable** to a particular team owner. To the owner leading the home run category by a wide margin, but is within a few stolen bases of getting more points, the lower ***valued*** stolen base specialist can be considered more ***valuable*** than the Triple Crown hitter.

4. A player only has value if he is one someone's roster at some point.

This is the most esoteric rule in the group. Strict application of the rule is unnecessary for no other reason than it is impossible accomplish. There is no way to predict the exact set players that will be drafted, hence only assign values to those players. Furthermore, it is near impossible to predict what players will get hurt or be sent to the farm, and who will replace them. Fortunately, the minute fluctuations in the cumulative statistics introduced by ignoring this rule are trivial. Ultimately, the calculated values are not significantly altered, and are certainly well within the range associated with each value.

Although the following discussion may seem in contradiction to the situation discussed in Rule 2, it really is not if you keep the working definition of value in mind. What Rule 4 assumes is values are calculated as if all of the teams draft players in the positive pool and never make a move all year. When the system is used to compute end of the season values, the assumption is each of the positive valued players spent the entire year on a roster.

5. All statistics in the same category count the same.

This is the valuation equivalent of "a win in April is just as important as a win in September". On the surface, the statement seems more rooted in common sense than it is qualified to be deemed a 'rule'. But for some reason, there are several examples of flawed valuation systems that go against this rule. In a way, those systems try to do too much. Not only do they attempt to assign value, they attempt to apply it as well.

Remember, value is not what makes you win. Rotisserie category points are what makes you win. Value represents potential. Just as is the case in other facets of life, potential can be wasted. In order to achieve a rotisserie point, you need only to have one more statistic in a particular category than the next guy. Let us use home runs as an example. If Team A has 150 home runs, then Team B needs 151 to garner one more point. Say Team B finished with 158. Some will argue that it was home run number 151 for Team B that deserves a greater share of the value, for it gained the point. They will continue to contend that home runs 152-158 are worthless, of no value since they were not enough to earn another point.

Rule 5 states that all 158 home runs are equal and values are computed as such. It is not how much value you have, but what you do with it. Home runs 152-158 have value, that value was simply not manifested into an additional rotisserie point, but remained as potential.

6. A player's value is his usefulness above and beyond replacement.

This, friends, is the defining element of the Replacement Level Valuation System and is thus why we have chosen to name the system as such. In short, why should we pay for something we can get for free? The replacement level represents a body of statistics common to everybody.

Think of it in there terms. Let us pretend we are having a home run derby league where there are 10 sets of three names, and everyone has to pick one name from each set to make up their team. Every single participant elects to choose Barry Bonds. Should you get excited when Bonds hits one into the drink? No, because every single one of the participants in the league just got credited with one home run. This is the same principle as replacement levels.

The System

In a nutshell, each player is awarded value in each of the league's scoring categories. This value is proportional to their contribution to the pool of positive valued players. The contribution is only that above and beyond replacement. The individual category values are summed into a total player value.

Recall that the amount of money in the economy is fixed and defined by the number of teams and the budget per team. In similar fashion, the player pool is defined by the number of teams and the number of roster spots allocated to hitting and pitching. The number of players that are assigned positive value should be exactly equal to the number of teams multiplied by the number of roster spots. Furthermore, the money should be split equally amongst the hitting categories and amongst the pitching categories, unless the league rules state otherwise.

Here we have a major conundrum. In theory, even though the hitting and pitching categories are scored independently, they should equally split the money, as they contribute equal points. However, in practice, more money is allotted to the hitting categories, emanating in what is conventionally referred to as the hitting/pitching split. The problem is, there is nothing specifically defining this split. It is calculated empirically.

Over time, the split settled into a fairly consistent ratio of 65% given to hitting, 35% given to pitching. There are many theories explaining the genesis of the split. And although the bottom line is that it does not matter how it evolved, only that it is applied, it is fun to consider. The explanation that makes the most intuitive sense is that the split is reflective of the difference in roster spots allocated to hitting versus pitching. Standard rules allot 14 hitters as compared to 9 pitchers. This results in a ratio of just about 61%. The inching upward to 65% can be attributed to the theory that hitting is more predictable than pitching and we are more apt to spend money on what is more predictable. Another theory espoused that does not seem as logical has to do with the ratio of counting categories in hitting versus pitching. In the original 4x4 rules, the counting categories would be home runs, RBI, stolen bases, wins and saves. The ratio categories are batting average, earned run average and WHIP. The suggestion, albeit flawed, is that the counting categories do not incorporate negative value while the sum value of the ratio categories must be zero; implying players helping your ratios get positive value while players hurting you ratios earn negative values. As such, only the counting categories should be considered when assigning the split. There are three hitting counting categories as compared to two for pitching, so the ratio is 60%. The reason for the extra 5% is unaccounted for, but again could be related to the perception that hitting is more stable than pitching.

More recently, drafts have trended towards a 70/30 split. Perhaps this is resultant of an amplified perception that pitching is more difficult to project. Here is an alternate explanation. The original 65% was determined when the majority of leagues abided by standard 4x4 rules. The recent proliferation of 5x5 leagues has likely altered the dynamics. One of the rules of thumb of 5x5 leagues is "closers are worth less", as the relative contribution of saves within the pitching categories drops from 25% to 20%. As such, bid prices for closers have dropped. This is s zero-sum economy. The money formally destined for closers must be shunted elsewhere. Instead of keeping it within the pitching categories and increasing the amount paid for starting pitchers, it is probable that the money gets tacked onto the hitting budget, widening the disparity in the split.

All this is well and good and makes for compelling barroom discussion, but is moot when it comes to the actual process of determining values. We have chosen to compute our published projection values using a 67/33 split and will continually monitor the situation, so our values will always reflect what is most useful in any given season.

