

Clutch Hitters Revisited
Pete Palmer and Dick Cramer
National SABR Convention
June 30, 2008

Do clutch hitters exist? More precisely, are there any batters whose performance in critical game situations consistently exceeds expectations, as established both by that batter's performance in less critical situation and also by the relative performance of average batters in critical game situations?

Thirty years ago one of us published a first investigation of clutch hitting (1), using 1969 and 1970 data (2) that at the time seemed the only play-by-play information that might ever become available. Its conclusions, that any clutch abilities were too small to be either detectable or meaningful, have been confirmed repeatedly (3) as much more data have emerged. However skepticism remains. The occasional stresses that all of us experience in our daily lives are certainly felt as negative influences on our own "clutch performances", and professional athletes in particular often talk about the challenge of contending with the pressures of critical game situations. Thus "clutch hitting" exemplifies the puzzling and fascinating conflicts that occasionally arise between human perceptions and the results of objective investigation.

For example, recently Bill James (4) has proposed that the existence of clutch hitters, as exemplified by David Ortiz's recent heroics, is obscured by "fog", that is, the unavoidable random variation in the performances of all players and game situations that underlie those objective investigations. Perhaps, he says, clutch hitting is a strong, and so more consistent and detectable, ability only for certain classes of players, identifiable by their personality type or overall hitting style.

Our response has been to perform several new studies, taking fullest advantage of the last fifty years of play-by-play that Retrosheet (5) now provides, that attempt objective answers to the following questions about clutch hitting –

- 1) Do batting performances actually decline in critical game situations, in general?
- 2) How thick is that "fog" that Bill reminds us about?
- 3) Over the last fifty years, what are the largest "clutch" or "non-clutch" career performances?
- 4) With respect to Bill's proposed player classes, how many players would need to be members in order to be persuasive about clutch hitting?
- 5) Would a "clutch hitting" skill reflect well on the players who possessed it?

For those readers who prefer conclusions without statistical equations or argumentative logic, here are the objective answers to those questions.

- 1) The average batting production by major league players with runners in scoring position, as measured by OPS from 1957 to 2007, is indistinguishable (if anything very

slightly higher) for the 15% of plate appearances when the game is late and close (as defined by Elias), once appropriate corrections are made for the substantial superiority of the pitchers offsetting the greater frequency of intentional walks and the modest superiority of the batters.

2) The “fog” of random statistical variation is much greater than most baseball fans realize, even those with an analytical bent. Indeed, that random variability has about the same effect on season outcomes such as batting titles and pennant winners as do differences in player skills. Considering also this lack of effect by game situation on overall batting and the small number of high pressure game situations that any individual batter encounters, objective evidence that some particular hitter’s superior record in clutch situation is a skill rather than a chance result is very unlikely ever to appear. To that extent, we agree with Bill James.

3) On a career basis, David Ortiz and a different clutch hitter nominee (5), Mark Grace, are both around a third of the way down a ranked list of “clutch performances”, among all players with 3000 plate appearances from 1957 to 2007. Positive but hardly to any notable extent. The highest career clutch performance was by Scott Fletcher, who on average contributed about ten more runs’-worth per year (or one game per year in the standings) to his teams than his season statistics imply. At the bottom of the ranked list is Richard Hidalgo, who contributed ten fewer runs’-worth. The overall distribution of the 897 career clutch performances is entirely attributable to chance deviations around our first result, that batters in general perform no differently in clutch situations.

4) The considerable thickness of the fog also makes identifying any class of players likely to possess a “clutch ability” into a considerable challenge. Inspection of the ranked list does suggest a tendency for power hitters to have experienced lower clutch performances. However this tendency may instead reflect slight weaknesses in the metrics we have used.

5) To repeat the concluding suggestion in the original study, but also now knowing that batters in general perform no differently in clutch situations -- if a batter’s performance in clutch situations truly exceeded that in other situations, might that not indicate that he was putting forward a less than maximal effort in non-clutch situations? Why would that be a desirable player characteristic?

We now describe, in general terms, the studies that lead to these conclusions. (methodological details of course being available from the authors.) First, perhaps the most important new finding. We compare the general performances of batters in tense game situations with performance in all situations, for two groups, one based on the Elias definition of “late and close” and considering all major league plate appearances from 1957 to 1979, and the other based on the Mills brothers’ model (see below) and considering only the previously mentioned group of 897 players who clearly had more successful major league careers. For this purpose we use the widely accepted OPS metric, originally proposed by one of us (6). Here are the results.

