baseballprimer

Baseball for the Thinking Fan

Win Share Chronicles

by Rob Wood & Tangotiger

An in-depth discussion of Bill James' Win Shares

ROB WOOD
And
TANGOTIGER
Copyright, 2003

TABLE OF CONTENTS

INTRODUCTION

Rob Wood and I decided to debate the merits of Win Shares. We brainstormed ideas, and came up with an agenda to form the basis of our discussion. The agenda is broken up into three sections: (1) Concept, (2) Methods, and (3) Mechanics. Within each section, we use the questions from our agenda to shed some light on the subject.   All comments made by me are on the left, and Rob's comments are on the right.

While not intended to be the be-all end-all of Win Shares, we hope to bring to the surface some thoughts on win distribution.

Thank you, Tom

DEFINITIONS

Value: past accountability of actual performance, after considering context

Ability: expected future performance, based on past performance, regression, and/or tools, in a context-neutral environment

Marginal utility: the change in the outcome, after a variable is introduced, into a known given context; this variable captures the marginal utility

1. CONCEPT OF WIN DISTRIBUTION

A. Do players "contribute wins" or do they "contribute towards winning" or do they "contribute towards trying to win"?

I think in a team sport, a player can't contribute a win. The best he can do is contribute towards a win. However, this would mean that a pitcher pitching a no-hit 1-0 loss can't contribute towards a win, since his team didn't win. So, I rely on the concept that a player's performance contributes towards trying to win.

Let's agree that the goal of win shares is to estimate a player's actual contributions to actual team wins. That is, it is designed as a value measure (looking backward), not an ability measure (looking forward).

In an ideal world, with infinite data breakdowns and infinite understanding of how runs are created and the probabilistic relationship between runs and wins (by half-inning, say), we'd probably agree that the best value measure of this sort would be a "value added" approach. For example, calculate the win prob of the team before and after each plate appearance over the course of the season for each player. Pitchers are a little trickier, and something like my win values stat would be needed.

That preamble is my introduction to responding to your initial question. I agree that in baseball a player cannot contribute wins directly or on his own. A player contribute towards a win, or even better a player's performance contributes towards trying to win.

This means, I think we agree, that a player's performance *even in a loss* can contribute towards increasing the chance that his team would win the game. As prototypical examples, we can keep in mind that a player who hits a two-out grand slam home run when his team was trailing by 3 runs in the top of the ninth definitely helped his team try to win the game, even if the team's relief pitchers give up two runs in the bottom of the ninth to lose the game. Or a starting pitcher gives up only one run in a 1-0 loss actually contributed to his team winning, even though the team's lack of offense led the team to defeat.

Tying it back to win shares, James' system predominantly uses seasonal data (not

	detailed play by play data) so that he perforce uses performances accomplished in losses as well as wins. I am entirely comfortable with this given what I said above and given the limitations of the data.
I think we are in agreement here, and I think I accept all your last post. Sticking to the topic 1a therefore, we can say that the concept of win distribution is based on "contribute towards trying to win". And if you move away from that, you are introducing some error somewhere. And that going to 162 games, these errors sorta/kinda balance out. My more specific point though is that if a player contributes towards trying to win, then on a game-by-game basis, you might have +1.5 wins contributed for the winning team and +0.5 wins contributed for the losing team, *IF* you take the sum of the parts of the contributions of the players. To continue then, I don't think that the concept of win distribution, as Win Shares is doing it, is even sound. The sum of the parts for an average team won't add up to 81 wins, but to something like 120 wins. Since wins contribution is based on marginal utility, and win shares is based on total utility, we've got a problem here. To convert from marginal to total is possible, but that would mean introducing an extra parameter to make the conversion. And that would mean that you could end up with negative win contribution in an absolute sense.	
	I don't see why the total of the marginal win contributions should be greater than the total number of wins over the course of the season. In fact, I thought they would be exactly identical. Here's my reasoning. Forget about win shares for the moment and think about an idealized value-added approach. Then each team has a 50/50 chance of winning before the first pitch. After every plate appearance the odds change a little, and the differential between the before and after can (ideally) be assigned to the various players involved in the play. Pitcher, batter, fielders if any. This, of course, is a marginal approach which we both embrace, I believe.

At the end of the game, one team wins and the other team loses. At that point, the win probability needles (meters) will be at 1 for the winning team and 0 for the losing team.

Thus, the sum of the marginal contributions among all the players on the winning team must be 0.500 (the ending 1.000 minus the original 0.500), and the sum of the marginal contributions among all the players on the losing team must be -0.500 (0.000 - 0.500).

Wait a minute, I may see a problem, but it's not the problem you see. Suppose a team goes 100-60. Then my description above suggests that its win contributions will be $100*.500 + 60*(-.500) = 20$. Of course, this is because they are 20 games (full games) above their "expected" 80-80 record. Thus, to get this total to be 100, we must add in the expected 80 win points. We may be getting ahead of ourselves, but maybe this is related to the absence of "loss shares".

Anyway, I don't think that your point is valid. In every game, the total number of "win contribution points" has to be a fixed number (be it 1 or .5 or any other number). I just don't see how it could be 1.2 in some games and 0.7 in other games.

But there does seem to be something strange going on in the "conversion" of marginal win contributions into total win contributions.

Yes, I think you are seeing the same thing as me, but in different ways.

If you decide that yes win contribution is marginal, then you have to decide if you want to have win AND loss contributions, or do you want to have positive and negative win contributions.