It is time to get into the nitty-gritty and start doing some calculations. The following represents the amount of value a player earns from a specific category:

PV$(cat) = CAT$ x (PS – RS) / (Pool), where

PV$(cat) is player value in that category

CAT$ is the portion of the economy assigned to that category and is equal to the total money in the economy (number of teams multiplied by the budget per team) multiplied by the appropriate split factor (.67 if this is a hitting category, .33 if it is pitching) divided by the number of scoring categories (4 in 4x4, 5 in 5x5, etc.).

PS is the raw player statistic.

RS is the replacement level statistic. There will be an ensuing discussion focusing on this in a bit.

Pool is the total stats in the pool, accounting for replacement. Basically, it is the sum total of all the (PS-RS) comprising the positively valued pool.

Using real numbers should clarify any confusion. Here are the parameters of a typical league:

12 teams
$260 budget per team
14 hitters
9 pitchers
4x4 scoring

Here are the resultant outcomes intrinsic to valuing players in this league:

There are 168 hitters (12 teams x 14 roster spots) in the positive valued pool.

There are 108 pitchers (12 teams x 9 roster spots) in the positive valued pool.

There is a total of $2090.40 allocated to the 168 hitters (12 teams x $260 per team x .67) and a total of $1029.60 allocated to the 108 pitchers (12 teams x $260 per team x .33).

Each hitting category is allotted $522.60 ($2090.40 / 4 categories) and each pitching category is allotted $257.40 ($522.60 / 4 categories).

Relating to the equation just presented, the Cat$ is $522.60 for a hitting category and $257.40 for a pitching category.

By means of example, let us make up some realistic numbers and calculate the HR$ for a player who hit 35 HR. Methods revealed later set a replacement value to 4 HR and the cumulative, unadjusted total of HR by the 168 hitters in the positive pool is made to be 2200.

(PS – RS) = 35 – 4 = 31
Pool = (2200 – (168 x 4)) = 1528
PV$(HR) = $522.60 x 31 / 1528 = $10.60

Remember when it was suggested that players cannot earn negative value in a counting category and that was deemed flawed? Let us try another example. Same league, same parameters, except this time we are calculating the WIN$ for a pitcher with 2 wins, with a replacement level of 4 and an unadjusted pool of 1000 wins for the top 108 pitchers:

(PS – RS) = 2 – 4 = -2
Pool = (1000 – (108 x 4) = 568
PV$(WIN) = $257.40 x (-2) / 568 = -$.91

This pitcher lost almost a dollar in the wins category. It makes intuitive sense when you consider that everyone in theory is picking up 48 “free wins”, and this team is already down two wins. Just to get even, another pitcher needs to win 6 games. Doing the math, 6 – 4 is 2 and plugging 2 into the above gives this pitcher a value of $.91, thereby balancing out the pitcher that won 2 games.

Performing the above operation for each scoring category and summing the individual values yields the final player value.

(continued…)

Ratio Stat to Counting Stat Conversion

Ratio categories present an interesting challenge to the valuation processes. Although these examples will use batting average, a similar description can be offered using ERA and WHIP, or any ratios stat for that matter. A .300 hitter who goes 130 for 500 should be more valuable than one who goes 3 for 10. Conversely, a .250 hitter that goes 100 for 400 should be less valuable than one who goes 3 for 12.

There are two popular published methods that address this issue. The first is championed by Art McGee and is based on principles first put forth by Alex Patton. The hits and at bats of an average team in a league are determined. It is assumed the player in question fills the last roster spot, so those average hits and at bats are prorated to one less roster spot. The hits and at bats of the player being valued are added in and the new average is computed. The counting stat is the difference between the adjusted average and the league average. The second is presented by John Benson. A baseline average is subtracted from the player average, and the player's number of at bats multiplies this number. So in the example given earlier, although both players hit .300, the resultant difference after subtracting the baseline average is multiplied by 500 in one case, and 10 in the other. The extra at bats render a higher converted counting stat. Benson recommends using the league median average as the baseline. This is not the baseline of the roto league, but of the entire player pool, positive and negative included. He cites years of empirical evidence that the league median average approximates the batting average of a typical last place roto team, and further hypothesizes this is the logical baseline, as players above it will raise their team batting average above the last place team's, earning value.

While this is mathematically logical, it was presented with no other theoretical basis. Why multiply by the number of at bats, other than it makes the difference in averages bigger and is a handy, available number? Enter John Mosey. Mosey has introduced the concept of "extras". The notion is that you multiply the baseline batting average by the at bats of the player in question. This leaves the amount of hits that the baseline player would have based on that amount of at bats. The counting stat, or "extra" is the amount of hits the player in question totaled minus this baseline hit amount. The funny thing is, when you write this formula out and algebraically manipulate it, it turns out to be identical to the Benson method. Mosey suggests using the entire pool's mean batting average as the approximation to the last place team. Mosey's theoretical basis for extras makes this method both logically as well as mathematically acceptable.

Reader feedback has expressed dissatisfaction with setting of the baseline average to the anticipated average of the last place team. This has prompted an investigation of the ideal baseline average. The study is ongoing, but preliminary results appear to hint that using the batting average of all the players in the positive player pool is better to use than that of the last place team. The difference, however, is statistically insignificant. That is, it does not matter which baseline is employed as they both lead to the same result.

When Benson and Mosey suggested using the median and mean of the entire player pool to represent the average of the last place team, they had deep leagues in mind, using either AL or NL only. This presents a problem for leagues of other formats, especially the shallower mixed leagues.

This is sufficient reason to recommend choosing the baseline to be the average of the positive player pool. This can easily be estimated by ranking the player pool by at bats and determining the batting average of the top 168 players. The number 168 refers to the example used previously. It should be customized according to the specific league's parameters.

