An Age-old Question

© Mitchel Lichtman

Introduction

Recently there has been some renewed controversy over the best way to ascertain the peak age of offensive performance for the average MLB player, as well as what that peak age is. Traditional sabermetric wisdom says that peak age is around 27. JC Bradbury, in a recent study (http://www.informaworld.com/smpp/content~content=a909810792), concluded that peak age is around 29, with a gradual decline after that, at least for players who played for at least 10 years in the majors (I don't know if that is consecutive years) and amassed at least 5,000 PA. His methodology was to put everyone's career (who met those requirements) in a "blender" so to speak and then fit a least squares career trajectory around the data.

Phil Birnbaum, Tango Tiger, and other very smart and well-respected analysts have taken a great deal of exception to JC's study and his conclusions, especially as he appears to generalize his results to *all players*, as opposed to only those whose careers are similar to the players in his sample -- namely those with long careers and many PA. To tell you the truth, I am not qualified to judge the merits of JC's methodology and I don't fully comprehend his responses to the criticisms of Birnbaum, et al. If you want to read Phil's series of posts critiquing JC, start here and work your way backward:

http://sabermetricresearch.blogspot.com/2009/11/bradbury-aging-study-re-explained-part.html

You can do the same for JC's responses to Birnbaum and others:

http://www.sabernomics.com/sabernomics/index.php/2009/11/aging-and-selective-sampling/

As well, here are some good data and discussions on The Book blog about JC's assertions and about aging in general:

http://www.insidethebook.com/ee/index.php/site/the_ten_year_aging_curve/

http://www.insidethebook.com/ee/index.php/site/peak_age_by_length_of_career/

http://www.insidethebook.com/ee/index.php/site/article/the_latest_and_last_i_hope_jc_thread_on_this_site/

And finally, here is the link to the BBTF discussion about JC's latest article in the Huffington Post – the article sort of summarizes his findings on peak age:

http://www.baseballthinkfactory.org/files/newsstand/discussion/huffington_post_bradbury_is_your_favorite_free_agent_over_the_hill/

The basic argument against JC's findings is that it is not surprising that players who have had a long and prosperous career would have a higher peak age than all MLB players, including those who fizzled out, had a cup or two of coffee, were career part-time players, etc. The reason is this: If we assume that different players have different "true" aging curves (with different "true" peaks), which is probably a good assumption, those players who peak early and/or decline quickly are likely to be out of baseball before they amass a large number of career PA, such that JC's sample consists of players who tend to peak late and decline less rapidly than the average player.

In deference to JC, in one of his posts linked to above, he states that he gets the same results when he removes the 10-year career requirement. He says, "I estimated the impact of age after dropping the sample to 1,000 (minimum number of career PA) and eliminating the 10-years of play requirement, and I found that the peak age remained at 29."

Anyway, I will let you wade through the back-and-forth discussions and arguments among JC and the sabermetricians. I think that this issue contains a lot of muddy water.

The Delta Method for Creating Aging Curves

As some of you know, Tango, I and others have traditionally used the "delta method" of computing aging curves. You can read about that method as well as some of the results here on Tango's web site:

http://www.tangotiger.net/aging.html

http://www.tangotiger.net/agepatterns.txt

http://www.tangotiger.net/AgingSelection.html

Let me now briefly explain the basics of the "delta method", why it is a good method to determine aging patterns, and why it has some shortcomings.

The "delta method" looks at all players who have played in back-to-back years. Many players have several back-to-back year "couplets," obviously. For every player, it takes the difference between their rate of performance in Year I and Year II and puts that difference into a "bucket," which is defined by the age of the player in those two years. For example, say a player played in 2007 and 2008 and he was 25 years old in Year I. And say that his wOBA (or linear weights or whatever) in Year I was .320 and in year 2 it was .340. The difference is plus 20 points and we put that in the 25/26 (years of age) bucket. We do that for all players and for all of their back-to-back year "couplets." For example, if that same player played again in 2009 and his wOBA was .330, we would put minus 10 points (.330 - .340) into the 26/27 bucket.

When we tally all the differences in each bucket and divide by the number of players, we get the average change from one age to the next for every player who ever played in at least one pair of back-to back seasons in. So, for example, for all players who played in their age 29 and 30 seasons, we get the simple average of the rate of change in offensive performance between 29 and 30.

Now, here is the tricky part (and later, one of the problems with this simple method). Is it fair to weight every "couplet" (back-to-back seasons) equally, such as in my explanation above (where we just added up all the rate differences in each bucket and divided by the number of players)? Or should we give more weight to players who amass more PA in one or both seasons in a couplet? There seems to be merit to both approaches.

On the one hand, let's say that we wanted to compute the average weight change from one age to the next for all MLB players. We probably want to weight each couplet the same. There is no reason to weight one player more than another just because one player gets more playing time than another. On the other hand, when we talk about the "average" player in terms of performance, like BA, or OPS, we usually mean the "average player," weighted by PA or the like. It also depends on how you frame the question. If you ask, "What is the average BA in the major leagues?" the answer is obviously total hits divided by total AB, for all players combined. But if you ask, "What is the BA of the average MLB player?" would the answer be the same or would it be the simple average of everyone's BA, regardless of whether they had 10 or 500 AB?

To be honest, at first I thought that the correct answer in terms of aging curves was to weight every player equally. After all, if we want to know how a typical player ages, we might want to average everyone's aging pattern, and I could think of no compelling reason to give less weight to players with little playing time. However, after lots of consideration, I decided that using some kind of weighting procedure by playing time is the much better way to do it. After all, there are many players who get "cups of coffee" in the major leagues, are September call-ups only, are limited part-time players, etc. What if these players do not have the same aging curves as most players who get considerable playing time and have at least reasonably long careers in MLB? Do we want these fringe players to greatly affect our results?