To go back to your example, if a hitter hits a slam in the top of the 9th, he might have added say +.8 wins (going from .1 to .9 win probability). But if they lose the game in the bottom of the 9th, the pitcher/fielders need to get docked -.9 wins. Or, you need to add +.9 losses. Doesn't matter either way (though my preference is to have the 2 separate).

So, if you have win and loss contributions, the winning team might have +1.5 win contributions and +.5 loss contributions (all numbers positive).

If you have only win contributions, then the winning team ends up at +1.0 wins, with the provision that a player is allowed to get NEGATIVE wins.

And you can also make this start at 0 wins 0 losses, as opposed to .5/.5. It works out to the same thing, but the former is more easily explained.	
	Yes, we are in agreement. After our discussion, here is how I think about it. Each team is allocated .500 win pts before the game begins. The allocation among its players is not obvious, but just for simplicity let's suppose the position player allocation is evenly distributed among the position players in the starting lineup, with presumably the starting pitcher getting more than the typical position player. As we mentioned, then during the game the win contribution points accrue to all players, some positive and some negative. So it is possible that some players wind up with significant positive contributions and some with negative contributions.

B. Do we expect the sum of the parts to add up to the whole?	
Yes, the sum of the parts should add up to the whole, if the totality of the game is considered, meaning wins and losses or positive and negative contributions, etc. If you get on base, and no one drives you in, then there needs to be a balance somewhere.	
	I agree completely. Sum of parts needs to add up to the whole.

C. Do losses exist? Does all the above apply to losing as well as winning?	
The game is made up of wins and losses, and your performance contributes either to helping your team trying to win, or facilitating (obviously without your intent) the other team's ability to win. I'm also comfortable saying that losses don't exist, but that games do. That maybe you don't help the other team in trying to win, but you use up your team's "time" (outs in baseball, for hitters anyway). Either losses as something tangible, or games as a context that needs to be allocated to individual players, needs to be considered. I think we'll end up with a $L = G - W$, so that even if losses don't tangibly exist, we can define as such.	
	I agree. I don't think we need to have a separate process to track losses or loss shares. These should just be the flip side of wins or win shares (the complement of games played).

To anticipate, this means that it is important to distinguish a player who has 25 win shares in 100 games versus a player who has 25 win shares in 160 games. But I imagine we'll get to this later. |

2. METHOD TO WIN DISTRIBUTION

A. What level granularity: do we distribute wins at a game level, in- game, or multi-game level? Or any/all?

I think we touched on this. We should probably validate the measure against the most granular possible, but settle for something that handles something larger, at the risk of introducing some error. So, distribute win contribution at a play-by-play level, but accept that we'll find a shortcut at a seasonal level, with an acceptable error range.

B. Can we distribute wins to players on losing teams (for any/all of the above 3 granularities)?

Yes, as discussed. There's a flip side to everything, and we should keep this in mind at all times.

C. If win contribution is marginal, then you end up with negative numbers. Is that ok?

As discussed, negative win contribution, or positive loss contribution, or whatever, is expected. There's no reason to "zero" things out (at this point anyway).

D. Do losses exist? Does all the above apply to losing as well as winning?

As discussed, we can look at losses as tangible, or the game context, as long as at least one of these is considered alongside wins at all times.

I believe we agree on all these points. Well stated.

1 AND 2 WRAP UP

It seems that we're on the same track, that we are considering marginal impact, but we also want to be able to account for absolute impact.

The path we are following seem to be - start the game at .500/.500, allocate marginal wins (positive or negative), such that at the end of the game one team is +.500 and another is -.500. Total marginal impact for both teams combined is 0. Since 1 win is allocated combined, split that 1 win across all players somehow, maybe based on PAs (for hitters, pitchers, and fielders, somehow), ensuring each team gets half an absolute win. The winning team will end up with 1.000 and the losing team will end up with 0.000 wins.

We can also keep in mind keeping the two ledgers separate, by keeping all the plus in one ledger and all the minus in another ledger. It's still the same thing. Kinda like in hockey where you are a +1, but that would be based on being on the ice for 3 GF and 2 GA for one player, or 1 GF and 0 GA for the other player. The first player probably had more ice time. In any case, it works out to the same thing based on marginal impact. Or keeping all the marginal totals in one ledger and all the redistributed absolute totals in another ledger.

Keeping it separate may open up other possibilities, going forward (like resetting the scale to the better accepted 0/0 instead of the mathematically sound .500/.500). Anyway, just something to keep in the back of the mind.

3. MECHANICS OF WIN DISTRIBUTION

A. *Value v ability*

Win shares is designed specifically to be a pure "value" measure, meaning that it looks backwards and attempts to allocate the win contributions of past games. Of course, conceptually this can be done on an on-going basis, but the idea is the same. Win shares is not an ability measure, though value and ability are typically highly correlated.

i. how do we attribute luck? does the player get it, or do we need to account for luck?

There are several flavors of "luck". But I don't see how any of them should be explicitly considered in the win share system. I believe that the players should be given all the credit of every play. Maybe you have a specific scenario in mind here.

ii. how much of the player's performance needs to be distilled to their (probable) true talent level, as opposed to the context or luck?

Related to what I said above, I believe the win shares system ought to be purely about actual performance in actually contributing to a team's win. Thus, I don't see any need to hitch it to ability or true talent level. Clutch hitting, I suppose, is the typical example here. I believe that a player should be given full credit for driving in the winning run in a clutch situation, even if we fervently believe that he has no special ability to hit better in the clutch the next time.

