The moderating impact of positions on the curvilinear relationship between age and performance

By Dante McClure

Abstract:

This project was intended to analyze the impact that Age had within the variables of performance regarding various statistics such as Points, Rebounds, and Assists. We hypothesized that factors such as increasing age would have a greater impact on certain statistics than others, such as Points or Blocks. We used data gathered from MVP Candidates since the 1950s, and used linear and nonlinear regression to extrapolate variables such as peak performance age, and trends regarding player performance on the independent variable of age. The results gave us an indication that defense-related performance would take a larger toll as a player got older when compared to its offensive counterpart. Throughout the project, we had to make sure that the data would be of use to NBA executives or managers, and not just any random statistic. When looking at other NBA statistics websites, we found that not much analysis was done on the true impact of age regarding specific skill categories, and thus we have tried to narrow down each graph to variability specific enough to where we can accurately test our hypotheses and come to decisive conclusions.

Introduction

When NBA players may reach their peak performance has always been an interesting question among not only NBA analysts but also among many NBA fans. However, due to many covariates that can play a role in the performance of NBA players, there has not been, other than some anecdotal evidence, a clear answer to this question. A systematic effort to examine the relationship between age and performance will help NBA franchises deal with players' ability and help negotiate effectively with their players.

As players get older, do they perform better or worse? This is not a simple question because the relationship is usually not linear. For example, if we assume that the relationship is positive and linear, that means as players get older, their performance will continue to increase. Knowing that players on

average retire at around the age 35, it is clear that the positive, linearity cannot be assumed. Does, then, performance decrease as players get older? This negative linear assumption can also not hold true because their performance increases until they reach certain age after which their performance may level off and start to decrease. Therefore, I postulate that the relationship between age and performance in the NBA among the players is nonlinear in that performance may increase as players get older up to a moderate level of age before it starts to decrease when they get much older. Given the situation, I hypothesize that the relationship between age and performance of NBA players is curvilinear in that performance will increase initially as age increases up to a moderate level, after which the performance will start to decrease as players get older. More importantly, I hypothesize that the curvilinear effect of age on performance will vary by positions. For example, regardless of the age, the number of assists is unlikely to vary for players who play center positions no matter how young or how old they get. Below is a summary of the research questions that will be explored in this report.

Research question 1: At what age do NBA players reach their peak performance?

Research question 2: Have there been statistically significant differences in performance since the hand-checking was no longer allowed? (Note: Hand-check rule change was implemented in 2004)

Research question 3: Does the relationship between age and peak performance vary by positions?

Methods

Data

A secondary dataset was obtained from NBA.com, which contained performance indicators of NBA players over their full career including such indicators as positions, years in the NBA, age at the time of

draft, games played, field goal made, field goal percent, free throws made, total rebounds, 3 pointers made, and many others. From the original 5,463 cases in the data, after removing outliers and invalid values, a total of 4,893 cases were retained for the analyses. Due to too many players in the NBA not playing the minimum number of games each season, I limited the sample to NBA players who were considered to be NBA's MVP at least once. A total of 281 NBA players were retained as a result. 22 percent of the cases were in point guard positions, followed by 16.6% by shoot guards, 21.4% small forwards, 20% power forwards, and 20% centers. The number of years that the NBA players played ranged up to 23 years with age ranging between 18 and 43. The highest number of games played by an NBA player was 1,429.

POS_N

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 PG	835	17.1	21.9	21.9
	2 SG	634	13.0	16.6	38.5
	3 SF	817	16.7	21.4	59.9
	4 PF	765	15.6	20.0	79.9
	5 C	768	15.7	20.1	100.0
	Total	3819	78.1	100.0	
Missing	System	1074	21.9		
Total		4893	100.0		