Thus, the formula for calculating extra hits is

ExH = PH – (Baseline BA x PAB), where

ExH is extra hits
PH is player hits
PAB is player at bats
Baseline BA the average of the positive pool

To determine the extras for the standard pitching categories of ERA and WHIP, the following equations are used:

ExER = (Baseline ERA x PIP / 9) – PER, where

ExER is extra earned runs
PIP is player innings pitched
PER is player earned runs
Baseline ERA is that of the last place team

ExWH = (Baseline WHIP x PIP) – PWH, where

ExWH is extra hits plus walks
PWH is player hits plus walks
Baseline WHIP is that of the last place team

Note the fact that the baseline element of the equation is subtracted from the player's contribution in the batting average conversion, while this is reversed for the pitching operation, due to a better ERA and WHIP being a lower ERA and WHIP.

The ERA and WHIP of the positive player pool is again employed as the baseline. To estimate the composition of the pool, the players are ranked in order of twice saves plus wins. The example from before would lead to taking the ERA and WHIP of the first 108 pitchers ranked by twice saves plus WHIP.

<u>Replacement Values</u>

As mentioned earlier, the use of replacement values is the defining element of the system, and thus deserves special attention and explanation. The intent is to convince everyone that the replacement values are more than a fudge factor, as opponents of the system contend. They will be given a mathematical, practical and theoretical definition. It is this theoretical definition that will serve to distinguish them from fudge factors, especially after numerical limits are imposed upon them.

Replacement values have thus been defined as representing those statistics available for free. To take this a step further, we must fast forward and explain a little about what happens on the spreadsheet.

In brief, the spreadsheet is automated to conduct all the calculations described earlier. The raw data is the player statistics.

1. The statistics are adjusted after replacement.

2. The adjusted pool totals of the individual categories are determined.

3. The player's value per category is computed.

4. These individual category values are summed into a total value.

5. The player values are sorted in descending order.

Rewind back to replacement value theory. The missing element is the numerical representation of replacement values. While the following definition may appear to be non-exact and border on violating the "no assumptions" provision discussed earlier, the definition will be refined to yield a more exact figure.

The replacement value is the sum average of the highest players in the non-positive pool after the descending sort mentioned above. The subjective part is how many players should be included in the replacement set. Studies are underway to determine the optimal number. Preliminary results suggest the number should be about 10% of the total number in the positive pool, perhaps a mite less. By means of example, earlier we identified a pitcher pool of 108 positively valued pitchers. The replacement pool should consist of about 11 pitchers, meaning the replacement set is the average of the statistics of those 11 pitchers. So using our pitching example, the replacement pool can be composed of pitchers 109 through 119. The average amount of wins, saves, etc. from these 11 pitchers are subtracted from every pitcher.

The above is a numerical representation of the replacement value set. We need to refine it. Recall that the number of players in the positive pool is bound by the number of teams and the number of roster spots per team. One of the functions of the replacement value will be to affect this boundary. That is, the replacement values will be adjusted (within soon to be defined limits) to set the pool at the precise number of players. The highest ranked non-positive player will be valued at $0. Please note, part of this is NOT that the first positive player necessarily be valued at $1. While this is most often the case, it is not a rule, or even convention of valuation. A player earns what a player earns. Obviously, the bid price of the last players drafted at each position will be $1. This does not mean the valuation system MUST value players at each position at $1, although, as stated, it most always happens. It only means that someone might be the recipient of a $1 bargain in the end game.

A phrase was just used that opens a can of worms. It was just revealed that the replacement values will be adjusted. At this point, two new terms must be introduced. The non-adjusted replacement value will heretofore be referred to as the 'theoretical replacement value', or TRV. The adjusted replacement value will be labeled 'actual replacement value', or ARV.

Fast forward once again to the value computation on the spreadsheet. The key to the entire operation will be repeated sorts of the data to establish the positive pool. Before the first sort is done, the ARV is by definition the TRV. As will be explained, the TRV are presently represented by a spreadsheet formula. . The data is now sorted. Perhaps the data was fortuitously set up in order of value, thus there is no shifting of players. Chances are, the order is shuffled; knocking some players out of the positive pool and altering the replacement set which in turn changes the TRV. The data now needs to be resorted. This procedure is repeated until one of two situations arises. The data either stabilizes so that repeated sorts lead to no change or it enters a repeating loop. Sometimes the loop takes as many as twelve sorts to cycle back to the beginning.

In order for the data to be useful, it must reach a stable state upon consecutive sorts. To get out of a loop, adjustment of the ARV is necessary. Recall the replacement set was defined by the players as ordered before the initial sort. This order has now changed. Before adjustment, the replacement values are in fact the original TRV. The TRV change as the pool of replacement players changes. Ultimately, we are going to require that the ARV be within a defined limit of the TRV. It is this operation that eliminates the notion that replacement value are little more than fudge factors. The term fudge factor connotes they can be selected to force the data into a state deemed acceptable. By imposing a limit defined by the TRV, the ARV are quite real.

Back to the repeating loop: one or more of the replacement values is adjusted and the data is resorted until it stabilizes or enters another loop. Adjustments are made to the ARV until the data indeed stabilizes.

At this point, remember an important function of the replacement values, namely to set the size of the player pool. Once the data stabilizes, it must be determined if the pool is of the proper size. If not, the data must be readjusted. It is sorted and adjustments are made until it is stable and the size of the pool is again determined. This process is repeated until the data is stable and the pool is of defined size. If the player pool has too many positive players, the correction is to raise one or more of the ARV. Too few players require lowering one or more.