Clearly, there is a fair amount that can be done in this area, but this is all outside the scope of the win share system and what it is attempting to measure/capture.

iii. do coaches, GMs receive any credit for the performance?	
	As above, I believe the win share system should allocate all the credit (blame) to the players on the field. Thus, no credit to managers, coaches, GM's, etc.

i. through iii.	
Since the first two topics were more of a conceptual, and this topic is more related to Win Shares, it's probably good that we discuss Win Shares specifically. Yes, I agree with all of your points. While there would be a segment of readers that would prefer to distill out the luck portion of any performance (say Pat Tabler with the bases loaded, or Shane Spencer's Sept 1998, or Eddie Murray being "fortunate" to have so many opps with men on base, or Mariano Rivera being used in so many high-leverage situations), I don't think it's so clear-cut. That is, you can make an argument on either side. And even if I believe that luck should be distilled away from the performance, this is not a knock on Win Shares specifically. Win Shares goes in with the idea that all performance is attributed to the players in some fashion or other. I'll accept that Win Shares should be evaluated upon that premise.	

B. Context

i. if Warry Lalker has more value to a team playing at the Astrodome than playing at Coors, do we need to worry about this?

I think it's important here to note that Win Shares, as a value system, attempts to share responsibility of performance among all players playing together in a particular context, of actual past games. This means that we cannot necessarily take that player throw him into another context, and expect that player to have the exact same value. Therefore, we should not expect a value system such as this to answer "what if" scenarios, with much certainty. Those "what if" scenarios relate to ability, and the interaction of those abilities of a particular context. Win Shares is not equipped to do so, even if we could kinda use it as a guideline.

Therefore, if you have a certain player that is more optimal in a context than a similar player is sub optimal in another context (say Wade Boggs at Fenway), then we don't care in this system. If Dodger Stadium was built for Sandy Koufax, then all the better for him.

As a value system, we want to know how much of the actual performance to attribute to the players. Since every player on the team has the same playing environment, everything will balance out. And since luck does not get its own share, it has to balance out among the players.

ii. what context is a player's performance analyzed under?

The player's performance is analyzed under the context of the actual playing environment. If Nolan Ryan pitches at Coors, it's understood that that environment kills all pitchers (some more than others). We adjust for the context, but equally, for all players playing in that context. You don't give different adjustments for curveballers or fastballers or lefties or righties. So, Coors' baseline might be 5.5 runs / game in 1995 while the Astrodome's baseline context might be 2.8 runs / game in 1975 (just examples). The performance of the players are analyzed under those conditions.

i. and ii.	
	I agree completely with your comments about win shares only being valid within the player's context. However, I am not sure we know how big a deal this is. I don't think I have seen studies that show that certain types of players are more valuable than others given a specific situation. I have always believed that as a general principle good players are good players no matter what the context. I know the Wade Boggs argument, but I am comfortable with the win shares system crediting him with his performances in Fenway Park. The flip side is Joe DiMaggio who suffered greatly hitting in Yankee Stadium. Let's say he hit a lot of balls into Death Valley in Yankee Stadium's left-center field that were caught but would have been home runs (at least extra base hits) in other stadiums. JoeD's relative lack of performance should be reflected in his win shares. JoeD would have been more valuable in Fenway Park or virtually any other AL park.
A more important example might be say Vince Coleman, who's value is extremely high when: the run environment is low and there are not a lot of HR. Put Vince Coleman at Coors in the 90s in front of HR hitters, and he doesn't steal as much. Maybe not Coleman, but say Brett Butler or maybe Paul Molitor. Every event is dependent on the context. Change the context dramatically, and every subcomponent of a player's skillset changes. A single is worth less with Pedro on the mound than when a bad pitcher is on the mound, etc. The balance doesn't remain the same. So, you can have two equivalent but far different players, say Willie McGee and Jack Clark, and depending on their contexts, one would generate more runs for his team than the other.	
	Having said that, you are definitely correct that there is a set of issues surrounding park effect "splits". Some parks are notoriously asymmetrical, favoring LHB over RHB, say. How that should be taken into account in evaluating LHB and RHB is a tricky issue that I am open to suggestion. In truth, I am not sure that the win share system goes beyond an overall run context (meaning no splits).
I agree with everything else you said, and I agree with this statement as well. In terms of value, the run context, rather than all the subcontexts, is what should count.	

iii. how do we establish the playing time context?	
First off, how do we establish a single playing context, or can we, for hitters, pitchers and fielders? Suppose we make it PA or BIP based. In a high-scoring environment, there are more of those to go around than in a low-scoring environment. So, do we give out game shares based on percentage of PA racked up in that game, such that the total game shares for each team is 1? Are pitchers game shares based on IP? How do you split the game shares around? Can we?	

iv. assuming the talent level is stagnant across years and eras, how do we make the playing context equivalent across teams, years, eras, like wins.	
Similar to above, there are fewer PAs in 1967, but there's always 1 game, 1 win.	

v. what about change in playing conditions, where talent level is not stagnant?	
James' timeline adjustment is rather simple.	