Analytical procedure

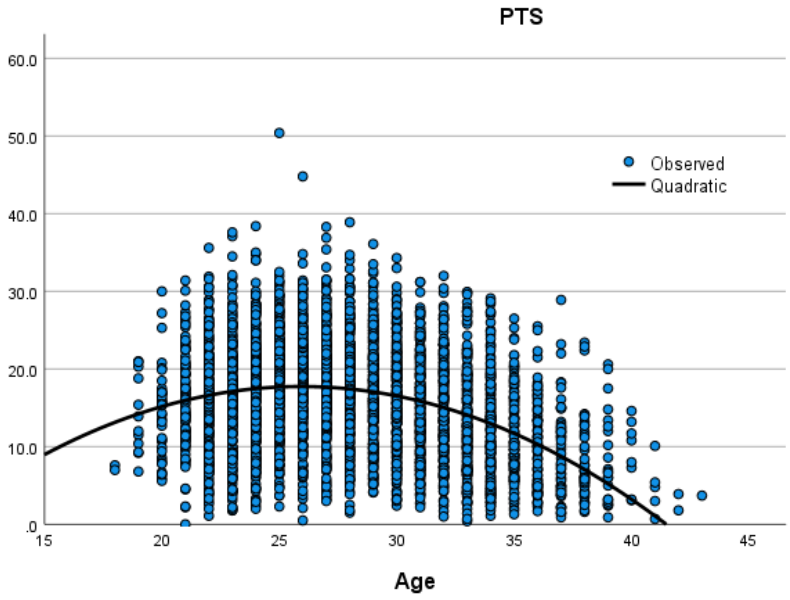
I used a nonlinear regression model to examine whether the relationship between age and performance can be better characterized as quadratic as opposed to linear. I ran the nonlinear regression model in a hierarchical fashion such that the original term for age was entered in the model, and quadratic term for age was then entered in the second model. This will test whether the hypothesis that quadratic effect better explains the outcome than the linear effect will be supported. Furthermore, to examine whether the curvilinear effects of age on performance vary by positions, a moderated curvilinear regression model was used to test the moderating impact.

Results

The preliminary curvilinear regression model showed that the relationship between age and performance in terms of average points made was better characterized as quadratic. As shown in the graph, until players reach moderate level of age, the performance initially increases, which starts to decrease when they get older. Based on the calculation derivative, I found that the peak performance in terms of average points made was reached when players were 26.1 years of age.

Coefficients

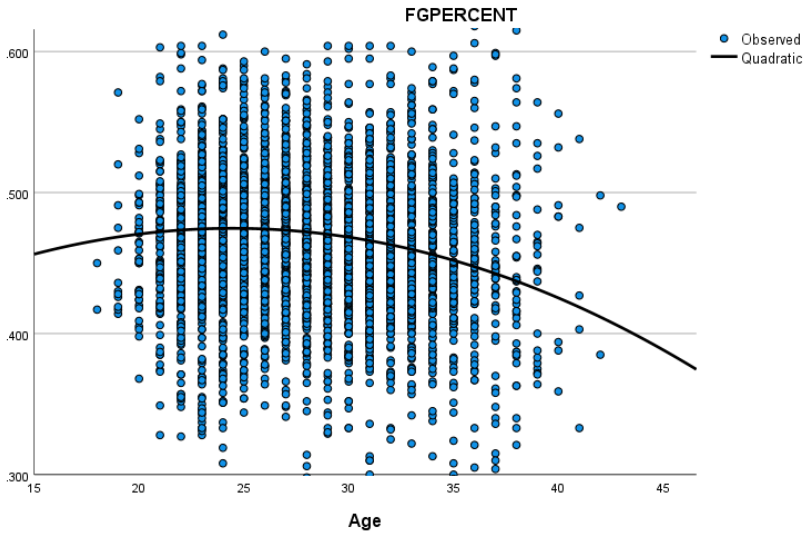
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	3.807	.270	2.407	14.078	<.001
Age ** 2	-.073	.005	-2.693	-15.747	<.001
(Constant)	-31.609	3.859		-8.190	<.001



Additional nonlinear regression models were tested using other performance indicators. For the age-field goal percent relationship, the quadratic equation was statistically significant (B= -.001p = .007). However, from a practical standpoint, the actual effect was minimal as shown below:

Coefficients

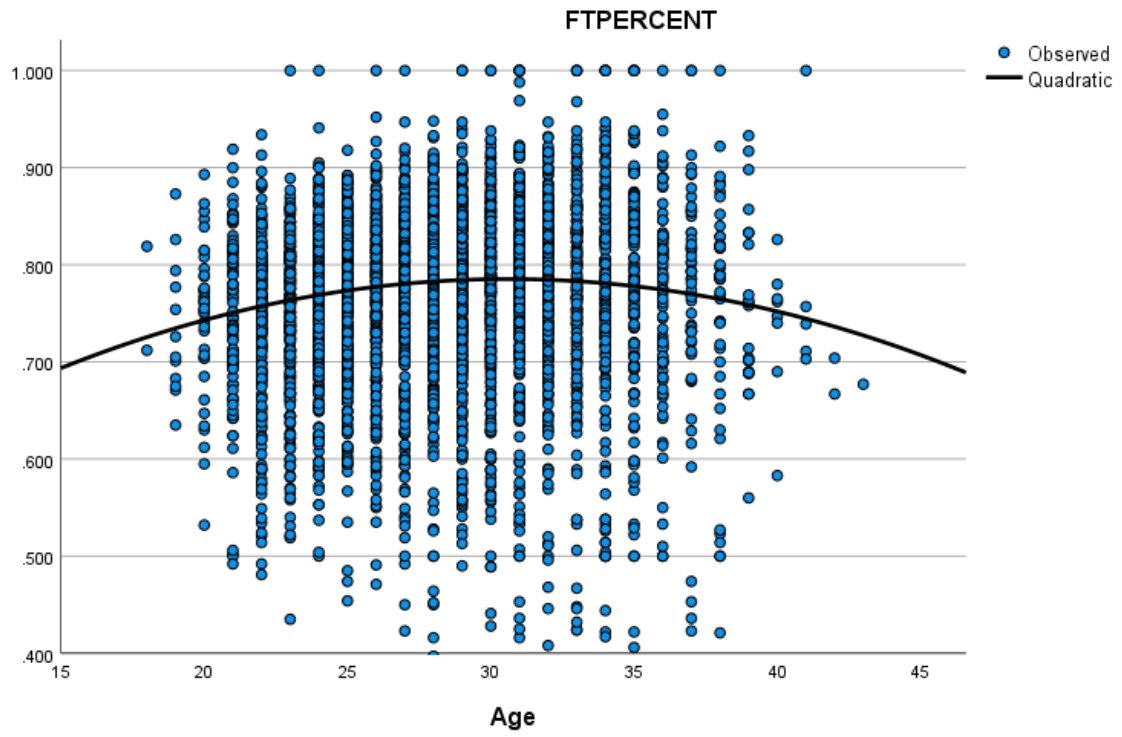
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	.010	.003	.870	3.877	<.001
Age ** 2	.000	.000	-1.032	-4.596	<.001
(Constant)	.352	.037		9.579	<.001



For the hypothesized curvilinear relationship between age and free-throw percentage, result showed statistically significant relationship (B = -.0004, p < .001).