If we decide to weight each couplet by playing time, to give more weight to those players who play more often, how do we do that? Do we use the combined number of PA (or the average of the two numbers) from both seasons? How about the lesser of the two numbers, or their harmonic mean (they are almost the same thing)? I don't really know which is correct. Traditionally, researchers like Tango and even myself have used the "lesser of the two numbers" to do the weightings. The harmonic mean, by the way, of N numbers, A, B, C, etc., is N / (1 / A + 1 / B + 1 / C, etc.).

Why have we used that weighting method? I am not really sure. For one thing, it reduces the impact of a large difference when that difference is "caused" by a small number of PA in one or both years. For example, let's say that one player at age 26 had a wOBA of .300 in 500 PA and at age 27, he only got 4 PA (say he was injured for the rest of the year) with a wOBA of 0. Now we have a difference of minus 300 points for that couplet. If we are weighting by the average of 500 and 4 (252), we might significantly affect our entire result because of that one .300 point outlier caused by a sample of only 4 PA in Year II.

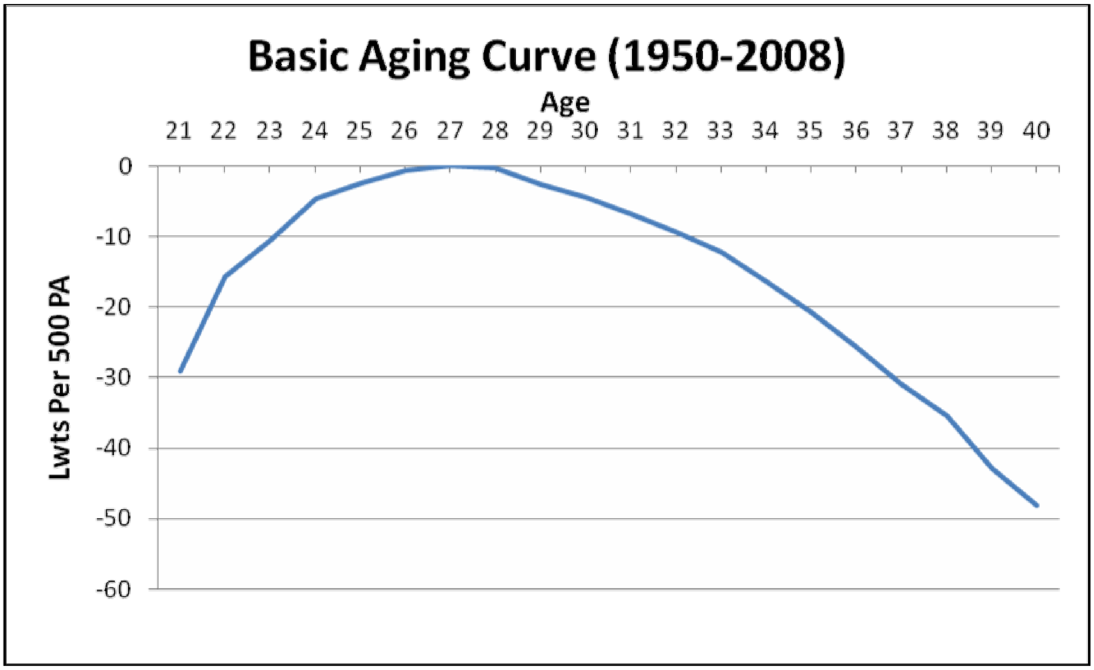
On the other hand, if we weight that 300 point difference by the harmonic mean of 500 and 4 (8), or just 4 (the "lesser"), the weight is so small that we might as well not even use that data point. Plus, given a large enough sample and given the fact that there is no bias in those large differences created by a very small number of PA in one or both years, such anomalous differences should "even out" in our sample and we should have little to worry about. And even if they don't "even out," if we have enough players, even a single plus or minus 300 points with a large weighting (like 252) shouldn't be all that problematic. As it turns out, whichever method of weighting we use makes very little difference in terms of the final results.

The following tables use the delta method, with each couplet weighted by the average of the two PA. That is the standard I am going to use for the remainder of this article. The numbers in the penultimate column represent the average change in performance in linear weights per 500 PA from one age to another for all players in MLB who have played in at least one pair of consecutive seasons at those ages.

Table I: Average change in offensive performance from one age to the next (1950-2008)

Age Year I/Year II	N	Year II – Year I, Lwts Per 500 PA	Cumulative Difference
20/21	142	-4.3	-29.2
21/22	366	13.4	-15.8
22/23	727	5.1	-10.7
23/24	1224	5.9	-4.8
24/25	1776	2.3	-2.5
25/26	2104	1.8	-0.7
26/27	2140	0.7	0.0
27/28	2088	-0.4	-0.4
28/29	1953	-2.4	-2.8
29/30	1775	-1.8	-4.6
30/31	1583	-2.3	-6.9
31/32	1357	-2.6	-9.5
32/33	1142	-2.8	-12.3
33/34	923	-4.1	-16.4
34/35	733	-4.5	-20.9
35/36	547	-4.9	-25.8
36/37	390	-5.1	-30.9
37/38	262	-4.6	-35.5
38/39	171	-7.5	-43.0
39/40	101	-5.2	-48.2

Chart I: Aging curve, using the "delta method" weighted by the average of the two PA (1950-2008)



As you can see, the peak age is 27-28 (more of a plateau from 26 to 28). After that there is a gradual decline of around 2-3 runs a year until age 33, then a steeper decline (4-5 runs a year) until age 38, after which the decline is steeper still.

Comparing Eras (pre and post-1980)

There has been some suggestion in the research that the aging curve is substantially different in the modern era, due to advances in medicine, higher salaries, and perhaps PED use. Let's split the data up into two arbitrary eras, pre and post-1980.