B. All	
	Okay, now we seem to be getting into the nitty gritty. Leaving the realm of conceptual approach to how win shares are actually calculated. As we stated previously, ideally win shares would be calculated on a per plate appearance basis, and trapping the win shares to be always equal to one summed over all players in each game. I think in this approach, the "context" issue is moot. Or rather the context is fully reflected in the differential win probabilities before and after each plate appearance. A home run in the AL 1968 is worth more than a home run in AL 2000. But since win shares predominantly uses seasonal data rather than detailed play-by-play data, there may be a question on what basis the win share credits should be doled out. For hitters, I think James does the right thing. He calculates runs created for each player, and then attempts to subtract out "baseline" runs based upon the player's outs consumed and the league average runs per out times 0.52. This leaves the player's marginal runs created and is the player's offensive claim points. We probably don't want to discuss the specifics right now, but I think this is the right approach. For pitchers, he does a similar thing based upon a pitcher's IP and runs allowed (differentiating earned and unearned runs allowed). He credits pitchers with claim points for runs allowed above a minimum threshold level based upon the league average. Fielding is a lot more complicated in the win share system, but again I think the approach is correct. Individual fielders are given credit based upon their putouts, assists, range, etc. (not directly on innings played since this is unknown for most of history). Let me get your thoughts on the above before we discuss era effects, talent level, timeline, etc.
I think we'll be discussing this topic for a while. We said that we always need the flip side to wins (losses) or something to go along with wins (playing time context).	

So, I'd like to talk about the following: 1 - hitters on the same team 2 - hitters on different teams 3 - hitters in different eras 4 - pitchers 5 - fielders 1 - If you have hitters on the same team, there won't be an issue regarding playing time context. If they both have 600 PA or 500 PA, they both have the same impact in terms of playing time. 2 - If you have hitters on different teams, 600 PA might not mean the same thing. Getting 600 PA at Coors or a high-run context or 600 PA at the Astrodome or a low-run context means that the Coors guy only played in 140 games while the Astrodome guy played in 150 games (or something like that). Therefore, we need to have a "playing time" converter that will adjust among teams of different PA levels. Essentially, you can take % of team PAs. Or what I'll call PA shares.	
	I don't see why we need this. Win shares are already denominated in wins (I know I am repeating myself). The win shares calculation already takes into account the player's run context. For example, it knows that 600 PA in Coors is not the same as 600 PA in the Astrodome. Or rather, 100 runs created in Coors is not the same as 100 runs created in the Astrodome. Am I missing something here? Maybe you are saying that before we calculate win shares, we have to think about the "basis" of the playing time context. I'll comment more in this vein below.
3 - This is answered by the above. Again, getting 600 PAs in 1967 is not the same playing time as getting 600 PAs in 1994. GAMES is what we want, and not PAs. Baselining things so that you compare against his own team's total PA will be a good start.	
	Yes, I agree, and I thought that this is essentially what James does now. Well, I may be beginning to see your point. Are you hinting that the runs created basis is essentially a plate-appearance perspective and not a games-played basis? If so, I agree.
	Though I don't see the problem with converting runs created into a "wins" number by taking into account how many extra runs it takes to generate an extra win in the ballpark. This is how Pete Palmer comes up with TPR as you know. Saying the same thing another way, I don't see how you can do anything with a "games played" perspective. Runs are scored via plate appearances. So looking at the outcomes

	of plate appearances is the proper way to estimate the runs contribution of a player's offensive performance. Right? I guess converting runs into wins is the part you are questioning.
The marginal runs to marginal wins computation (and I'm using marginal as I defined it at the beginning of this series, in terms of marginal utility, and not the James-version) as Palmer does it is correct.	
4 - Pitchers on the other hand are not given PAs, but outs. Using team IP as the baseline is fine with me. As it turns out, since all teams and all years have essentially the same number of total IP, we don't have to worry about this too much. But, we should do the calculation anyway, and come up with IP%.	
	I agree here. Though some may want to do per-game analyses here (as I do in win values). Especially since we have the ability to "parse" a pitcher's performance on a per-game basis. But I don't quibble with seasonal data analyses using innings pitched as the basis.
5 - Fielders I think would get the "balls in play" / "team balls in play" as their BIP%. Course, not always easy to calculate, but at least we get to see how much fielding time they actually put in. A SS has more fielding time than a 1B. Catchers are another issue. But, let's leave fielders alone for now. Win Shares does a 2-step process in 1 step. While we agree that you want marginal wins, and that we want to somehow redistribute the "absolute" portion of wins somehow, Win Shares does this in a 1-step process by using the James-version of "marginal". This I believe is where the problem lies, and I recommend that we take a step back, and stick to a 2-step process first. Let's get it right, before we try to make it easy.	
Going back to your paragraph before that, let's say you have a league-year that scores and allows 3 runs / game, and you have another league-year that scores and allows 6 rpg. Assume that you have an average team with average players in both leagues. Now, the team in the low-scoring league will have about 36 PAs per game, and the team in the high-scoring league will have about 45 PAs per game (just picking numbers here).	

Let's say that the hitters in each team are all equals, and they play in all games. These hitters will have zero marginal wins, and they will share in part of the .500 absolute wins / game. Giving the hitters .27 wins / game (I'm using .27 for ease; you can make it whatever you want), that means each hitter gets .03 absolute wins / game to add to his zero marginal wins / game, for a total utility of .03 wins / game.

And this applies to players on the low and high-scoring teams.

However, you have players on one team getting .03 wins/game on 4 PAs/game, and players on another team getting the same .03 wins/game on 5 PAs/game.

Stretching this out to 162 games, and you've got 14.6 win shares for both players, but one does it in 648 PAs, and the other in 810 PAs. However, these PAs are equivalent because they both represent 162 games in the context of their league.

Now, since we've agreed that win contribution needs a companion (be it loss shares, or some playing time context), what playing time context can we use so that we can say that these two players provided equal contribution in the same playing time?

Now, it gets worse. What if these two guys were on the same team? What if you have two guys both with 14.6 win shares, but one did it in 648 PAs and the other did it in 810 PAs. Here, it's clear that they did not play 162 games each.

Therefore, we need either loss shares or game shares. We need something to go along with win contribution, so that you can provide the proper dimension. PAs, unadjusted, won't work. They'll work on an in-team basis, they might work on an in-year multi-team basis, and will not work at all on a year-to-year basis.