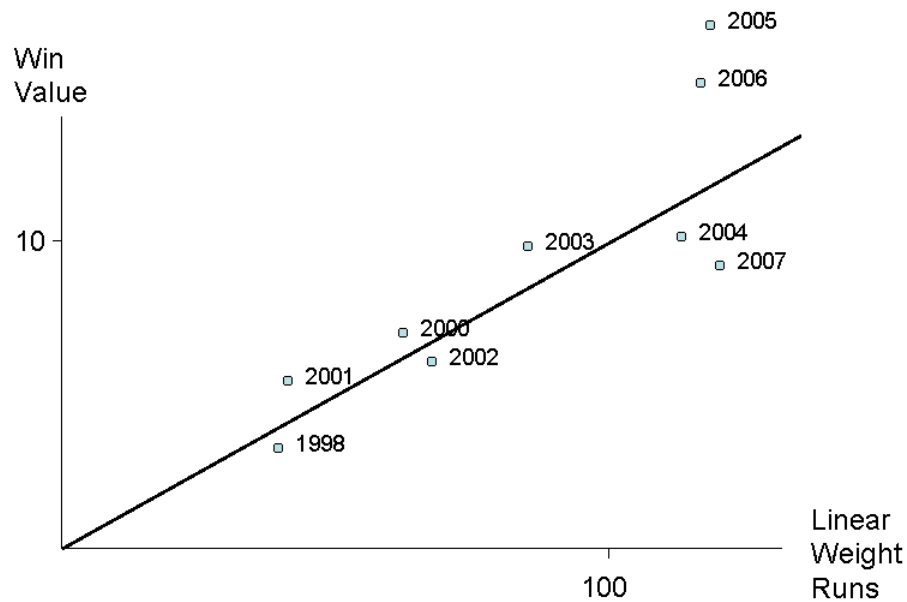
Group	OPS: Tense Situations	OPS: All Situations
All players, Elias “late and close” (15% of appearances)	.704	.715
891 players, tensest 10% of plate appearances (Mills)	.779	.771

Certainly factors other than tension affect the average outcome of “late and close” plate appearances. The best available pitchers will be used (depressing OPS in tense situations), but pinch hitters are used more and intentional walks occur twice as often (raising OPS in tense situations). However, we estimate that these factors roughly cancel one another, so that, as the tabulated OPS values themselves suggest, major league hitters’ is little if at all affected by game situation tension.

Further investigations require a working definition of pure clutch performance, and we follow other analysts in taking the approach that the Mills brothers (2) pioneered. The probable outcome of any game fluctuates at-bat by at-bat until the game ends, and this probability fluctuation constitutes a “win value”, a positive or negative contribution that a batter makes to his team’s chance of winning, as the outcome of each and every plate appearance. The sum of these win values over every appearance in a particular batter’s career becomes a direct measurement of that batter’s total win value. Division of that total performance by the number of plate appearances yields the batter’s average win performance.

However there is a general propensity for superior hitters to have superior win performances in any game situation. To isolate “clutchness”, a possible tendency for a hitter to be more effective in critical game situations than otherwise, this general propensity must be removed from that hitter’s total win value. Linear weights (also introduced by one of us) provide a thoroughly tested measure of situation-independent hitting skill. We can then compare these two measurements, win performances depending only on how the probability of victory changed after each plate performance, and linear weights depending only on the fundamental counting statistics. Correlation of the 897 players’ seasonal win performances now available with their corresponding linear weight totals yields this general propensity, as depicted by the diagonal line in Figure 1. The “clutchness” or “non-clutchness” of a player in a particular season then becomes the distance by which his total win performance in that season lies above or below the particular position on that diagonal determined by his linear weight total. Figure 1 also shows these clutchness results for each season of David Ortiz’s career. His reputation for clutchness was built in 2005 and 2006. His hitting in 2007 was even slightly better overall (the 2007 point is farther to the right of the graph) but less timely (the 2007 point is also lower), yielding a lower-than-average “clutchness” result.