Table II: Average change in offensive performance from one age to the next (1950-1979)

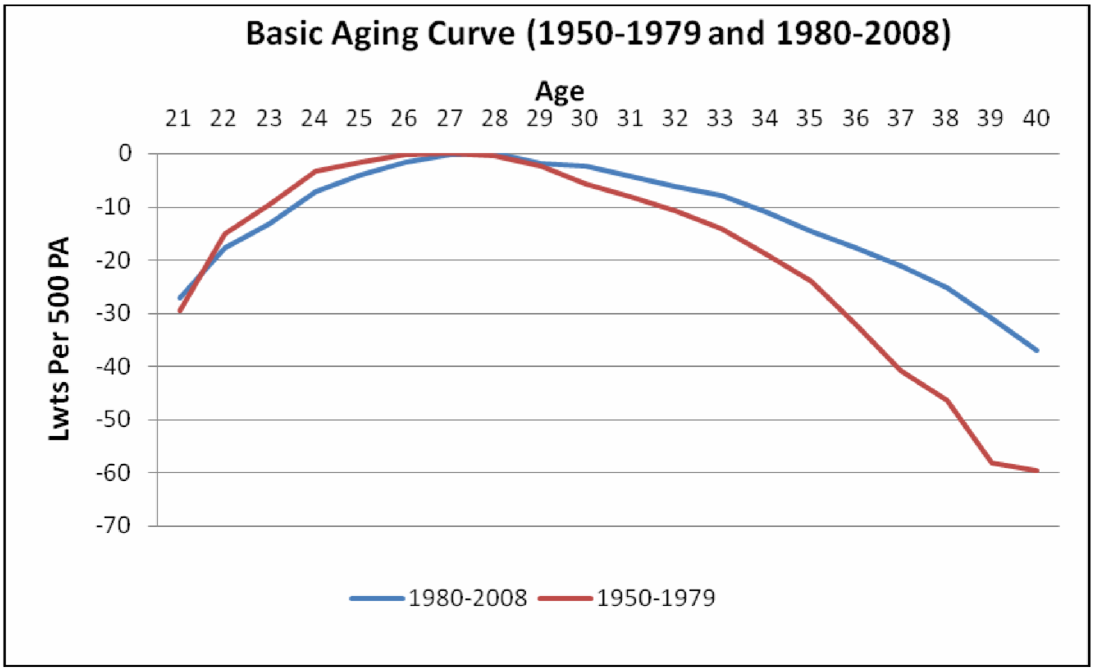
Age Year I/Year II	N	Year II – Year I, Lwts Per 500 PA	Cumulative Difference
20/21	115	-7.7	-29.6
21/22	226	14.4	-15.2
22/23	398	5.6	-9.6
23/24	588	6.3	-3.3
24/25	783	1.7	-1.6
25/26	880	1.4	-0.2
26/27	868	0.2	0.0
27/28	852	-0.5	-0.5
28/29	793	-2.0	-2.5

29/30	711	-3.3	-5.8
30/31	627	-2.3	-8.1
31/32	538	-2.6	-10.7
32/33	426	-3.4	-14.1
33/34	342	-4.9	-19.0
34/35	272	-5.1	-24.1
35/36	189	-8.0	-32.1
36/37	124	-8.6	-40.7
37/38	79	-5.7	-46.4
38/39	55	-11.7	-58.1
39/40	31	-1.5	-59.5

Table III: Average change in offensive performance from one age to the next (1980-2008)

Age Year I/Year II	**N**	**Year II – Year I, Lwts Per 500 PA**	**Cumulative Difference**
20/21	27	18.4	-27.1
21/22	140	9.4	-17.7
22/23	329	4.6	-13.1
23/24	636	6.0	-7.1
24/25	993	3.0	-4.1
25/26	1224	2.4	-1.7
26/27	1272	1.5	-0.2
27/28	1236	0.2	0.0
28/29	1160	-2.0	-2.0
29/30	1064	-0.4	-2.4
30/31	956	-2.0	-4.4
31/32	819	-1.8	-6.2
32/33	716	-1.7	-7.9
33/34	581	-3.1	-11.0
34/35	461	-3.7	-14.7
35/36	358	-3.0	-17.7
36/37	266	-3.4	-21.1
37/38	183	-4.0	-25.1
38/39	116	-5.9	-31.0
39/40	70	-5.9	-36.9

Chart II: Comparing the aging curve in two eras (1950-1979 and 1980-2009)



Indeed, in the 1908-2008 era, peak age is a little later (28 versus 27) and the decline after that is significantly more gradual. Today's 35 is the equivalent of yesterday's 33, and players at 40 in the modern era are more productive than those at 37 prior to 1980. As you can also see from charts II and III above, there are many more players in their 30's in the post-1980 era.

<u>Survivor Bias</u>

As many of you know, and JC points out in one of his responses, the "delta method" suffers from something called "survivor bias." What is "survivor bias" and how does it affect the aging curve and peak age of offensive performance? At every age, there are a class of players who play so badly in any one season that they don't play at all (or very little, which I call "partial survivor bias") the following year. (Obviously there are also players who *don't* play poorly in Year I but also don't play in Year II.)

These players tend to be unlucky (and usually untalented of course, at least at that point in their careers) in Year I. Therefore the rest of the players who do play in Year I *and* the following year tend to have been lucky in Year I. This is survivor bias. Any player who "survives" to play the following year, especially if they rack up lots of PA, whether they are good, bad, or indifferent players, true-talent-wise, will tend to have gotten lucky in Year I. In Year II, they will revert to their true talent level and will thus show a "false decline" from the one year to the next.

Obviously players who do not survive from one year to the next are not in our sample. If they were (if they were allowed to play the next year), those players would tend to show a "false" improvement and thus would "balance out" those players showing a false decline.