And to make matters worse, we haven't even touched on balancing hitting pitching and fielding.

Do you still agree with a prior statement of yours that (I think) says we need to figure out the 25 win shares in 600 PAs not being the same as 25 win shares in 500 PAs?

Very good job of explaining the issue. I had a blind spot and couldn't see it until you described it the way you did.

Okay, different eras and, presumably, homeparks will exhibit different average number of

	plate appearances per game. But if that is the issue, why don't we simply normalize by the average number of plate appearances by the player's team in the season.
Yes, exactly what I want to do.	
	I imagine this is obvious, but let's walk through a couple of examples just to make sure I am following. One case is a high offense. Player A has 25 win shares in 600 plate appearances, where the team averages 45 plate appearances per game. The win shares is 25 and we are seeking the loss share. Player A's context is 600/ (45/9) = 120 full games, if I did the math correctly. Player B, on the other hand, has 25 win shares in 600 plate appearances in a league-park context in which his team averages only 36 plate appearances per game. Then B's context is 600/(36/9) = 150 games. Wait a minute, am I crazy or am I thinking about this backwards? Which player had the better year, taking into account their context and the loss share context? Maybe it's too late at night for me to make sense of this. I think I need help here.
No, I think you have it right. A guy's PAs have less game impact, if there's more of them to go around. In your example, the guy managed to get 25 win shares while playing only 120 equivalent games. The other guy got 25 win shares playing in 150 equivalent games. I'm happy to leave it as "equivalent" games, and move on to pitchers and fielders. For pitchers, I'd take equivalent games based on innings, since this is what a pitcher is allocated. So, pitcher innings / team innings * 162. Ok, so let's stop here for a moment. The group of hitters will have a total "equivalent games" of 162 * 9, while the group of pitchers will have a total equivalent games of 162. So, I'd suggest that a batter's equivalent games should be divided by 9. So, 120 equivalent games for a batter is really 120/9, or 13.33 games, or 13.33 "full equivalent games". Either that, or multiply the pitcher's games by 9. The reason is that 1 pitcher faces all the batters.	

However, thinking it through, it's the defense that should be multiplied by 9, and therefore, maybe the pitcher's totals should be multiplied by 6, leaving the other x3 to the fielders. Thinking even more, you can make each individual pitcher's multiplier different, based on how much they rely on their fielders. Pedro might by x7 and Tommy John might by x5, or something. Let me know what you think about all this, and if you prefer the x9 or /9 approach.	
	The cleaner the better, at least for now. So let's not introduce different factors for pitchers who rely on their defense to differing degrees (Pedro vs Tommy John) yet. I have a slight preference to scale down a hitter's context to the individual; that is, dividing by 9, rather than scaling up a pitcher's context. Fielders should be handled similarly.

iii. how do we establish the playing time context?	
As discussed - For hitters: PA/teamPA * 162. For pitchers: IP/teamIP * 162. We didn't discuss this, so For fielder: PlaysMade/teamPlaysMade * 162. We can discuss "playsmade" and what that means in a future topic on fielding, if you're ok with that. So, what we have here are hitting shares, pitching shares, and fielding shares. We still haven't tried to make those equivalent across each other. Again, I think we have another topic on this, so let's wait for that to get started.	
iv. assuming the talent level is stagnant across years and eras, how do we make the playing context equivalent across teams, years, eras, like wins.	
Done by the above.	
v. what about change in playing conditions, where talent level is not stagnant?	
I've done some work on this in the past. It's very very hard to try to establish this. Essentially, I've found some small improvements from post-war era, but not much. Pre-WWII is a different matter, as is the difference between NL and AL at that time. I would propose that we either say that the timeline is a different animal altogether, and that we don't look at it, or we spin it off into its own major topic altogether. The James adjustment is horrible, and we can discuss the merits of that if you like.	

C. Separation of pitching/fielding/hitting

i. if a team's pitchers has a lot of strikeouts, fielders have less impact overall. how do we account for that?	
Ok, so I guess now we have to talk about this! I'd start off with something like: if there are 40 PAs / game, 30 of which involve fielders, and 10 of which don't, then we should take the 30 and split it as 15 for fielders and 15 for pitchers. I don't know what the split should be, and we might be close to an answer soon. So, on a league-level, the split is around 62.5% of defense shares are for pitching, and 37.5% for fielding. This changes for every pitcher, and we can follow this approach for each pitcher. For fielders, just take the team level, and follow-through as well. The whole thing should add up, I think. However, now we are talking about PAs instead of IPs. I'm not sure if that is correct, and I'd like to hear your thoughts on the matter.	
	Yes, I agree.
It's almost like we're trying to find the pitcher's share of the pitcher/fielder dynamic based on PAs. But then we apply that factor to the pitcher's share based on IP relative to team IP. So, we end up with something like: Pedro: 213 IP / 1400 team IP * (15 solo PA + 12 halfPAs)/(15+12+12 total PAs) * 162. I don't know if that even makes sense, but I think this is where I'm going to end up with.	
	I agree with the concept, but the numerical example puzzles me. Is that right? What are the 15 & 12 numbers? And what units will the result be in? Is this the formula where K, BB, and HR will appear as the fielding-independent plate appearances?
The 15 "solo PAs" would be Pedro's K+BB+HR, etc. The 12 "half PAs" is Pedro's share of the 24 BIP. There are 39 PAs in all in this example. So, while a typical pitcher's share of all PAs was 62.5% as mentioned earlier, Pedro's share in this example is 69%. If Pedro pitched 213/1400 of all innings, this means that with Pedro+fielders, they used up 15.2% of all innings. 15.2% * .69 = 10.5%. So, Pedro's individual share of the Redsox	

innings is 10.5%. If we multiply this by 162, his "defense game share" is 17 games. I'm not sure how logical all that is, but it seems like a good starting point, anyway.	
	Okay, I see what you are saying. I agree that this method of accounting can be used to derive each pitcher's "full game equivalents" over the course of the season.