Coefficients

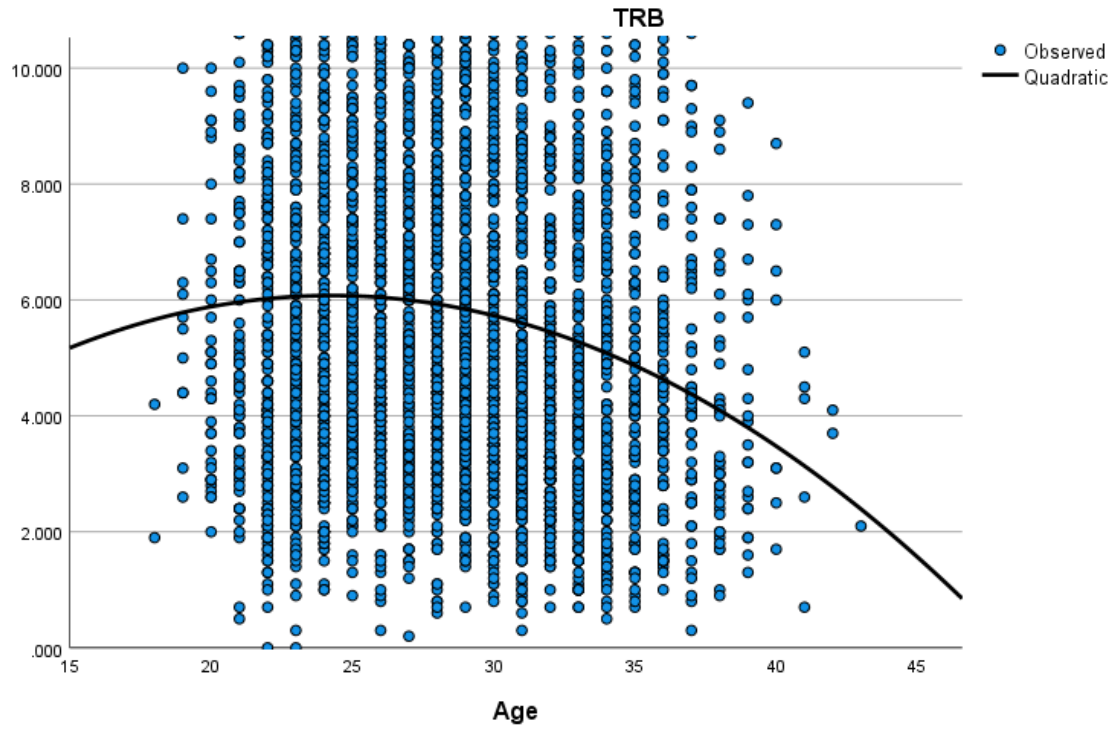
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	.023	.005	1.116	4.920	<.001
Age ** 2	.000	.000	-1.058	-4.663	<.001
(Constant)	.432	.067		6.471	<.001



The analysis also supported the hypothesis that the relationship between age and total rebounds is nonlinear as shown below:

Coefficients

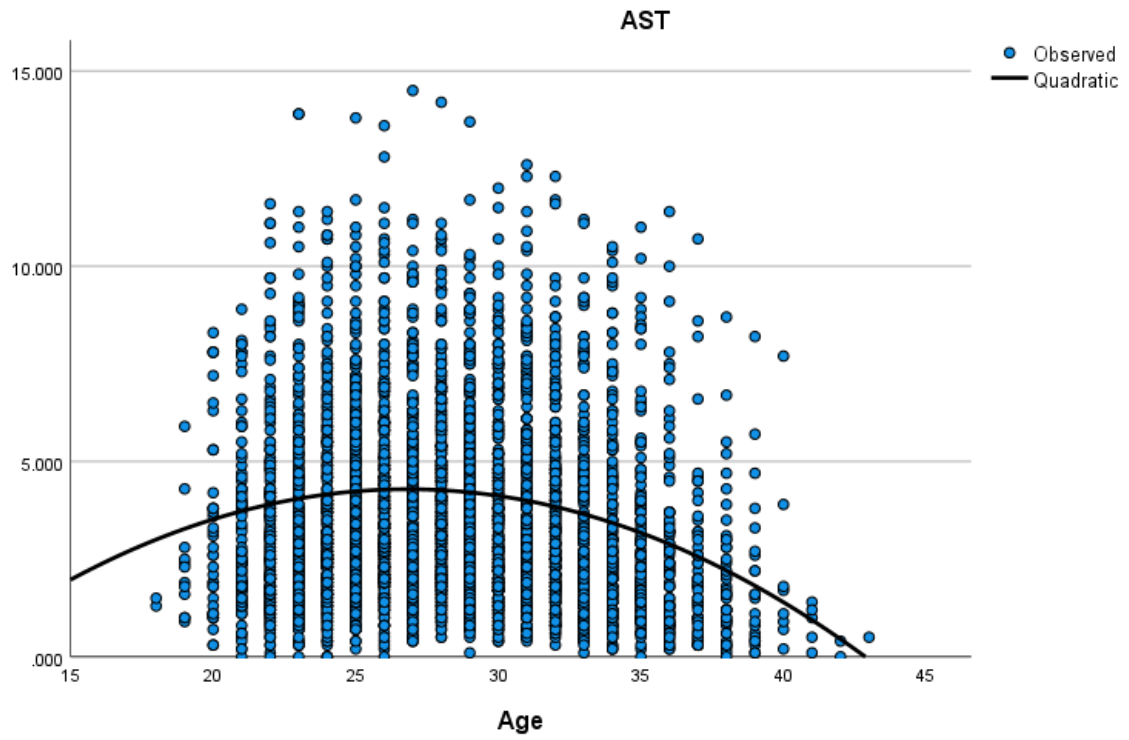
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	.510	.154	.746	3.314	<.001
Age ** 2	-.010	.003	-.892	-3.960	<.001
(Constant)	-.108	2.193		-.049	.961



The analysis also supported the hypothesis that the relationship between age and assists is curvilinear, such that as players got older up to a moderate age level, their assists increased, before decreasing as they got older.