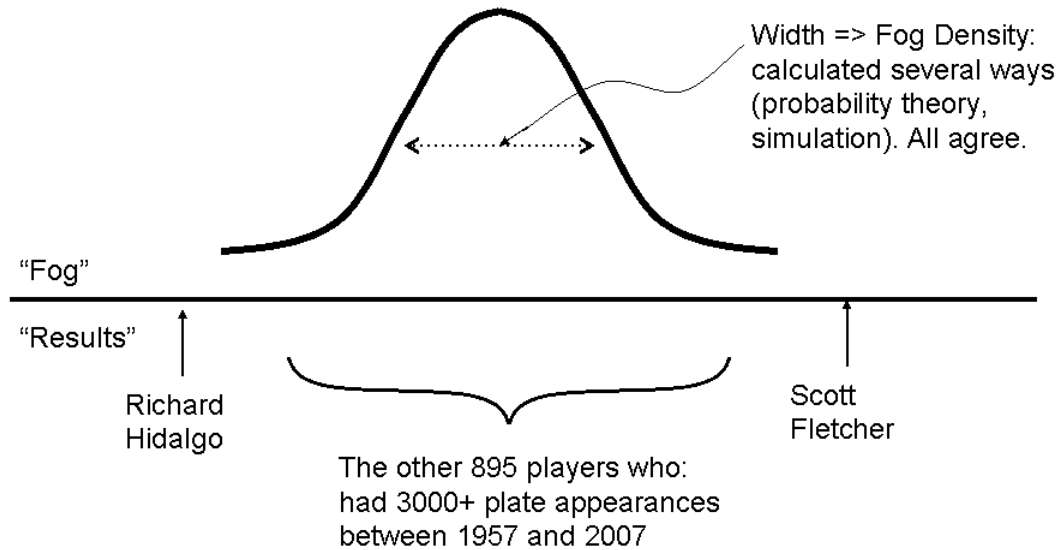
Figure 1. David Ortiz's Clutch Performances by Season



However, no one doubts that baseball is a “percentage game”, that because of unavoidable and therefore random variations such as “bad bounces”, good performances can produce bad results and vice versa. To what extent might David Ortiz’s historically superior clutch performances in 2005 and 2006 be matters of good fortune? This type of question motivated us to estimate the general magnitudes of these inherent and inevitable random variations in both win value and linear weight performances, using a variety of computer modeling approaches.

Another way of expressing this central concern is shown in Figure 2, which compares the “clutch performance results” we have just defined (at the bottom) with a curve representing our various estimates (which all agree within 10% or so) of these random variations in clutch performance (or, in Bill James’ language, the density of the “fog”). As the graph suggests and as probability theory confirms, there is nothing in the overall distribution of “clutch performance results” for the 897 most active batters over the last fifty years that proposes anything other than random fluctuations as their “cause”.

Figure 2: Comparing the “Fog” to the Clutch “Results”



Because of both the importance of the result and also the uncertainties of the result from any individual computer model, the largest component of this study involved devising and applying several independent computational approaches. The agreement among their results is encouraging.

Here are brief descriptions of these approaches:

- Direct simulation, or actual “play” of 1220 games with 18 similar batters and using the same outcome probabilities regardless of “game situation”.
- Comparing actual “clutchness” distributions with a performance distribution that is undoubtedly random, in this case the remainder from dividing the day of the month of the game by 10. actual seasonal win-performance-to-linear-weight comparisons by “clutchness”
- Considering that Ortiz’s two greatest “clutch” seasons were consecutive, perhaps suggesting that “clutch ability” is real but fleeting, and recapitulating a key component of the original study, we calculated the correlation coefficient between “clutch performance results” in all consecutive seasons of at least 250 BFP’s each for every one of those 895 most successful batters. The resulting r^2 value of .002 is as inconsequential as it looks. For comparison, the r^2 value for OPS itself is 0.43.
- Presumably batters feel pressure most strongly at the unproven starts of their major league careers. However the batting average of these 895 batters in their first hundred BFP, .260, is not very different from their collective career batting average.

Thus the results of the original study (1) are yet again confirmed, this time by every analysis we can devise and based mostly on fifty seasons of major league play. Over this period there is no convincing evidence that any fluctuation of any batter's performance in tense situations had any cause beyond random variation. Furthermore, because batting performance is unaffected by game situation overall, there would be no honor in being identified as a clutch hitter. Why shouldn't a major league batter give his best effort regardless of the game situation?

The conflict continues between the feeling of tension in critical game situations, expressed even by many of the most experienced and successful major league batters, and the actual outcomes.

References:

1) [Do Clutch Hitters Exist? By Richard D. Cramer \(SABR Baseball Research Journal 1977\)](#)

2) Mills, E. G.; Mills, H. D. Player Win Averages, A.S. Barnes, Cranbury, NJ (1970). (1970 data provided by Pete Palmer.)

3) <http://www.geocities.com/cyrilmorong@sbcglobal.net/ClutchLinks2.htm>

4) [Mr. Clutch Big Papi, Chipper, Pujols come through when it counts by Bill James. Underestimating the Fog by Bill James](#)

5) Tango, T. M.; Lichtman, M. G.; Dolphin, A. E. The Book: Playing the Percentages. TMA Press (2006).

6) Thorn, J.; Palmer, P. Total Baseball (1989).