If you are confused, here is simplification of the process, which I think will help you to understand the principle: Say there are 100 marginal players in MLB at any one age, and their true talent BA is .220. Let's say that they get 200 AB in Year I and half of them end up with a BA of .180, and the other half, .260. Since in 200 PA, their "sample" BA will be all over the place (the SD of BA in 200 AB is 29 points), but centered on .220, this is a plausible, albeit simplified, scenario. And let's say that only those 50 players who hit .260 were allowed to play the next year. After all, if a marginal (or old) player, talent-wise, has a very unlucky season, he is often benched or retires the following year.

So now we have 50 players remaining who hit .260 in Year I and are allowed to play and amass another 200 AB in Year II. What will they hit, on the average, in Year II, if they neither improved nor declined in true talent (say they were around age 27)? .220. After all, we already said that they were true talent .220 hitters. But what would the delta method say about them? It would say that they declined by 40 points (.260 in Year I and .220 in Year II)!

If all 100 players were allowed to play the next year, and they all hit .220, as they should, then half would decline by 40 points, and half would improve by 40 points (.180 in Year I and .220 in Year II), and the net change for all 100 players according to the delta method would be zero, as it should be. That is why survivor bias produces more decline (and less improvement) than it should at every age interval.

Also keep in mind that survivor bias is merely a subset of "partial survivor bias," whereby some players who are unlucky in one year get limited playing time the following year. When we weight by the lesser of the two PA or the average of the two PA, partial survivor bias will also create a "false decline" for all age groups. In the above example, if the true .220 batters who hit .180 in Year I were allowed to get 50 PA in Year II while the remaining batters were allowed to rack up 200 PA in Year II, we would have the following average "delta," if we used the average of the two PA for the weighting:

((-40 x 200) + (40 x 125)/ 325)

That is a 9 point "false" decline.

Any kind of correlation between performance and future playing time, either within a season or from season to season, is going to cause problems with our results when using the delta method to compute aging curves. The most dangerous and prevalent kind of bias is partial or full survivor bias, which causes a "false" decline at all age intervals.

How large is the class of players who do not survive from one season to another? From age 20 to 24, for every hundred players that play back-to-back seasons (and are included in our sample), there are 7 more who play for one year and don't come back at all the following season (and are not in our sample). From age 25 to 28, it is 11 players who do not survive for every 100 who do. From age 28 to 35, it is 16/100 and after age 25, it is 30/100! Almost 25% of all players after the age of 35 have played their last year (or at least don't play the very next year) in MLB!

How bad is their performance in Year I, such that they are not permitted to play the following year? From age 20 to 24, these non-survivors average -34 runs per 500 PA in Year I. It is no wonder that there is no Year II for these players! Of course if there *were* a Year II, they would likely hit quite a bit better than -34, due to regression towards the mean. A player who is allowed to get at least a few MLB PA is likely not a true -34 hitter. At age 32-35, players averaged around -18 runs in their likely final years. Obviously these were probably very good players in their prime and in their careers (otherwise they would not have made it that far). Again, if they were allowed to play one more year, even at such an advanced age, on the average (not everyone) they probably would have played a little better in Year II than in Year I (although probably not so much as with the younger players).

Keep in mind that it is not the absence of these players that is causing a problem. It is the fact that the remaining players – all the players in our sample – tend to be (on the average) slightly lucky in Year I and thus will show a "false decline" in Year II, over and above the "real" aging-related increase or decrease from one year to the next.

Here is some data on all players who played at age X but not at age X+1. For each interval, for example, 25/26, the player played at the first age but not at the second.

Table IV: Players who played in Year I but not in Year II (1950 to 2008)

Age	N	Year I PA	Year I Lwts Per 500 PA	Career PA	PA Last 3 Years	Career Lwts per 500	Last 3 Years Lwts per 500
20/21	19	45	-32	46	46	-32	-32
21/22	24	14	-42	26	26	-34	-34
22/23	53	43	-29	76	76	-27	-27
23/24	90	49	-37	84	84	-32	-33
24/25	178	41	-25	94	88	-25	-25
25/26	231	55	-28	130	117	-23	-23
26/27	321	52	-26	183	144	-20	-21
27/28	324	63	-26	315	208	-18	-20
28/29	339	62	-25	405	235	-15	-17
29/30	320	71	-23	637	259	-12	-16
30/31	283	80	-20	961	360	-11	-14
31/32	274	77	-25	1351	373	-10	-15
32/33	251	92	-19	1724	433	-8	-13
33/34	246	88	-20	2153	449	-8	-12
34/35	232	102	-18	2930	505	-5	-11
35/36	210	106	-14	3323	508	-1	-8
36/37	150	107	-17	4077	553	-1	-7
37/38	127	119	-12	4597	586	3	-4
38/39	97	148	-13	5398	664	2	-4
39/40	69	128	-11	5329	594	3	-5

As you can see, in almost every case, even for the younger players, their "last" year (Year I) performance is worse than the average of their last 3 years (including Year I). Presumably if they were to play the following year, their numbers would improve, after regressing their last 3 years toward some mean (sort of a rough Marcel projection).

So how can we account for the fact that these “unlucky” players would have been in our sample save for the fact they weren’t allowed (or able) to play the following year? We can assume that they got some playing time in Year II, “pencil in” a conservative projection for them, and then re-do (with these players now included in our sample) one of our delta methods. I say a “conservative projection” because the fact that they were not allowed to play in Year II, for whatever reason, implies that they were not very good players, true talent-wise, and we certainly don’t want to regress their career or last 3-year numbers toward that of a “league average” player. Clearly these players do not belong to the population of the typical MLB player, *per se* (even though technically they are MLB players of course).