All this time, the adjustments are made keeping the ARV within reasonable limits of the TRV. I suppose you want to know what is meant by a reasonable limit. This is another study in progress and will likely be a centerpiece for next year's Annual. A sufficient working number is between 5% and 10%. Preliminary results are showing that each category will have its own limit. By means of example, a 10% adjustment to the replacement level in home runs leads to a larger change in value than a 10% change in the replacement level for stolen bases.

Let us rewind again back to replacement value theory and review. The replacement pool of players is represented by a defined number of non-positively valued players, but may include players at the tail end of the positively valued pool. The set of replacement values must lead to a stable set of data upon repeated sorts, lead to a player pool of the proper size and be within reasonable limits of the average of the statistics defining the replacement player pool.

While it is true that there are infinite sets of replacement values that will satisfy the first two criteria, if all three are followed, the argument that replacement values are little more than fudge factors has little, if any merit. In the future, the determination of the optimal replacement value set will be further refined.

REPLACEMENT VALUES AND POSITION SCARCITY

When presenting the Rules of Valuation, an important point was made but has not been sufficiently explained. Specifically, "the positively valued player pool must contain sufficient players at each position so that each team could indeed field a legal team of positively valued players."

What happens if values are calculated without consideration of the above statement? The most obvious result is there is a paucity of catchers.

That is, in a 12-team league that requires the inclusion of two catchers, there are not 24 that naturally come out with positive value. The number is closer to 18. This is an oversight many people ignore, but if the rules dictate that each team must have two catchers, a proper value system should have an ample supply of catchers. This situation is commonly referred to as position scarcity.

The way rectify this catcher scarcity is to dedicate a set of replacement values just for the catcher pool. Recalling the league parameters we set up earlier, there were 168 hitters. So the catcher replacement levels are chosen so that there are 24 catchers with positive value and the non-catchers replacement level is chosen so that there are exactly 144 players with positive value.

Let us take a look at how this change affects things mathematically. Earlier we derived the following three equations for a player in a pretend league:

(PS – RS) = 35 – 4 = 31
Pool = (2200 – (168 x 4)) = 1528
PV$(HR) = $522.60 x 31 / 1528 = $10.60

Everything remains the same, except that the non-catcher replacement level would naturally shift to around 4.5 while catcher replacement level for HR ends up around 2. This makes sense, as there were likely catchers in the original replacement which are no longer there and there low HR total dragged the average down. The total is now broken into two components.

Pool = 2200 – (144 x 4.5) - (24 x 2) = 1504

The PV$(HR) for a non-catcher hitting 35 HR is

$522.60 x 30.5 / 1504 = $10.60, same as above.

The PV$(HR) for a catcher hitting 35 is

$522.60 x 33 / 1504 = $11.47, a marked gain.

This is an ideal occasion to point out a little quirk with regards to player valuation. Very few leagues with a cap like $260 allow bids in increments of anything other than a whole dollar, thus values are expressed as such. If the above example were for a home run derby auction league, the value of the non-catcher would be rounds up and expressed as $11. The value of the catcher would be rounded down and also be expressed as $11. The problem is, the catcher is really worth almost a buck more than the non-catcher, $.87 to be exact.

Based on numbers from last year, it can be expected that presently catcher values will incur an increase of $2 or $3 when a separate set of replacement values are used. Keeping in mind this is a zero-sum economy, others players must lose value. Assuming there are 24 catchers at an average increase of $2.50, there is $60 of value must be accounted for. This is distributed amongst the 144 non-catchers, each of whom loses $60 / 144 or $.41. Depending on the round-off phenomena just demonstrated, several of these non-catchers will not appear to lose any value, while others will drop a buck.

In 12-team mixed leagues, the final value of catchers is presently increased by $7 to $9 using a dual replacement set system. Here, the catcher pool assumes a total increase of, on the average, $8 times 24 or $192. Dividing this up between the remaining 144 players and it is clear each drops a buck, maybe two on round-off.

Note the inclusion of the qualifier ‘presently’ in the above examples. The described increases are not a hard and fast rule, but more a rule of thumb based on the present major league player pool. The distribution changes from year to year and should be investigated when assigning player values.

What about the other so called scarce positions like middle infielders? In both deep and shallow leagues, mainly due to a rising number of players with multiple position eligibility, there are ample players with positive value to fill out legal rosters for everyone. So according to this application of scarcity, the only position that one needs to worry about is catcher. There is no scarcity at any other position.

Please realize some define scarcity in other ways, relating to the quality of the player pool at the individual positions. This is not a function of valuation so it need not be incorporated into the calculations. It falls under the mantra of “what you do with the values”.

The above explanation of scarcity has been a point of contention regarding this system, as some contend that even though the positively valued player pool is composed of a sufficient distribution of players, each position should be compared against replacement level players at their respective positions. You know what? They are right. But here's the catch: doing this does not significantly alter the values. What happens is there are positions that favor speed and positions that favor power. Those that favor speed have a higher replacement value for stolen bases and a lower replacement value for HR and RBI. The opposite is true for the positions that favor power. The end result is these effects cancel each other out. The individual category contributions are indeed changed, but the sum total for the majority of the player pool does not change, and if it does, it is by a buck in deep leagues, maybe two in shallow ones. Recalling that values are ranges, it may not be necessary to devote the extra time necessary to calculate the values with individual replacement sets for each position. Your bids will not be altered.

An exception is when retro dollar values are calculated and individual category values are needed for strategy studies. Here it is worthwhile to utilize different replacement values for each position, lumping DH only players in with the outfielders. The process is the exact same as described for using two replacement values in terms of dividing up the total of the stats in the player pool. You will just have as many as six constituents encompassing this pool.