Coefficients

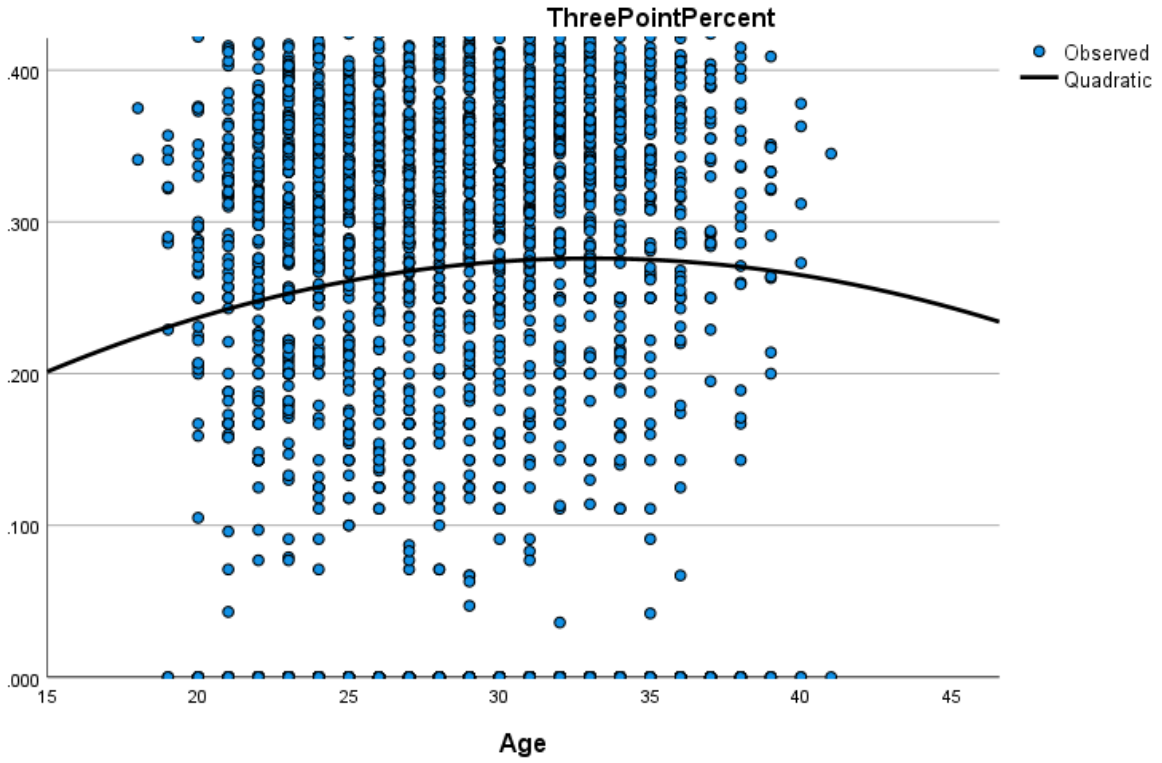
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	.891	.122	1.629	7.283	<.001
Age ** 2	-.017	.002	-1.762	-7.879	<.001
(Constant)	-7.655	1.745		-4.387	<.001



The analysis also supported the hypothesis that the relationship between age and three-point percentage is curvilinear (B = .0003, p = .041).