Here is the same chart as above. This time I added a column, which is the average Year II projection for the pool of players at each age group. The projection is their last 3 years lwts per 500 PA, weighted by year (3/4/5) added to 500 PA of league average lwts for that age minus 5. In other words, I am regressing their last 3 years lwts (weighted) towards 5 runs lower than a league average player for that age. That “minus 5 runs” is the downgrade I used to generate a “conservative” projection. The final step in the projection is to add in an age adjustment. I realize that we don’t *know* the exact adjustment until we re-run the data, but it is not that critical to get it right.

Table V: Players who played in Year I but not in Year II, including their projection for Year II

Age	N	Year I PA	Year I Lwts Per 500 PA	Career PA	PA Last 3 Years	Career Lwts per 500	Last 3 Years Lwts per 500	Year II Projection
20/21	19	45	-32	46	46	-32	-32	-20
21/22	24	14	-42	26	26	-34	-34	-13
22/23	53	43	-29	76	76	-27	-27	-14
23/24	90	49	-37	84	84	-32	-33	-14
24/25	178	41	-25	94	88	-25	-25	-10
25/26	231	55	-28	130	117	-23	-23	-15
26/27	321	52	-26	183	144	-20	-21	-13
27/28	324	63	-26	315	208	-18	-20	-12
28/29	339	62	-25	405	235	-15	-17	-13
29/30	320	71	-23	637	259	-12	-16	-13
30/31	283	80	-20	961	360	-11	-14	-15
31/32	274	77	-25	1351	373	-10	-15	-16
32/33	251	92	-19	1724	433	-8	-13	-13
33/34	246	88	-20	2153	449	-8	-12	-15
34/35	232	102	-18	2930	505	-5	-11	-14
35/36	210	106	-14	3323	508	-1	-8	-12
36/37	150	107	-17	4077	553	-1	-7	-12
37/38	127	119	-12	4597	586	3	-4	-8
38/39	97	148	-13	5398	664	2	-4	-12
39/40	69	128	-11	5329	594	3	-5	-13

Even though the projections are in all cases better than their last year (Year I) of performance, they are still close to replacement level (with an exception at 37/38 for some reason), so you can see why MLB teams were not too enthusiastic about allowing them to play (at the major league level) anymore.

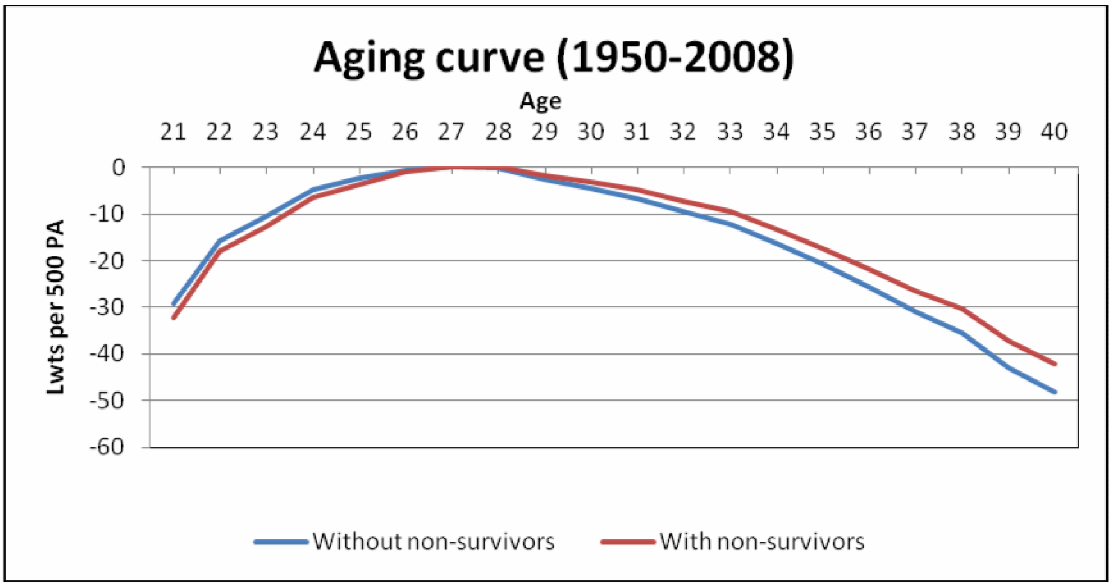
So now that we have taken care of our survivor bias problem, all we have to do is redo our aging curve by using one of our delta methods. Before we look at those numbers, you may be wondering, "What am I going to use for the number of PA in Year II for those players who didn't actually play in Year II?" I am going to use their Year I PA, though not for any compelling reason. It is going to generally be a small number, as you can see from column 3 (the average number of PA in Year I) in the above chart.

Table VI: Aging data, including non-survivors (1950 to 2008)

Age	**N**	**Year II – year I, Lwts Per 500 PA**	**Cumulative difference**
20/21	161	-4.1	-32.3
21/22	390	14.2	-18.1
22/23	780	5.3	-12.8
23/24	1314	6.1	-6.7
24/25	1954	2.8	-3.9
25/26	2335	2.1	-1.1
26/27	2461	1.1	0.0
27/28	2412	0.0	0.0
28/29	2292	-1.9	-1.9
29/30	2095	-1.3	-3.2
30/31	1866	-1.8	-5.0
31/32	1631	-2.3	-7.3
32/33	1393	-2.3	-9.6
33/34	1169	-3.8	-13.4
34/35	965	-4.1	-17.5
35/36	757	-4.4	-21.9
36/37	540	-4.7	-26.6
37/38	389	-3.9	-30.5
38/39	268	-6.8	-37.3
39/40	170	-4.9	-42.2

In the graph below, you can see the difference between the aging curves when we do not include the non-survivors, and when we do – i.e. with and without a correction for survivor bias.