While on the subject of multiple replacement values, a few other points should be considered. A hierarchy should be set up to assign players with multiple position eligibility. Presently, this hierarchy is C>SS>2B>3B>1B>OF/DH. The positions of SS and 2B are interchangeable, as they have very similar replacement set. They can even be combined into a single position of middle infielder is desired. Presently, the difference in quantity of positively valued 1B is sufficiently different from 3B to warrant each having individual replacement values. Actually, 3B is becoming a talent deficient position and is on the verge of being on a par with the middle infielders in terms of the quality and quantity of positively values players.

According to the definition of scarcity with respect to valuation and having ample players of each position in the pool, the best manner to determine the number of players making up each positional pool is to initially run the values with a single replacement level. Isolate the positive pool and sort it by position. Count up the number of players at each position. In standard 12 team leagues, if there is no scarcity, there will be minimum 36 total middle infielders and 36 corner infielders, with at least 12 at the respective positions. There should be 24 catchers. It is OK if there is an excess of corner or middle infielders as they will fill the utility spot. You now have a basis for setting the size of each constituent pool. By means of example, the first step is forcing 24 catchers into the pool, so the replacement set for catchers may average the statistics of catcher #25 - #28, valuing catcher #25 at $0. Say there are 17 3B and 23 1B in the pool. Use 3B #18 - #21 for the replacement set and value 3B #18 at $0. Similarly, use 1B # 24 - #28 as replacement and make 1B # 24 be worth $0. Now say there are 17 2B and 19 SS in the pool. The replacement set and the first $0 player are established in an analogous manner. This leaves us with 68 players at OF/DH whose replacement set is represented by OF/DH #69 - #75 with the first $0 OF/DH being #69. Note the sum total of the six positions necessarily equals 168. Each of these positions has their statistics adjusted according to their corresponding replacement set, and it is these adjusted stats that comprise the total player pool. Also note that as the individual pools were reduced in size, we strayed away from making the replacement pool being 10% of the size of the positive pool. For the most part this would mean that the replacement set consisted of only two players and that is not enough, the minimum should be four to insure outliers do not overwhelm the data. The actual number of players used is subjective, the above numbers are realistic but can be tweaked to taste until the optimal number is formally determined.

In summary, values can be calculated with a single replacement value to determine the distribution of players within the pool. This distribution changes from year to year. Individual positional replacement values can be used to regulate the size of each positional pool. In practice, using a dual replacement system including a replacement set for catchers and one for all non-catchers will suffice for the majority of applications.

The Spreadsheet

Understanding the principles of valuation theory is well and good, but it is useless if it cannot be implemented into a spreadsheet for automated calculations. What will follow is a series of guidelines and tips designed to ease the conversion from the abstract to the concrete, that is, from theory to actual values.

Everyone has his or her own level of familiarity with spreadsheets. This presentation is by no means the only way to set up the spreadsheet. The suggestion is to follow along, perhaps even parroting the directions onto a real spreadsheet. Once a comfort level is reached with the system, feel free to incorporate alterations, as you see fit.

Understanding that the beauty of this system is its flexibility and adaptability, to avoid confusion, all the guidelines and tips will be based upon a sample league with the following parameters:

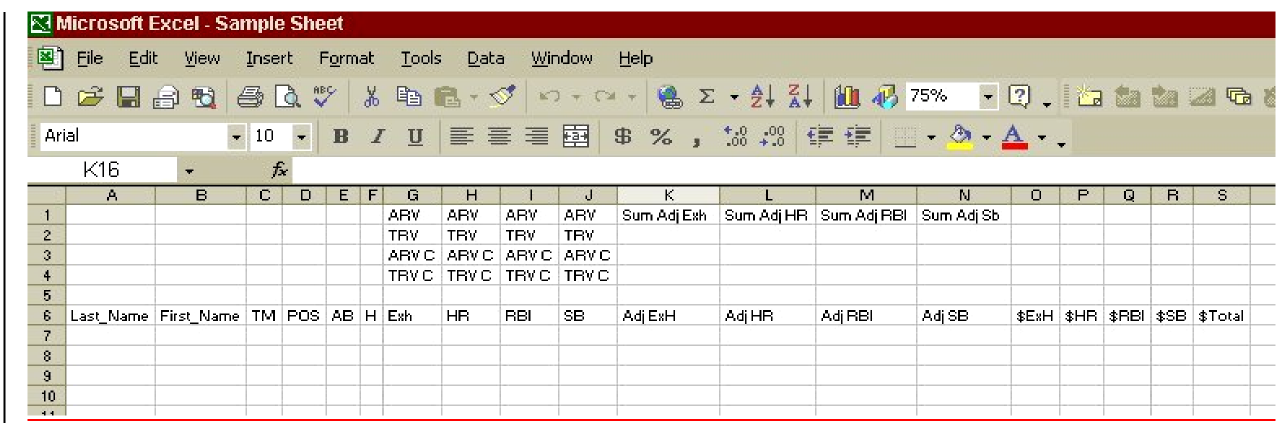
12-teams
4x4 scoring
14 hitters
$260 cap per team.

The data will include the following elements:

Last Name, First Name, Team, Position, ABs, Hits, HRs, RBIs, and SBs.

A space will be left between hits and HR so that the extra hits can be calculated. A baseline average of .270 will be used to generate extra hits. A baseline ERA of 4.10 will be used to calculate extra runs and a baseline WHIP of 1.33 will be used determine the extra hits plus walks. These baselines are representative a 12-team AL only. They should be customized just like the rest of the parameters.