Coefficients

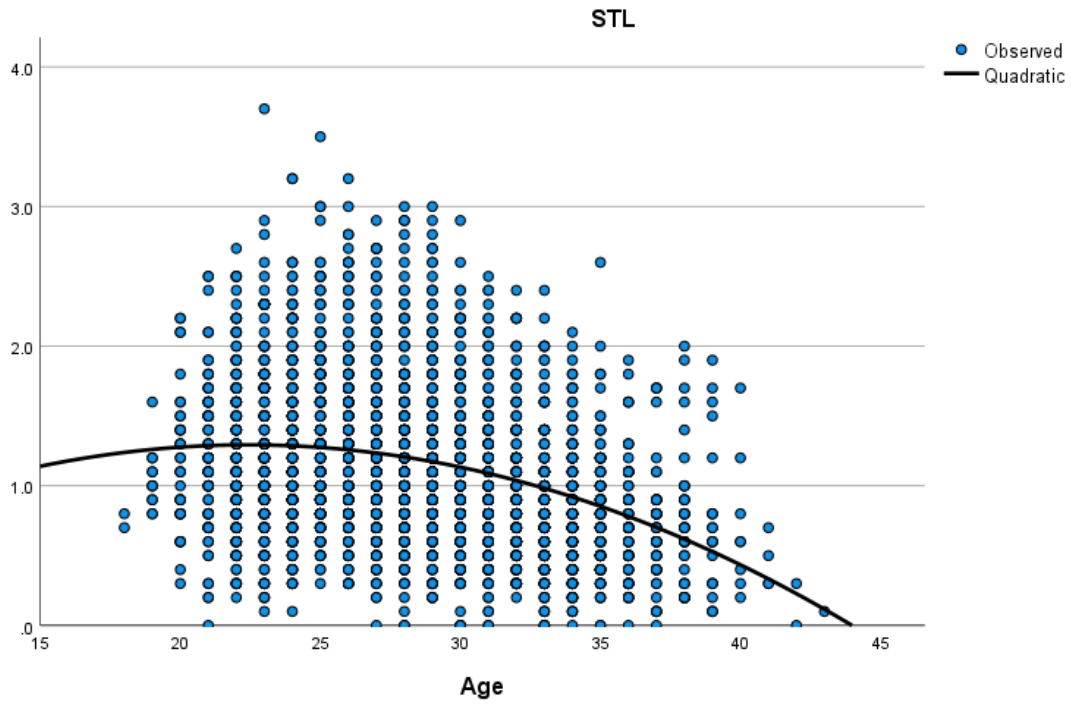
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	.015	.008	.442	1.940	.053
Age ** 2	.000	.000	-.387	-1.701	.089
(Constant)	.026	.111		.233	.816



When the model was used to predict steals as another aspect of performance, the hypothesized curvilinear relationship between age and steals was statistically significant, such that as players got older to a moderate level, their average steal increased initially, which started to decline as they got older.

Coefficients

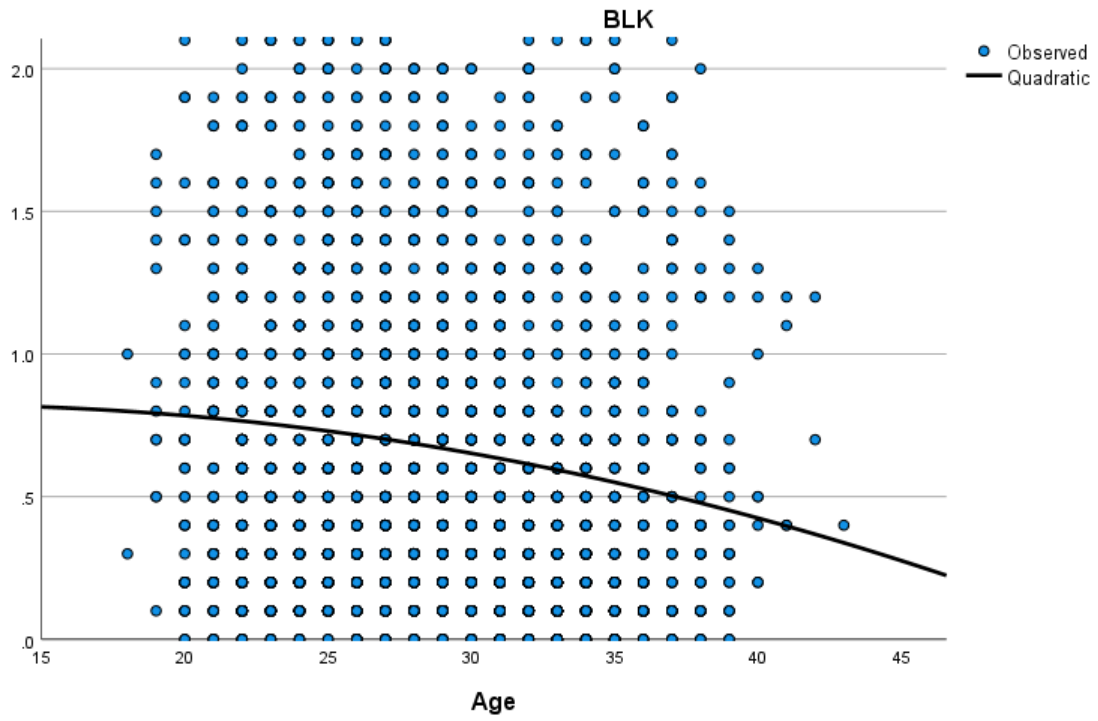
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	.125	.026	1.033	4.778	<.001
Age ** 2	-.003	.000	-1.336	-6.179	<.001
(Constant)	-.106	.372		-.284	.777