Ok, so now we have the equation to get "full game equivalents" for hitters, pitchers, and we'll think about one for fielders (probably based on plays made, but that would not be too fair to the catcher).

Anyway, so the sum of the parts on offense will work out to 162 games, and the sum of the parts on defense will work out to 162 games (with around 100 or so going to the pitchers).

However, that gives us 324 games. Since there are 162 offensive games and 162 defensive games and 162 real games, if we want to the sum of the parts to add to the whole, we therefore need 81 offensive games and 81 defensive games. Do you agree?

If that is the case, then this would mean that the average hitter would have 9 game slices(81/9). If the average hitter has 9 game slices, that would mean that the average hitter has 4.5 wins contributed.

But does this mean that a hitter has to top off at 9 wins created, or +4.5 wins above average? I think we can both agree that this is not true. That Bonds is actually over 10 wins above average. So, Bonds would have say 15 wins created, but with only 8 or 9 game slices.

Too me, this is acceptable, because of the way runs and wins are created. They are not created in a vacuum. They are based on interdependence, and therefore we should NOT expect things to add up nice and neat. We should expect things to add up, but not all nice and neat.. The result of this is that some players will have more wins than game slices. And since the difference of game slices and wins is losses, Bonds has negative losses.

I believe that Bill James has gone through this whole exercise. That he couldn't fathom negative losses or having more wins than games, and therefore said "I couldn't get it to work". He can't get it to work because of his insistence that wins be bounded by 0 on one side and games on the other.

However, wins is better conceptually as wins above average + share of average wins

So, if Bonds is +10 wins above average, and if his PA slices gives him 4.5 wins, then Bonds in this case is 14.5 wins in 9 games.

Let's spend a minute on this topic since it may prove important. Before discussing James' seasonal version of win shares, let's think about the idealized game-by-game play-by-play win share system. Therein each player is credited with how much his accomplishments helped his team (try to) win the game. For hitters, this can be thought of as the difference between his team's win prob before and after each plate appearance. Then, by construction, the sum of the players' (hitters, pitchers, and fielders combined) win shares will be 1.0 for a game the team wins and 0.0 for a game the team loses.

Consider a hypothetical great game by Babe Ruth. Suppose the Yankees are losing by 3 runs with two outs in the bottom of the ninth when the Bambino hits a tying 3-run home run. This is a lot of win prob; let's call it 0.4 (0.1 goes to 0.5, say). Next time Ruth comes up the Yankees are again down by 3 runs with two outs and he hits a game-winning grand slam! This is a ton of win prob; let's call it 0.7 (0.3 goes to 1.0, say).

Adding 0.4 + 0.7 = 1.1 which, of course, is greater than one. Does this make sense? That is, can any player properly be credited with more than one "win" in any one game?

We agree that negative idealized win shares make sense; for example the pitchers who gave up those dramatic home runs to Babe Ruth to turn a winning lead into a loss will be rightfully credited with negative win shares.

But something may be troubling about crediting a player with greater than one win in a game.

Yes, the way it works out, in marginal terms it's fine. You can have say a game where the lead ALWAYS changes. So you keep oscillating between .3 and .7 win prob after every half-inning. So, you get +.4 after one inning, and then another +.4 after another inning, and another +.4 , etc, etc. On the flip side though, that is balanced by being -.4 for losing your lead each time.

So, this gets back to what I was talking about last week, about keeping "positive and negative" ledgers. You may have such a game where you have lots of lead changes, so that you might have +2.6 win contributions in one game for one team, but also -2.6 win contributions for losing the lead, so that going into the top of the 9th, in a tied game, it works out to +/- 0 win contributions total.

On a MARGINAL basis, when the game is over, the team will have contributed +.5 wins, relative to average (that +.500 may have been 1.1 positive moves and 0.6 negative moves).

But you also have the .500 wins that you started the game with, which you redistribute based on the playing time share that we last talked about.

I see no reason why a player cannot contribute more than +1.00 win contributions.

So, to continue, if Bonds gets 1/9 of the hitter shares, and the hitters get 1/2 of all player shares, then Bonds gets 1/18 of the .500 wins that you started with.

Bonds' win absolute contribution is therefore +.09 (let's say) +.500/18 = .12 wins per team game. At 150 games, he's contributed 18 wins. His "game slices" is going to be around 9 games.

And therein lies the "problem".