On the following page is a snapshot of a typical spreadsheet. The stepwise procedure for calculating values with a single replacement set will be initially described. At the conclusion, the necessary adjustments to incorporate multiple replacement sets will be explained, as well as a few tips for setting up a pitching spreadsheet.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1							ARV	ARV	ARV	ARV	Sum Adj Exh	Sum Adj HR	Sum Adj RBI	Sum Adj Sb					
2							TRV	TRV	TRV	TRV									
3							ARV C	ARV C	ARV C	ARV C									
4							TRV C	TRV C	TRV C	TRV C									
5																			
6	Last_Name	First_Name	TM	POS	AB	H	Exh	HR	RBI	SB	Adj ExH	Adj HR	Adj RBI	Adj SB	$ExH	$HR	$RBI	$SB	$Total
7																			
8																			
9																			
10																			

Step 1. Setting up the spreadsheet

In row 5, set the headings as shown in the example above, except for the column marked ExH. Then type in or copy and paste your data into the spreadsheet, starting at A7 with Player name and going across. Then insert a column between H and HR for Extra Hits (ExH).

Tip: You may also wish to consider clicking on row 7 so that the entire row is selected and using the "freeze panes" command under your windows menu. This will allow you to scroll up and down the data while being able to see all of the headings at the same time.

Step 2. Converting Ratio Stats to Counting Stats

To calculate extra hits, enter in =F7-.270*E7 or "Hits-.270*ABs". Copy and paste this formula down the column for each hitter in the player pool. For pitchers, you will need to insert two columns; one column for extra earned runs and one column for extra walks plus hits. On your pitcher spreadsheet you would enter
= (4.10 * IP / 9) - ER into the ExER column. You would then enter
= (1.33 * IP) - (Walks + Hits) into your ExWH column. And then you would copy both formulas downwards for every pitcher in your pool

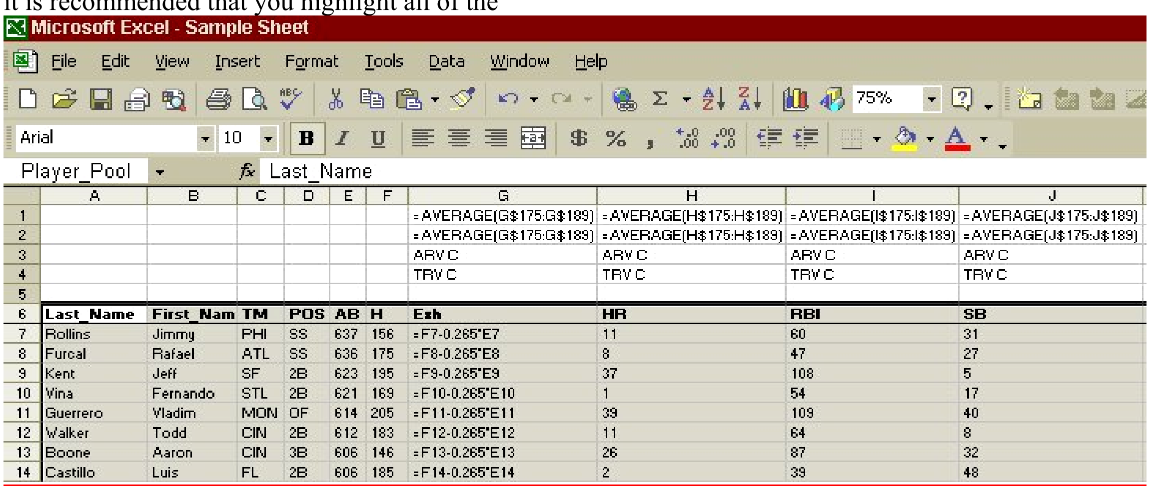
Step 3. Ordering the Data

The order of the data before the first sort is crucial. Something that fairly well approximates the eventual pool is needed, or else the initial set of replacement values severely skews the data. If the data is not ordered correctly, players who should have positive value could up at the end of the pool with a negative value. This is especially critical with pitchers, as closers are particularly susceptible to this phenomenon. In order to prevent this from happening, you need to rank the hitters in order of descending At-bats. For pitchers, you will need to insert an additional column and insert the formula =(2 x saves) + wins. Then sort all of the pitchers in descending order on this column.

Tip: As you will be sorting the data set on multiple occasions during the valuation process, it is recommended that you highlight all of the data within the set, including the column headings. Then go to the insert menu, select name, and then define. Insert a name of your choice, such as "Player Pool". From then on you use the drop down arrow in the "Name Box" (located typically just below where the Font Type is listed), and select "Player Name". It will then automatically highlight the data you need to sort and will save valuable time, freeing you from have to manually highlight all of the data each time you need to sort it.

Step 4. Calculating Replacement
For this exercise, our sample league contains 168 players occupying rows 7 through 174. The replacement pool will consist of 15 players, using the hitters listed in cells 175 to 189 (or technically batters 169 through 183)

The cells in row 1, our row for Adjusted Replacement Value (ARV), are reserved to hold the replacement set. In G1, the cell that will hold the replacement set for Extra Hits, (note that it corresponds to the headings below



	A	B	C	D	E	F	G	H	I	J
1							=AVERAGE(G$175:G$189)	=AVERAGE(H$175:H$189)	=AVERAGE(I$175:I$189)	=AVERAGE(J$175:J$189)
2							=AVERAGE(G$175:G$189)	=AVERAGE(H$175:H$189)	=AVERAGE(I$175:I$189)	=AVERAGE(J$175:J$189)
3							ARV C	ARV C	ARV C	ARV C
4							TRV C	TRV C	TRV C	TRV C
5										
6	Last_Name	First_Nam	TM	POS	AB	H	Exh	HR	RBI	SB
7	Rollins	Jimmy	PHI	SS	637	156	=F7-0.265*E7	11	60	31
8	Furcal	Rafael	ATL	SS	636	175	=F8-0.265*E8	8	47	27
9	Kent	Jeff	SF	2B	623	195	=F9-0.265*E9	37	108	5
10	Vina	Fernando	STL	2B	621	169	=F10-0.265*E10	1	54	17
11	Guerrero	Vladim	MON	OF	614	205	=F11-0.265*E11	39	109	40
12	Walker	Todd	CIN	2B	612	183	=F12-0.265*E12	11	64	8
13	Boone	Aaron	CIN	3B	606	146	=F13-0.265*E13	26	87	32
14	Castillo	Luis	FL	2B	606	185	=F14-0.265*E14	2	39	48

that you typed in previously) type in the formula: =average(G$175:G$189). Then take this formula and copy it across for your other

categories. For our example, that would be HRs, RBIs, and SBs.
Row 2 is held for our Theoretical Replacement Value (TRV). Simply copy the contents of the 4 cells in ARV directly below to row 2 for the TRV.

