Journal of Physical Education and Sport ® (JPES), Vol.20 (2), Art 100 pp. 686 - 689, 2020
online ISSN: 2247 - 806X; p-ISSN: 2247 – 8051; ISSN - L = 2247 - 8051 © JPES

Original Article

Predicting per game performance through per minute performance in basketball

JOSE A. MARTÍNEZ
Business Management Department, Technical University of Cartagena, SPAIN

Published online: March 31, 2020
(Accepted for publication: February 21, 2020)
DOI:10.7752/jpes.2020.02100

Abstract:
This paper introduces a way to relativize per-game Player Total Contribution PTC/G by minutes played. It is a form of obtaining a prediction of PTC/G (PTCpred) in order to compare the performance of basketball players who have played different minutes along the season. A sample of 5060 NBA players was collected and analyzed using linear and non-linear variables. In addition, a Maxima code for computing standard error is provided. Results shows that PTCpred could be valuable for analysts, media and fans. Finally, several limitations are discussed in order to avoid mistakes in its practical use.
KeyWords: *basketball, NBA, player total contribution, statistics, box-score, player productivity*

Introduction

Usually the performance of basketball players is provided per-game. Not only points, rebounds, assists, or the remaining box-score variables are showed in rankings, but global performance measures as Win Score or Efficiency (among others) are also provided.

In a series of papers (Martínez, 2019a;b), the Player Total Contribution (PTC) index was presented as an improved method to obtain a global performance index based on box-score variables:

$$PTC = 1\,PTS + 0.91\,BLK + 0.58\,DRB + 0.92\,ORB + 0.86\,STL + 0.48\,AST + 0.23\,FD - 0.91\,MFG - 0.57\,MFT - 0.86\,TOV - 0.23\,PF$$

Where: PTS: points made; BLK: blocks made; DRB: defensive rebounds; ORB: offensive rebounds; STL: steals; AST: assists; FD: fouls drawn. MFG: missed field goals; MFT: missed free throws; TOV: turnovers; PF: personal fouls made.

PTC can be computed by game (PTC/G) or by minutes played (PTC/MP), just dividing PTC by games or minutes, respectively. Initially, it is advisable to provide the two indexes to obtain a complete picture of a player performance, but the problem arises when two players played different minutes. Consider this case showed in Table 1:

Table 1: Example of player performance for the NBA 2018/19 regular season

	PTC/G	PTC/MP	Per game minutes played
James Harden	29.773	0.810	37.0
Anthony Davis	29.642	0.897	33.0
Boban Marjanovic	8.787	0.748	11.7

James Harden obtained a slightly better PTC/G than Anthony Davis, but Anthony Davis was better in PTC/MP. This means that Davis played less minutes than Harden, and that minutes played influenced PTC/G. Therefore, Davis was better than Harden in per-minute performance, but as Harden played more minutes this could have affected the PTC/G variable.

The aim of this research is to answer the question of what would be the PTC/G predicted once known the PTC/MP, i.e. the aim of this study is to provide a way to compare PTC/G between players who have played different minutes along the season. Consequently, a unique index of global performance (PTCpred) is built.

Obtaining PTCpred is also important to project the per game performance of players who do not play many minutes. An example would be Boban Marjanovic (Table 1), who performed very well in the 11.7 average minutes played that season. Comparing the PTC/G of Marjanovic with Harden or Davis is unfair, but using PTCpred could be a way to obtain a better picture of that comparison.

Corresponding Author: JOSE A. MARTÍNEZ, E-mail: josean.martinez@upct.es

Material & methods

We computed the PTC/G and PTC/MP of the 2000 and the 2006-2019 seasons using the NBA official statistics. For the 2001-2005 seasons fouls drawn were not correctly computed in the web site of the NBA, so those seasons were initially discarded. In addition, we employed some filters to obtain a more realistic picture of players, discarding players who played less than 1/3 of games and less than 12.7% of minutes. For an 82 season game, this means discarding players who have played less than 28 games and less than 500 minutes. A total of 5060 observations were obtained as a sample.