I wrote the following last year http://www.geocities.com/tmasc/winsloss.html and I think we are at the point where this concept (based on the Mills Brothers) makes more sense than Win Shares.

ii. if an infielder happens to pitch in front of a groundball staff, and an equivalent infielder happens to pitch in front of a strikeout/FB staff, do we expect their "value" to be the same? do we care?	
We discussed this kind of issue in the past. We are concerned about value and what has transpired, and so we don't care. Barry Larkin has less value on a staff with Pedro, RJ, Clemens, Dibble, et al than on a staff with TJ, Kaat, et al.	
	Yes, I agree since we are talking about a value measure (not ability).

iii. what is the proper split between hitting/pitching/fielding? is this constant across teams/years/eras? is the split dependent on the type of team you have?	
For off/def, the split should be 50/50. If you look at the stdev of RS and RA across eras, you will see that this is true. As well, I've found in the past that a .250 OBA guy against a pitcher/fielder with a .350 OOBA will result in the same matchup as the reverse. To move away from that, even to get to 52/48, will need some 'splaining. I found that the James argument basically boiled down to "cause it works better". My response to that is "it probably works better, because the system is flawed somewhere, and you have to compensate for it somewhere else".	
	I have done a bunch of research into this matter as well. I have come up with different splits, but I am definitely in favor of saying the offense/defense split is 50/50. Just so I am on the record, I don't think there is anything magical about the 50/50 split, but there are a lot of forces that tend to balance the game to that midpoint equilibrium.

iv. how do you measure fielding without PBP data? if you have PBP data, how do you measure fielding? v. how much does DIPS play into this? how much of the luck of allowing hits on balls in play will be attributed to the pitcher or fielder or discarded into a luck component?	
Very, very carefully. I think the ongoing work by Saeger, coupled with the work from MGL will go a long way in terms of establishing a baseline to work against. Bill James should have done this, esp with his easy access to STATS pbp, and I find it a failing in Win Shares that he didn't. I think that all of his fielding measures has so much guess work in it. Really, we can spend the next month picking it apart, because the weighting were provided with little justification.	
	Yes, and I am not very interested (or knowledgeable) in the real nitty gritty of PBP defensive measures. However, we'll have to spend some time on this later.

D. Equations

i. should we use RC, BaseRuns, or (static or dynamic) Linear Weights?

It doesn't make too much of a difference at the MLB team level (OBA of .300 to .400), since they all work pretty much the same at that level. My preference is to use the best one, and that would be BaseRuns. It doesn't have the level of illogic of RC or static LWTS, and its accuracy at that level is on par with these.

You can also try dynamic LWTS, but that's a bit harder to do.

I also have no problems with trying to adjust for the timeliness of performance (whether for relievers or for hitters with men on base). We are after value after all. James' adjustments are pretty basic. If you're going to do adjustments, do them right. Limiting it to BA and HR with men on base is ok, but not too good.

Really, I would use my LWTS by 24 base-out states, and come up with the proper adjustments. http://www.geocities.com/tmasc/lwtsrobo.html

And for relievers, Leveraged Index would be better if you have PBP. James' version should be baselined against this to see how accurate his is. Palmer's version should be baselined against LI as well.

4. LOOSE ENDS AND NEW TOPICS

A. The conundrum of win contributions being greater than game slices, on a player basis, for some players

This goes to the heart of the matter, and the reason that I think Win Shares fails. The process to separate a player's *marginal* contributions in his attempt in trying to help his team towards winning into an *absolute* win contribution value is bound to fail at some point. And that point is that you get absolute win contributions that are greater than absolute game slices. If this can be accepted, then I can accept the premise of Win Shares. I believe that the Mills brothers approach, albeit by leaving it as "win points" and "loss points" separately (and letting the reader the option of how to combine it), is the one that conforms to reality (marginal contributions), and in the end will prevail.	
	I am fine with where we left the possibility of "more win shares than game slices" issue, as well as the possibility of negative win shares in the ideal system.

B. The arbitrary weights in the fielding component

It is arbitrary, and James gives no real justification. I will point you to Mike Emeigh's series http://www.baseballprimer.com/articles/emeigh_2002-11-18_0.shtml that has tremendous analysis on the fielding component of Win Shares. Probably the best you can find.

Here is a snippet exchange I had with Mike and Charlie

Mike: The result of this choice is that fielders are penalized in Win Shares for every play made by a fielder at another position - plays which they themselves had no chance to make.

Tom: Actually, this makes sense in the James scheme. That is, James assigns the Yanks fielders 52.5 Win shares, while an average team would have received 41.3 Win Shares. Therefore, under this scheme, the players on the Yanks are being compared *to each other* as well as to the league. Therefore, an average fielder playing in an above average fielding team (as noted by the James system of 52.5 WS versus 41.3) should come out looking worse than his teammates, but still come out looking exactly like his average counterpart on an average team.

I'm not sure how clear all that was, or if I'm contradicting what Mike or James are saying. But, as best as I can figure it, that's what James is trying to do.

Just to be clear, this is what I think James is trying to do. I don't know if he is successful. One way to test it is to construct various "teams", and start to control for things.

For example, start off with an average team. Add 2 assists to your SS, and subtract 2 assist from your 2B. You'll probably have to subtract 1 putout from your SS, and add 1 to your 2B. Now, what is the effect? How much WS did your SS gain, and how much did your 2B lose? Do the same, but with your SS and 3B (this time, you don't need to worry about PO). What happened?

Now, do the same, but for a top team. Are the changes similar? The changes should be very similar, as putting Vlad on the Expos or Vlad on the Mariners as a hitter should have a very similar effect (though not exactly). This method doesn't apply to pitchers.

Because of the complexity of the James calculations, I find that it is easier to look at various marginal effects to note what is really happening under the hood.

Charlie: I have looked at Win Shares from what you suggest, and it is clear to me that James did NOT do this.

I said the idea of such adjustments makes sense. James just did not do them right

Tom: My objective was simply to determine if James did the claim points correct by position. Removing an assist from the SS and giving it to the 3B should have a net effect of zero. I don't know that it does. And even if it does, what is the degree of change? Is this change correct?

Mike: OK, that wasn't clear from the comment. I don't think the net effect would be zero, because James uses a 50-point scale for assists at 3B and a 40-point scale for SS assists, and because the LHP adjustment differs for 3B and SS.