(Note that the formulas in the samples are visible as a guide. When you type them in, the resulting figures will instead be visible.)

Step 5. Adjusting the Raw Data with the Replacement Set

The replacement value for each stat must now be subtracted from the raw data. In K7, input the formula =G7-G$1. G7 is the raw extra hit value and G1 is the corresponding replacement value. Next, copy this formula across to column N for all of the statistical categories and then downwards for all of the players.

K	L	M	N
Sum Adj ExH	Sum Adj HR	Sum Adj RBI	Sum Adj Sb
Adj ExH	**Adj HR**	**Adj RBI**	**Adj SB**
=G7-G$1	=H7-H$1	=I7-I$1	=J7-J$1
=G8-G$1	=H8-H$1	=I8-I$1	=J8-J$1
=G9-G$1	=H9-H$1	=I9-I$1	=J9-J$1
=G10-G$1	=H10-H$1	=I10-I$1	=J10-J$1

Step 6. Summing the Adjusted Statistics

Recall that the sum of the adjusted stats was defined as "the total of raw stats in the positive pool minus the product of the number of players in the pool, multiplied by the replacement value for that statistic". While it is possible to input that formula using the raw category stats, it is more efficient to just sum up the adjusted stats in columns K through N.

The pool size is the 168 players discussed earlier. To generate the sum of their adjusted stats, input =sum(K7:K174) into K1. Copy this formula to the right for the adjusted HRs, RBIs, and SBs through to N1.

K	L	M	N
=SUM(K7:K174)	=SUM(L7:L174)	=SUM(M7:M174)	=SUM(N7:N174)
Adj ExH	**Adj HR**	**Adj RBI**	**Adj SB**
=G7-G$1	=H7-H$1	=I7-I$1	=J7-J$1
=G8-G$1	=H8-H$1	=I8-I$1	=J8-J$1

Step 7. Calculating the Category Dollars

The amount of money allocated per category must be determined. For the purpose of this exercise, the hitting to pitching split will be set at 67:33. Just to recap, our sample league has 12 teams, each with a $260 budget and 4x4 scoring categories. Therefore the total money per category for hitters is equal to $522.60 (12 x $260 x .67)/4.

The player's portion of this total is his category value. Input =(K7/K$1)*522.60 into cell O7. K7 is the adjusted extra hits and K1 is the sum of the adjusted hits for the top 168 hitters. Next copy this formula across to R7 and down for all the hitters.

O	P	Q	R
$ExH	**$HR**	**$RBI**	**$SB**
=(K278/K$1)*522.6	=(L278/L$1)*522.6	=(M278/M$1)*522.6	=(N278/N$1)*522.6
=(K279/K$1)*522.6	=(L279/L$1)*522.6	=(M279/M$1)*522.6	=(N279/N$1)*522.6
=(K280/K$1)*522.6	=(L280/L$1)*522.6	=(M280/M$1)*522.6	=(N280/N$1)*522.6
=(K281/K$1)*522.6	=(L281/L$1)*522.6	=(M281/M$1)*522.6	=(N281/N$1)*522.6

Step 8. Calculating the Player's Total Value

The player's total value is the sum of his category values. Input =sum(O7:R7) into S6 to give you that value and then copy that formula downwards for the entire pool. You should also at this time set the format of this column to a $ format with no following zero's.

9. Sorting the Data Until the Values Are Correct

This is the key to the entire operation and admittedly takes patience. Over time, the process will become easier and easier with practice. The end results will be the most accurate set of values possible, which is definitely worth the wait.

The first step is to dedicate a pair of cells to the total value of the last player in the pool (player 168) and the total value of the first player in the replacement pool (player 169). This will serve as a guide for the repeating sorts. The process is finished when the ARV and the TRV are stable and within the specified range while the first player to be valued at $0 is the first replacement player.

Input =S174 into cell S1 and =S175 into S2. S1 is player 168 and he will have a total value of $1 when we are done. S2 is player 169 and he will have a total value of $0 when we are done.

Next, copy the data ranging from G1 to J1 (your ARV). Right click on G1 and select "Paste Special". Select the check box with "Values" next to it and press "OK". Note you are copying over the original formula and simply putting in the values that were generated from them. You will use this in relation to the TRV in which the formulas are left alone.

Now, using the name box, go to "Player_Pool" so that every player starting with their name through their $Total is highlighted. Next, go to the Data Menu, select "Sort" and sort by $Total, descending. Press OK. Keep repeating the sort until the data enters a repeating loop or becomes stable. Watch S1 and S2 as well as the relationship between ARV and TRV to see this happen.

If both S1 and S2 have values of $1 or more in them, it means you have too many positive players in the player pool. In order to cut down the number of positive players, you will need to **increase** the ARVs, but keeping it with +5 to +10% of the TRV as described earlier in this article.