Another model was tested using age as a predictor of block shots. For block shots, the decline would appear much faster compared to other performance indicators.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
Age	.011	.035	.068	.299	.765
Age ** 2	.000	.001	-.177	-.781	.435
(Constant)	.764	.502		1.521	.128



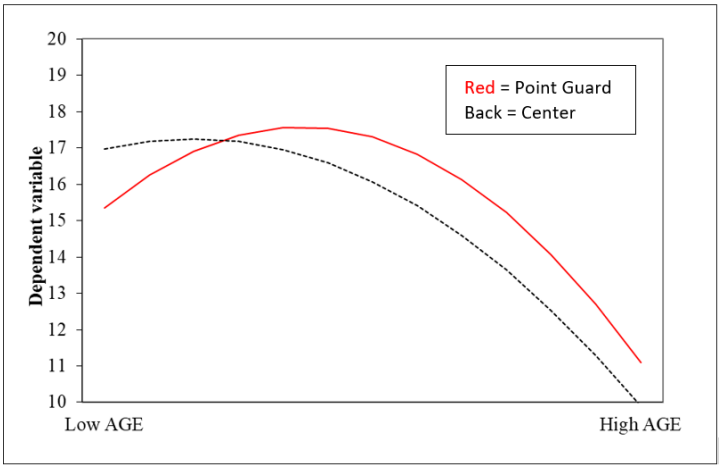
Finally, I also examined if the curvilinear relationship between age and performance outcomes by the NBA players is moderated by the positions they played.

The first model using PTS (average points made) as the dependent variable showed positions that NBA players played significantly moderated the curvilinear relationship between age and the dependent variable (see the table below):

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-31.186	3.864		-8.070	<.001
	Age	3.782	.271	2.391	13.973	<.001
	AGE_SQ	-.073	.005	-2.679	-15.659	<.001
2	(Constant)	-31.163	3.879		-8.034	<.001
	Age	3.781	.271	2.390	13.961	<.001
	AGE_SQ	-.073	.005	-2.679	-15.643	<.001
	POS_N	-.005	.074	-.001	-.070	.944
3	(Constant)	-33.479	3.976		-8.421	<.001
	Age	3.810	.271	2.409	14.067	<.001
	AGE_SQ	-.071	.005	-2.613	-15.104	<.001
	POS_N	1.229	.477	.252	2.577	.010
	AGE_POSN	-.043	.017	-.274	-2.619	.009
4	(Constant)	-62.657	9.232		-6.787	<.001
	Age	5.878	.650	3.716	9.048	<.001
	AGE_SQ	-.107	.011	-3.925	-9.509	<.001
	POS_N	10.522	2.697	2.156	3.901	<.001
	AGE_POSN	-.698	.188	-4.421	-3.718	<.001
	AGESQ_POSN	.011	.003	2.455	3.501	<.001