Chart III: Aging curve with and without non-survivors (1950-2008)

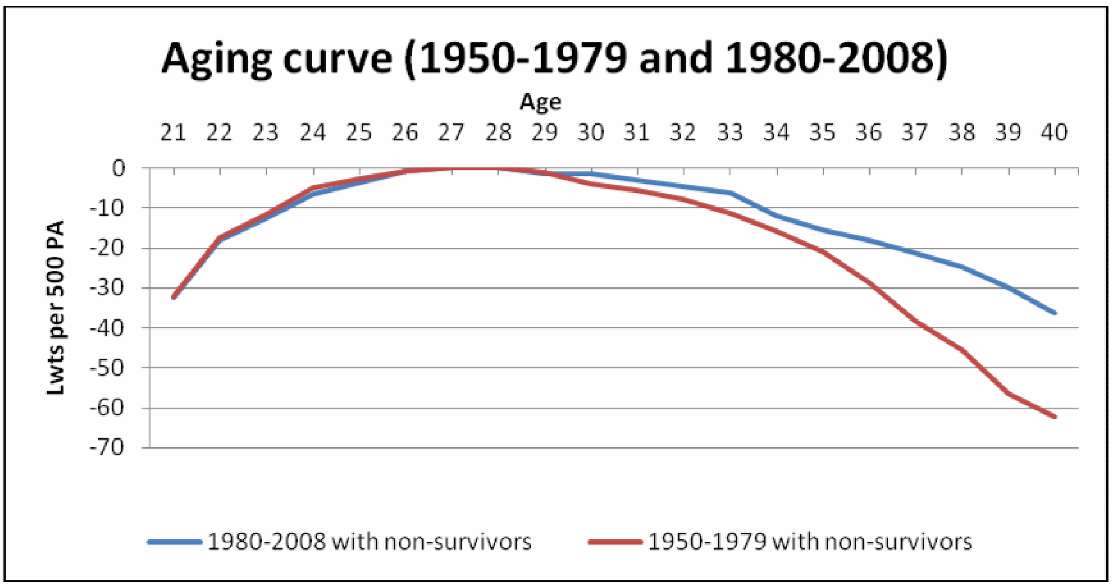


Once we account for survivor bias by creating a “phantom” Year II for the non-survivors in each of the age intervals, the average aging curve shifts slightly to the right and the decline after the peak is a little flatter. There isn’t a large difference between the two aging curves, however.

<u>Comparing Eras After Correcting for Survivor Bias</u>

Let's compare again the pre and post-1980 eras, after accounting for survivor bias. Without that adjustment, we found a slightly later peak and a much more gradual post-peak decline in the modern era.

Chart VI: Comparing the aging curves (with non-survivors) in two eras (1950-1979 and 1980-2008)



We find the same thing when comparing the two eras after adjusting for survivor bias – the curve is shifted slightly to the right, and the post-peak decline is much less steep in the modern era. So it appears that advances in medicine, better training, higher salaries, and perhaps PED use in the modern era, changes peak age only slightly, but significantly affects post-peak decline, keeping players in the major leagues far longer than in previous eras. In the post-1980 era, a player at age 38 is expected to have lost 23 runs off his peak. Prior to that, players on average lost 46 runs off their peak by the time they reached the age of 38. Clearly in the old days, only superstars remained in baseball into their mid to late 30's and beyond.

<u>Players with at least 10 years and 5000 Career PA</u>

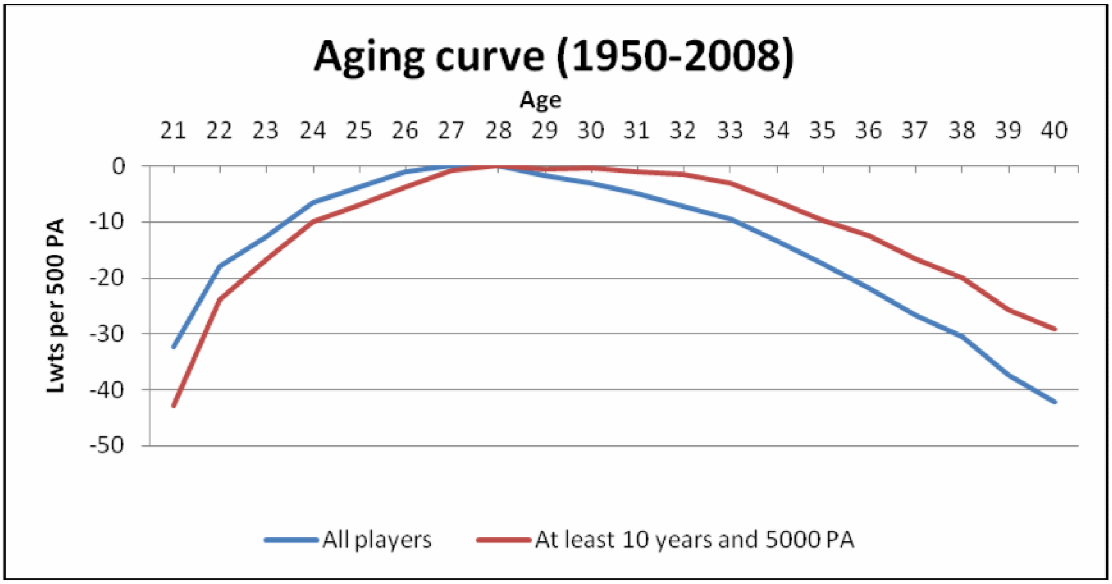
What if we do the exact same thing as above (use the delta method to construct an aging curve, and adjust for survivor bias), but we only use players who have played at least 10 years in the majors with at least 5000 career PA (10/5000), similar to the sample that JC used in his study? We might expect a later peak age and perhaps a more gradual post-peak decline. Presumably, when we use the delta method with this kind of a sample (player with long careers), our results are more representative of the "average aging curve" for this type of player, as we are not "piecing together" careers of various lengths, as we are when we include all players. In other words, our results should look something like JC's.