We employed a quadratic approximation to predict PTC/G (1)

$$PTC/G_i = \beta_0 + \beta_1 PTC/MP_i + \beta_2 PTC/MP^2{}_i + u_i \quad (1)$$

Where PTC/G_i is the Player Total Contribution per game for each i player, β are the weights of the PTC/MP_i (being β_0 the intercept), and u_i is a random error. It is a model assumption that $cov(x, u) = 0$. We also proposed a linear approximation, i.e. fixing $\beta_2 = 0$

In order to compute de standard error of the predictions we employed a propagation of error approach, based on the Taylor series approximation to the variance function. The procedure is described in Ku (1966).

Results

We estimated the model (1), and a linear approximation. The Breusch-Pagan/Cook-Weisberg test for heteroscedasticity was significant $\chi^2 = 993.24; p < 0.001$ for the quadratic fit and significant $\chi^2 = 969.14; p < 0.001$ for the linear fit, so we computed robust standard errors using the Stata 13.0 software. Results are showed in Table 2.

Table 2: Model estimates

	Quadratic fit			Linear fit		
	Coef.	*Std. Err*	*p*-value	Coef.	*Std. Err*	*p*-value
PTC/MP	25.39	1.99	<0.001	37.80	0.423	<0.001
PTC/MP^2	14.90	2.62	<0.001			<0.001
β_0	-2.25	0.35	<0.001	-4.63	0.143	<0.001

The explained variance of both models was similar (0.700 vs 0.697),). As Figure 1 shows, the quadratic model fitted better in the region of high performance (high PTC/MP and high PTC/G), however the gains in explained variance were negligible and the errors of the parameter estimates were larger than the linear fit. Therefore, the linear fit was finally chosen as the model to predict PTC/G.

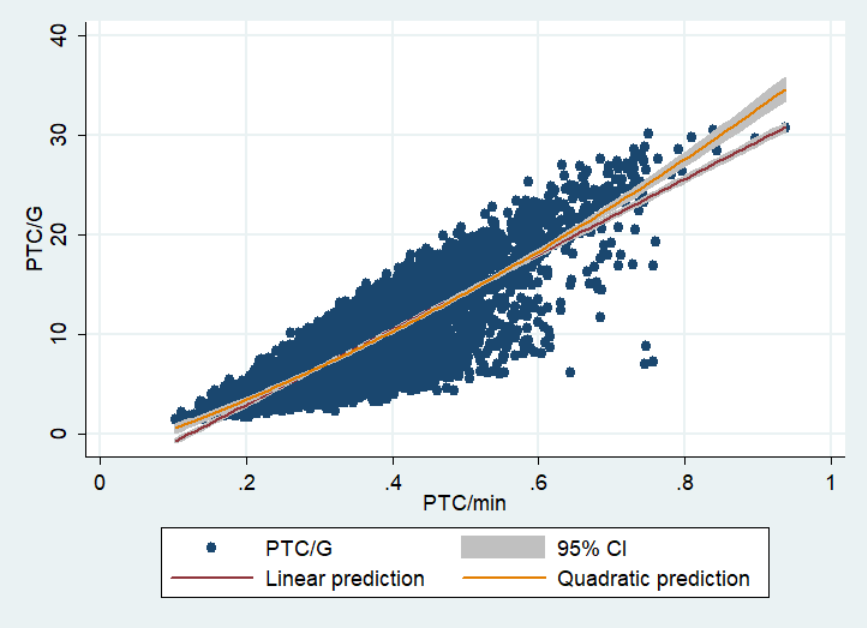

Figure 1: Quadratic and linear fit with 95% CI

Data of Table 2 can now be completed with PTCpred (Table 3). Now, PTCpred shows a different conclusion regarding the per-game performance of Harden vs Davis. Davis had higher per-game predicted production than Harden, once controlling by minutes played.
In addition, Marjanovic had a PTCpred close to the Harden performance, indicating that if the Serbian player had played the same minutes of Harden, his production could have been close to the production of Harden.