a. Dependent Variable: PTS

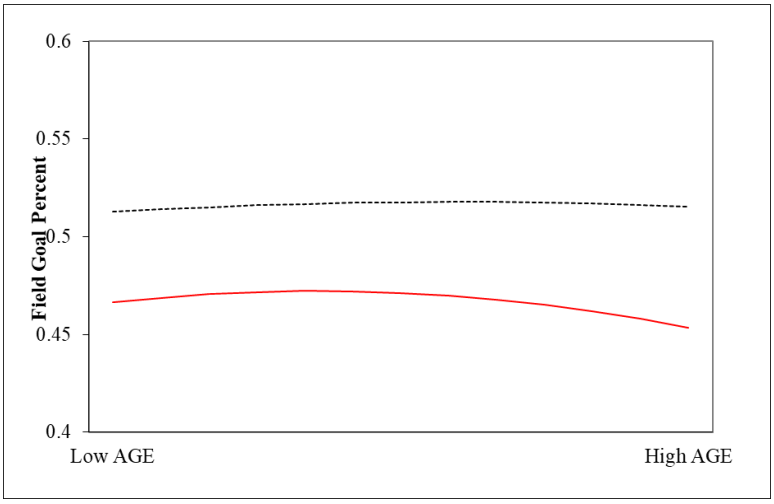


This graph shows that the peak happens at a relatively later age for point guards compared to players who play center positions whose peak performance is reached much quicker. The next analysis has shows that position also significantly moderates the curvilinear relationship between age and field goal percentage.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.409	.034		12.082	<.001
	Age	.005	.002	.375	2.048	.041
	AGE_SQ	.000	.000	-.454	-2.480	.013
2	(Constant)	.355	.032		10.969	<.001
	Age	.006	.002	.489	2.808	.005
	AGE_SQ	.000	.000	-.587	-3.366	<.001
	POS_N	.012	.001	.309	20.086	<.001
3	(Constant)	.375	.033		11.325	<.001
	Age	.006	.002	.470	2.697	.007
	AGE_SQ	.000	.000	-.659	-3.744	<.001
	POS_N	.001	.004	.034	.345	.730
	AGE_POSN	.000	.000	.298	2.792	.005
4	(Constant)	.213	.077		2.774	.006
	Age	.018	.005	1.354	3.236	.001
	AGE_SQ	.000	.000	-1.546	-3.678	<.001
	POS_N	.053	.023	1.322	2.348	.019
	AGE_POSN	-.003	.002	-2.508	-2.069	.039
	AGESQ_POSN	6.233E-5	.000	1.661	2.323	.020

a. Dependent Variable: FGPERCENT



For point guards in general, the field goal percentage (in red) was considerably lower compared to the field goal percentage for centers (in black).

Discussion and Conclusion

This report analyzed one of the most important questions that any NBA fan would be interested to hear. That is, when can we expect NBA players' performance to decrease? Is it as soon as they join the NBA? After 5 years? Or is it after 10 or more years? Answering this question can come in handy for any NBA insiders or NBA organizations who need to be smart about how they draft and keep their players.

Using a nonlinear curve estimation model, I found that the relationship between NBA players' age and performance is mostly nonlinear. More specifically, the curve estimation models showed that the average block shots, steals, points scored, and steals made started to decline after first few years in the NBA.

However, I noticed a big difference between offense-related performance and defense-related performance. For example, in terms of the offense-related performance indicators, NBA players' three point percentage did not decrease until later in their career, which was after the age of 35. The actual peak on thee point shooting percentage occurred on average between 30 and 35 years of age. The results for defense-related performance were, although nonlinear, more dramatic in terms of how early their performance started to decline. For example, for steals made, the peak was around age 25 after which it quickly declined. The pattern was the same for other defense-related performance indicators, such as blocked shots, which showed decline early in the players' career.

In conclusion, this report makes one clear message for NBA executive owners, which is that they must consider the players' arsenal when drafting or signing veteran players because their performance will depend on whether they are offensive-minded players or not. For offensive-minded players, NBA executives would continue to benefit from signing them as the return on investment is greater for longer period of time. However, for defensive-minded players, NBA executives, as the results showed, will be better off to be conservative when signing older players.