Table VIII: Players who have played for at least 10 years with at least 5000 career PA (1950-2008)

Age	**N**	**Year II – year I, Lwts Per 500 PA**	**Cumulative difference**
20/21	65	-0.8	-42.8
21/22	146	18.9	-23.9
22/23	247	7.1	-16.8
23/24	344	6.7	-10.1
24/25	424	3.1	-7.0
25/26	452	3.1	-3.9
26/27	465	3.1	-0.8
27/28	476	0.8	0.0
28/29	479	-0.6	-0.6
29/30	478	0.1	-0.5
30/31	478	-0.7	-1.2
31/32	472	-0.5	-1.7
32/33	459	-1.6	-3.3
33/34	439	-3.0	-6.3
34/35	411	-3.4	-9.7
35/36	367	-2.9	-12.5
36/37	306	-4.2	-16.7
37/38	236	-3.3	-20.0
38/39	172	-5.7	-25.7
39/40	116	-3.4	-29.1

Let's compare the aging curve for these players to that of all players, again, as determined by the delta method, weighted by the average of the two PA.

Chart VII: Comparing the aging curves of all players and only those with at least 10 years and 5000 PA (1950-2008)

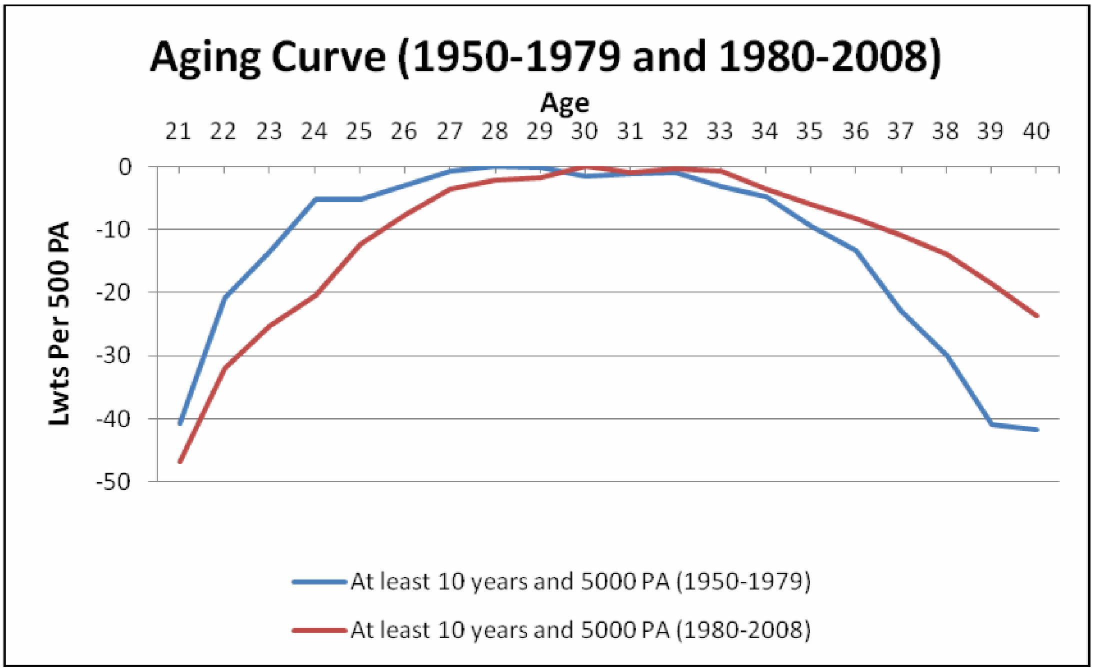


Indeed, once we restrict our players to those with at least 10 years and 5000 PA, as in JC's sample, the aging curve looks quite different. While technically the peak age is 28, there is a plateau from 27 to 30. After that, there is a slight decline until age 32 or 33 and a gradual decline thereafter.

Comparing Eras for Players with 10 Years and 5000 PA

What if we split our sample again into the two eras – pre and post 1980? We are only looking at players with at least 10 years and 5000 PA.

Chart VIII: Comparing the aging curves of two eras, for only those players with at least 10 years and 5000 PA (1950-1979 and 1980-2008)



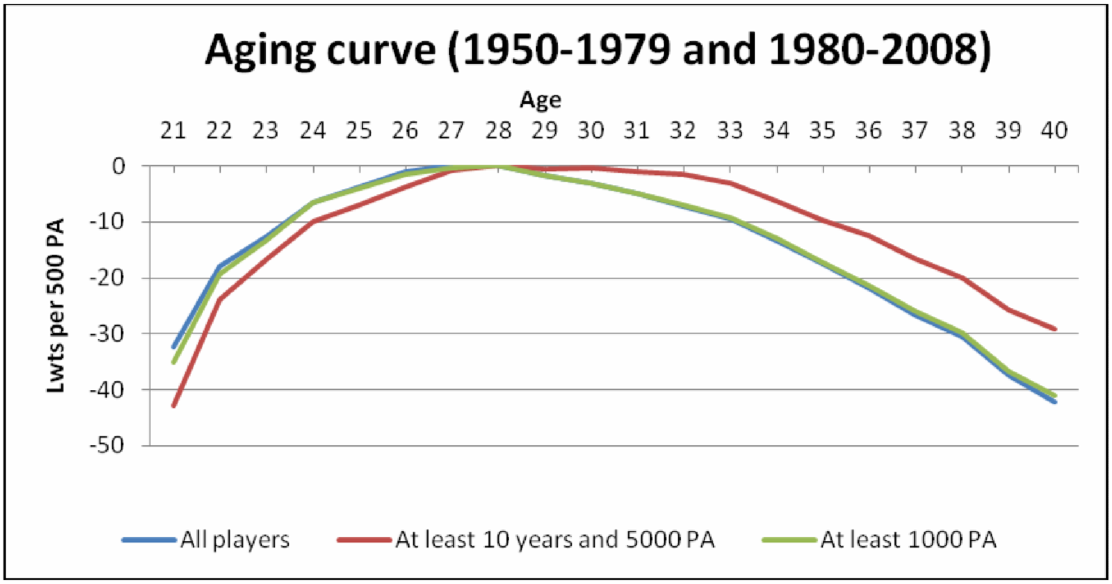
In the pre-1980 era, for players who have at least 10 years and 5000 PA in MLB, the aging curve is pretty symmetrical around a plateau stretching from around 27 or 28 to 32. From 21 to 27 or 28 is almost a mirror image of 32 to 38. In the modern era, the player with a long and prosperous career peaks at 30, stays relatively stable until age 33, declines gradually (around 2-3 runs per year) after that until age 38, and then declines by around 5 runs per year after that.