Table 2: Example of performance for the NBA 2018/19 regular season

	PTC/G	PTC/MP	Per game minutes played	PTCpred
James Harden	29.773	0.810	37.0	25.99
Anthony Davis	29.642	0.897	33.0	29.28
Boban Marjanovic	8.787	0.748	11.7	23.64

Computing the standard error
Martínez (2019b) wrote a script in Maxima to compute the standard errors of PTC/MP. Using Taylor series approximation as described in Ku (1966), the following code has been written (Table 3):

Table 3: Maxima code for computing the standard error of PTCpred

```
/*Program developed by Jose A. Martinez (Technical University of Cartagena, Spain) */
/*
Standard error computation using error propagation
First of all, employ the Maxima code proposed by Martinez (2019):
Martínez, J. A. (2019). A method to compute standard errors in per-minute performance metrics in basketball.
International Journal of Physical Education, Sports and Health,  6 (4). 45-51.
*/
/* Specify the equation*/
PTC_pred: beta_0+beta_1*PTCMP;
/* Compute the partial derivatives*/
partialder1:diff(PTC_pred,beta_0,1);
partialder2:diff(PTC_pred,beta_1,1);
partialder3:diff(PTC_pred,PTCMP,1);
/* Calculate the total error*/
totalerror:(((partialder1*standarderror_beta_0)^2)+
   ((partialder2*standarderror_beta_1)^2)+
   ((partialder3*standarderror_PTCMP)^2))^0.5;
/* Evaluate the equations in their corresponding estimated values*/
PTCpred: ev(PTC_pred,
beta_0= -4.63,
beta_1=37.80,
PTCMP=PTC_MP);
standarderrorPTCpred:ev(totalerror,
standarderror_beta_0=0.143,
standarderror_beta_1=0.423,
standarderror_PTCMP=standarderrorfinal,
PTCMP=PTC_MP,
beta_1=37.80);
/*Compute the confidence intervals*/
IC95lowpred:PTCpred-(Z*standarderrorPTCpred), numer;
IC95highpred:PTCpred+(Z*standarderrorPTCpred), numer;
/* End of program*/
```

For James Harden, and after applying the Maxima code shown in Table 3, the 95% confidence interval is (25.09 ; 26.89)

Limitations and further research

This research has some limitations that deserve to be commented. Firstly, the estimated coefficients should be recomputed each year, in order to incorporate new data and achieve a more robust estimation. Obviously, our sample size of 5060 observations was large, and new estimates with new data is expected to be very similar to the estimates provided in this study. In this sense, stability of estimates is expected.

Secondly, PTCpred slighltly underestimated the real PTC/G of James Harden. However, recall the PTC/G is also a variable with its own uncertainty when a player does not play all the games of the season. Therefore, both confidence intervals of the PTC/G and PTCpred should be compared in order to see if they overlap. And thirdly, prediction with players performing low minutes should be achieved with caution. The case of Marjanovic (see Table 1 and Table 2) was deliberately chosen because may illustrate a player profile

associated to certain physical features which are a barrier to play much more minutes. In addition, we have to stress that did not consider the individual profile of each player. This means that individual curves of performance could be different from one player to another. Even for the same player the relationship between PTC and minutes played could vary among seasons, depending of several factors such as age, skills developed, role in the team, etc. Figure 2 shows this relationship for the NBA career of James Harden. Globally, however, there is a positive relationship practically linear.