Similarly, if S1 and S2 are listed at $0 or less, it means you have too few positive players in your player pool. You will therefore need to **reduce** your ARVs, but keeping them within -5 to -10% of the TRV.

At this point you are done with your valuation for pitchers. You can also choose to finish here for hitters if you do not wish to use multiple replacement sets, but it is recommended that you do continue on, especially if you wish to have all of your catchers to with a positive value.

10. Values For Multiple Replacement Sets

The first step in using multiple sets of replacement value is to run the data with the single set as just described. This is used as a guide to set the size of each component pool. Using our model league, that would entail assuring that 24 catchers, 36 middle infielders with a minimum 12 each at 2B and SS, 36 corner infielders with a minimum 12 each at 1B and 3B, and 60 outfielders comprise the positive pool. In all likelihood, the catcher pool will ultimately be exactly 24 catchers with the necessary 12 utility players spilling over from the other positions.

In order to count the number of players per position in the positive pool, first copy the $Total column and use the paste special command to copy just the values, not the formulas, into the column immediately to the right of it. You can then sort all of the players by position and by your second $Total column, in descending value.

If a dual replacement system is to be employed with a separate set for catchers and non-catchers, rearrange the data so the pool of catchers is at the end of the pool. If more than two sets are to be used, it does not matter what order of the individual positional player pools are.

The first adjustment that needs to be done with multiple sets is that there needs to be space for all the necessary ARV and TRV. The model league is designed for the dual replacement system. (Where ARV C and TRV C are listed on the sample images is where you would place your replacement figures for catchers). If more replacement sets are being used, then you will have to devote more space for them and insert more rows as a result.

The next change is with the designation of the players within each replacement pool. What will be described is the method for using dual replacements, but you can extrapolate it to employ as many sets as desired.

Our sample league requires there be 24 catchers and 144 non-catchers in the positive pool. Strictly by means of example, let us say there are 260 non-catchers and 40 catchers in the entire pool. This would place the non-catchers in rows 7 through 266 and catchers in row 267 through 306.

The calculations of the replacement values for the non-catchers entails setting a new replacement set, perhaps using players 145 through 154, corresponding to cells 151 through 160. So now the formula is =average(G$151:G$160). This is still situated in cells G1 through J1, so no change is necessary when adjusting the raw data with replacement.

The calculation for the replacement set for catchers necessitates setting a distinct catcher replacement pool. Before it was suggested to have 4 catchers in the replacement pool so that will be carried over here. This sets catchers 25-28 as replacement and they reside in rows 291 through 294. Their replacement calculations have a dedicated set of cells. Our template uses cells G3 through J3 as the catchers ARV and G4 through J4 as the TRV. So, =average(G$291:G$294) is entered into cell G3. Presently the catcher values are adjusted by the ARV for the non-catcher pool. To fix this, highlight the adjusted extra hits of the first catcher, cell K267 and change the formula to =G267-G$3. Copy this formula in for the adjusted HRs, RBIs, and SBs for all of catchers as well.

If more than two replacement sets are used, the operations are exactly the same. Players qualifying at DH should be lumped into the OF pool. This can be done by sorting the two subgroups together or by changing the DH designation under POS to OF just for the purpose of efficient sorting. Remember to change them back to DH to avoid confusion during your auction or draft though!

The next change you will need to make deals with how the adjusted stats in the positive pool are calculated. There needs to be as many components as there are replacement sets. Based on our sample league and spreadsheet, the formula that needs to be entered into cell K1 to total the adjusted extra hits is =sum(K7:K150)+sum(K277:K290).

The final change as far as spreadsheet preparation is concerned is dedicating a pair of cells to the values of the lowest positive player and first replacement player in each position group. Based on our template, use the S column to input these numbers. For example, in our sample, you would enter =S290 into S3 and =S291 into cell S4 for the 24th and 25th catchers.

The spreadsheet is now ready again to be sorted. All of the same principles apply as they did in valuation with a single replacement set, except only those players in each individual positional pool are sorted. You may wish to highlight each of these different replacement sets of players and once again use the Insert/Define tool in order speed up your sorting. The ARV adjustments follow the same principles and the pools are precisely sized as before, just individually. Occasionally, since the positional pools are considerably smaller, the last positive player in a specific pool may be valued at something greater than $1. This occurs more frequently at 1B than anywhere else, but is not restricted there as it is pool dependent.

The values are accurate when all the individual pools can be sorted consecutively and the order remains the same within each one. Although they are separate, they are all linked by the calculation of the total of the adjusted stats. This number serves as the denominator in the category value equation for all of the players. So if any of the component pool contribution changes, the other pools may also change. Thus, it is very important that do each pool in succession.

After each pool has stabilized, is of proper size, and has ARV and TRV within defined limits, the data needs to be sorted together. First you can remove your second $Total column you created earlier. Second, it is recommend that you copy your entire worksheet onto a new worksheet, thus archiving the replacement value you just generated and reserving the work in case you find errors and need to adjust it later. Now on your new worksheet, you now need to copy your $Total column and copy over that same column using "Paste special" and insert just the values into the column without the formulas. If you do not do this, then your values will change. Once you have made this conversion, you can now take this worksheet and sort it by $Total, descending to generate your final values!

Closing Remarks

Admittedly, the calculation of customized values may appear as a daunting task. Understanding the theory and principles helps to ease the conversion from the abstract to a concrete spreadsheet, but patience is required as multiple sorting is tedious and it can take awhile to massage the ARV just right. With practice, the procedure can be perfected so that it takes less time to run a set of values than it does to watch the Sunday edition of Baseball Tonight.

Finally, as hard as we try, it is difficult to present this method so that everyone can completely understand the theory and operations. As such, we encourage you to take advantage of your site's message board to ask any questions regarding this system that you may have. Good luck.