Tom: There are *many* forces at work here! While James uses a 50pt scale for assists for 3b, and 40 for assists, the "available total points" for 3B is 24 and for SS is 38(?). So, from this standpoint, the 3B assists get 24 x .5 = 12, and 38 * .4 = 15.2 for SS assists.

I agree with Mike's assessment that there are many forces at work - breakdown of assists, po, dp within position, - claim points by position, - comparison relative to teammates, - whole team fielding compared to whole team off+def, and - whole team compared to league.

Whew! Ideally, moving an average guy from a great team to a bad team should have little impact on the WS of the player in question. (It might have some, because of the interaction, like hitters, but generally speaking it would be pretty small.)

However, given all these forces at work, I don't see the evidence that all these forces have been balanced out to the point where we can say that this is a true statement. In fact, I think it is a daunting task to determine the extent to which this is true or false. The number of forces and variables at work here would make it a small project until itself to determine the validity of the fielding portion of Win Shares. That's not to say that all is lost. All those arbitrary claim and point assignment that James make may actually be calculable through a more rigid analysis.

Then again, Clay and Charlie's approach is probably just as good and far easier.	
	Again, I defer to the experts such as you, Charlie, Mike, and Clay. My defense of James' win share defensive system is that he has a great idea. Start from the team stats and then work down to positions and then to individual fielders. Sure, he has a lot of arbitrary numbers in his allocation system. You know, arbitrary weights, arbitrary thresholds, etc. Every system James ever developed can be criticized for this. However, James has a great intuition as to what works and what doesn't work, and of course, knows baseball and baseball stats as well as anyone. To my mind, James successfully plays the role of innovator. He develops good ideas but doesn't quite implement them correctly. It is left to others, often with more mathematical or statistical acumen, to clean up the mess. Of course, I am speaking hyperbolically since I have infinite respect and admiration for Bill.

C. Real-life example of using extreme players (bonds, pedro, pitcher-hitting, horrible pitching).	
All these will show the boundaries of Win Shares, and the issue of win contributions and game slices.	
** Hitters ** We touched on this already, so I'll try to keep it short. The game slices of Bonds as a hitter, if he plays 162 games at the #5 hitter, is 162/2/9=9. You start with 162 games, divide by 2 for offense/defense, and divide by 9 for 9 hitters. So, Bonds has 9 "absolute" offensive games. An average hitter would have 4.5 "absolute wins" giving that playing time. A hitter with 9 "absolute wins" in these 9 absolute games would be +4.5 wins over average. I think it's easy to say that I can find many examples of a hitter that is far more than 4.5 wins over average. Bonds would be 10+ wins above average. Giving Bonds 15 absolute wins, with 9 absolute games is the problem.	

** Pitchers ** Let's look at pitchers. Defense (pitchers + fielders) get 81 "absolute" games. With 1458 innings at play, that gives you a rate of 18 defensive innings per absolute game played. With Pedro pitching 216 innings, that gives Pedro+his fielders 12 "absolute" games. While the average pitcher would end up with about 8 of those 12 games, a pitcher like Pedro, who uses his fielders less, will end up with say 9 "absolute" games out of the 12.

(In 2002 for example, Pedro had 787 PA, of which 307 did not involve his fielders, or 39%. Assuming that the other 61% is split evenly between the pitcher and the fielder, that gives Pedro a share of 39+30.5 = 69.5% of the 12 absolute games, or a total of 8.34. Let's use 9 for ease. The league average is 29% don't involve fielders, giving the average pitcher 64.5% of the share.)

Again, like the Bonds example, an average pitcher with 9 absolute games will have 4.5 absolute wins. And again, we can easily find pitchers who are better than +4.5 wins over average. And again, you have a pitcher's absolute win contribution being greater than his game contribution.

** Bad players **

Bill James also does come up with negative win shares for truly horrible and truly horrible pitching and hitting. He zeroes them out. Zeroes out!!! Are you bothered by that?

D. Summary

The system fails from a conceptual standpoint (trying to turn marginal impacts into absolute impacts), and from a mathematical standpoint (zeroing out on one end, and ending up with more wins than games on the other end). His splitting of offense/defense is done to make it "work better", instead of using 50/50. A ridiculous decision. His splitting of fielding/pitching makes no use of the Pedro issue, instead applying the same split for all pitchers on the same team. His fielding weighting is done arbitrarily, with no evidence that everything will balance out properly between positions.	
	I agree with your point. But I don't agree that win shares is a failure because of this. In all candor, Bill James is not a mathematician, and in the past has developed less than perfect implementations of other good ideas. It's kinda like runs created being "flawed" since it does not properly handle the incremental runs an extra home run provides. Base runs handles this properly and is therefore more mathematically sound than runs created. But a lot of people still use runs created and it is good for most purposes. I fully expect the win share system to evolve over time as improve in the areas that it needs improvement. I am not sure if Bill himself will frequently tweak the system, but I imagine if Bill doesn't many others will. For example, you probably know that Charlie Saeger and Joe Dimino have modified win share defensive machinery to correct some minor flaws. Similar advancements will surely come for win shares too.
I agree that James has excellent ideas, perhaps the best in the business. The implementation, the mechanics of his concepts are sometimes (many times) questionable. I'm sure some of his mechanics can be improved (for example, going back to 50/50 on off/def as we agreed, or in dividing the fielding/pitching based on balls in play per pitcher	

and not per team, etc).	
It is much too early to declare any kind of success on Win Shares. Given that Bill spent 4 years on this, he would have been best served going for peer review, whether at Primer or SABR, etc.	