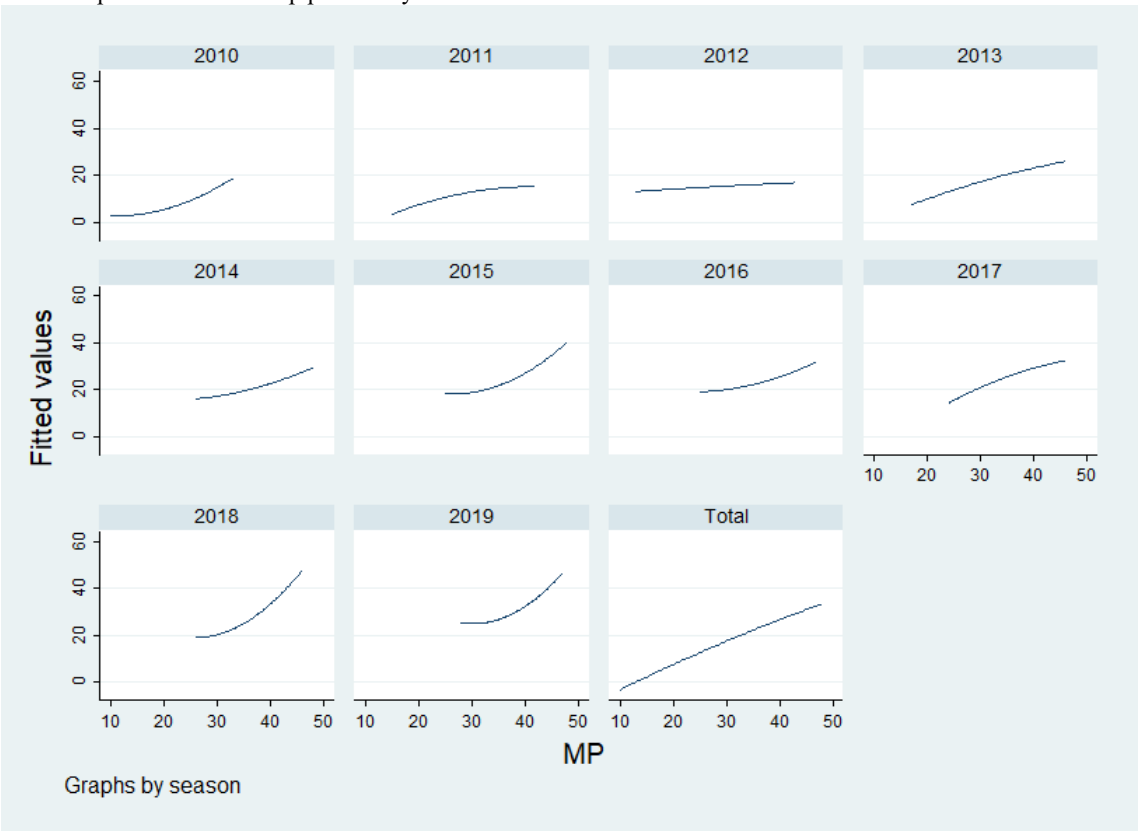

Figure 2: PTC/G predicted by minutes played of James Harden regular seasons

Conclusions

This paper has presented a way to relativize PTC/G by minutes played. It is a form of obtaining a prediction of PTC/G (PTCpred) in order to compare the performance of basketball players with disparate minutes played. Therefore, to obtain the value of PTCpred for every i player, we have to follow the next equation:

$$PTCpred_i = -4.63 + 37.80 PTC/MP_i$$

In spite of PTC/G and PTC/MP can be used to evaluate global performance, in some instances could be more practice to handle a unique index of production. In that case, PTCpred could be valuable for analysts, media and fans.

Funding: This study is the result of the activity carried out under the program Groups of Excellence of the region of Murcia, the FundaciónSéneca, Science and Technology Agency of the region of Murcia project 19884/GERM/15.
Conflicts of Interest: The author have no conflict of interest to declare.

References

Ku, H. H (1966). Notes on the use of propagation error formulas. *Journal of Research of the National Bureau of Standards-C. Engineering and Instrumentation,* 70 (4),263-273.

Martínez, J. A. (2019). A more robust estimation and extension of factors determining production (FDP) of basketball players. *International Journal of Physical Education, Sports and Health,* 6 (3). 81-85.

Martínez, J. A. (2019). A method to compute standard errors in per-minute performance metrics in basketball. *International Journal of Physical Education, Sports and Health,* 6 (4). 45-51.