Players with a Minimum of 1000 Career PA

Finally, what if we reduce the requirements to 1000 minimum career PA with no minimum number of seasons? We are essentially eliminating all those players who come up for a cup of coffee or two, or are career September call-ups or ultra part-timers only. JC mentions that when he does this, he comes up with the same results as with the more restrictive sample (10/5000).

The following chart compares all players to those with 1000 career PA or more to those with a minimum of 10 years and 5000 career PA.

Chart VII: Comparing the aging curves of all players, those with at least 10 years and 5000 PA, and those with 1000 career PA or more (1950-2008)



The aging curve of the sample of players who have at least 1000 career PA is almost exactly the same as that of all players. This is not too surprising as even a part-time player needs only 4 or 5 years to accumulate 1000 PA. Given the data above, I am skeptical of JC's claim that he got the same results (using his "least squares" method) when he increased his sample size by requiring only 1000 career PA (and no minimum number of years). Of course, it is likely that as you increase your requirements and include only those players with longer and longer careers and more total PA, you will see the aging curves shift to the right and flatten out after their peaks. In fact, if I bump the requirement to at least 5 years and 2000 PA, the curve moves slightly to the right (with a definite peak at 28) with a more gradual decline after age 31.

Summary and Conclusions

JC Bradbury, a Ph.D. econometrician and college professor who has written several books on the economics of baseball and hosts a website/blog on the same, www.sabernomics.com, recently published a scholarly paper in which he concluded that a certain sample of players peak offensively at age 29, and decline gradually after that – among other things. He appears to have generalized that conclusion, at least implicitly, to "all players" and not just those whose careers are similar to those in his research sample. He goes on to claim that when he reduced the requirements to only 1000 career PA and no minimum number of seasons, his results were the same. I am skeptical of that claim based on the research that I have done using the "delta method" corrected for survivor bias.

The "delta method" is one way to construct an aging curve for offensive performance (JC's uses another method – plotting each player's career trajectory and then using a least squares/best fit model to come up with a composite trajectory), although it is debatable which way is best and which method answers what questions. The question, "What does the aging curve of the average MLB player look like and what is the peak age of that player?" is not an unambiguous question.

It is also not clear what weighting to use for the performance differentials when using the delta method. I have explained three methods – weighting everyone equally, using the "lesser of the two PA," and using the "average of the two PA." I think that the latter two are the best methods, and I prefer the last one, although I think there is very little difference between the two.

I have also explained the problem of survivor bias, an inherent defect in the delta method, which is that the pool of players who see the light of day at the end of a season (and live to play another day the following year) tend to have gotten lucky in Year I and will see a "false" drop in Year II even if their true talent were to remain the same. This survivor bias will tend to push down the overall peak age and magnify the decrease in performance (or mitigate the increase) at all age intervals.

One way to account for and overcome survivor bias is to imagine that these players actually played another year and that their performance in Year II was at their true talent level. We arrive at that true talent level by doing a basic Marcel-like projection. Since these players are truly marginal players (and/or at the end of their careers), when we do the regression toward the mean, we use a very conservative mean (5 runs worse than we would ordinarily use).

When all is said and done, and we use the delta method and weight each interval by the average of the two PA, and we account for survivor bias, we pretty much find that which we already knew – players peak at around age 27-28, decline gradually until around age 35, and then decline a little more rapidly after that. We also find that if we split the data up into 2 eras, pre and post-1980, there is a slightly higher peak in the latter years (28, as opposed to 27-28 prior to 1980), similar declines until the mid to late 30's, and then a much shallower decline after that in the modern era. That should not be all that surprising given the advances in medical care, better training, higher salaries, and perhaps the use of PED's.

When we use the same methods, but only for players who played at least 10 years in the majors, we see, not surprisingly, significantly different trajectories. The aging curve is shifted to the right, there is a plateau from age 27-30, and a more gradual decline after that. My results are presumably not unlike those of JC. However, when I use the delta method (accounting for survivor bias) on a sample of players with a minimum of 1000 career PA only, the resulting curve is not much different from that of all players – perhaps a slight shift to the right.

So while I think that we have shed some more light on the subject of peak age and trajectories for the average MLB player (as well as some subsets), I think the issue still contains a lot of muddy water. We don't really know what the question is or even what it means once we articulate it. Practically speaking, we are usually interested in the (estimated) peak age and trajectory for individual players and subsets of players in MLB, for projection and salary purposes. To that end, it is probably more useful to frame these questions in a much more specific fashion, such as, "Given that a player has already played 5 years full-time in the majors, and is a fast and wiry player (and given his trajectory thus far), what is his likely peak age, and what will his future trajectory look like?" Those are the questions we really need to answer, and we probably shouldn’t concern ourselves so much with the nebulous concept of the “average MLB player’s aging trajectory” – whatever that even means.