Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v12n2p362

High performance of professional basketball players and the settings to measure their regularity: evidence from the Spanish ACB League

By Román Salmerón-Gómez, Samuel Gómez-Haro

Published: 14 October 2019

Electronic Journal of Applied Statistical Analysis
Vol. 12, Issue 02, October 2019, 362-379
DOI: 10.1285/i20705948v12n2p362

High performance of professional basketball players and the settings to measure their regularity: evidence from the Spanish ACB League

Román Salmerón-Gómez*[a] and Samuel Gómez-Haro[b]

[a] *University of Granada. Quantitative Methods Department, Economics and Business School. Campus Cartuja S/N. 18071. Granada (Spain).*
[b] *University of Granada. Business and Management Department II, Economics and Business School. Campus Cartuja S/N. 18071. Granada (Spain).*

Published: 14 October 2019

In this paper, we analyse the relationship between performance and regularity to avoid the problem to the use of averages in sports statistics given their sensitivity to extreme data. The aim of this work is to establish which statistics discriminate between the high and low performance and regularity of basketball players using a performance-regularity index. The sample is composed of players from the Spanish professional basketball league, ACB League, from 2014-2015. We divided the sample into players with low, medium and high performance and regularity through a k-means method and a discriminant analysis associating high numbers of rebounds, assists and steals with high performance and regularity and higher numbers of 3-point shots, turnovers and fouls with low performance and regularity. These results should help coaches and sports managers design strategies that improve performance and recruiting for their teams.

keywords: sport statistics, performance, regularity, basketball players, sport management.

*Corresponding author: romansg@ugr.es.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

1 Introduction

Professional basketball is one of the most followed sports in the world. In recent years, the application of statistics to this field has been increased by the desire of researchers and professionals to know the game more deeply and be able to make better technical and strategic decisions that affect final performance (Alamar (2013); Çene (2018); Drust (2010)).

A basketball game consists of a sequence of actions where players and teams adjust dynamically during periods of time. The actions and skills that develop during the game are varied, as well as the decisions that the players make based on their characteristics, their previous experience or other factors. Researchers and professionals have a large amount of information available on the actions of individuals and teams, and this analysis of the game has assumed a wide range of views and perspectives in recent literature. Starting from the individual contribution of players under different situations, game strategies, occupation of game spaces or probabilities of results, these analyses help us to identify individual or collective strengths and weaknesses to predict performance (Page et al. (2007); Metulini et al. (2018); Zuccolotto et al. (2017)).

In this paper, we work on the analysis of performance indicators available in the box-score of players in a season. Teams use these basic indicators as a measure of player performance, and they are useful for establishing patterns of play in the short term, such as choosing which players play at a particular time in a game, and in the long term, such as the negotiation of new salary or contract conditions (García et al. (2014); Ibáñez et al. (2008); Özmen (2016)).

Examples of player performance indicators include the TENDEX index of Heeren (1992), which is based on the difference between positive actions during the game (points scored, rebounds, fouls received, steals) and negative actions (missed shots, fouls made, turnovers); the Player Efficiency Rating (PER) of Hollinger (2004) or the Adjusted Plus-Minus (Rosenbaum (2004)). However, the way in which researchers and professionals use these data (and others indices) is still under discussion since these basic statistical measures ignore the context in which the player obtained them (Deshpande and Jensen (2016)), or depend too much on points scored (Berri et al. (2007)). The sensitivity of the average of these indices to extreme values has a consequence, it creates rankings where different players may have the same value according to a performance index even though the way they achieved these results differs substantially.

The aim of this article is to analyse the way that players obtain values under these performance indices through the concept of regularity, which refers to the frequency with which the player obtains these average values. Starting from research in the field (Martínez et al. (2017); Owen et al. (2007); Sampaio and Maças (2012); Schmidt and Berri (2001, 2002)) as a starting point, the authors develop the performance/regularity PR index (Salmerón-Gómez and Gómez-Haro (2016)) and establish three levels of performance and regularity: high, medium and low. Finally, the authors discriminate which statistical elements classify players in one of these levels, predicting player behaviour from their values in the statistical variables that classify them in one level or another.

The work is structured as follows: in section 2, we define the PR index as a measure of performance and regularity in professional basketball players and the statistical variable to which it is applied; in section 3, we analyse a set of players in terms of three levels of performance and regularity, and establish the factors that discriminate among these levels; in section 4, the authors discuss the implications that these results might have for the work of the coach and team manager; and finally, in section 5, we present the conclusions of the work.

2 Sports performance and regularity method

When professionals analyse the performance of a player during a season, regardless of the indices used, the tendency is to use the average value of the statistical variable used (points, rebounds, turnovers...) as a measure of central position that summarizes the observed values. However, the use of the mean as an indicator of measurement has certain limitations, such as its sensitivity to extreme values. This creates a situation in performance analysis in which different players may have the same values on a performance index even when the way they achieved these results differs substantially. This situation leads to the question of how to design a performance index that circumvents this issue.

Basing decisions exclusively on the performance or regularity of a player may not be appropriate if the right conditions are not present from the beginning. The PR performance-regularity index (Salmerón-Gómez and Gómez-Haro (2016)) attempts to avoid these recurring problems noted in previous literature (Deshpande and Jensen (2016)) by considering the regularity with which these performance analysis values are obtained.

2.1 The dataset

We did our analysis with data from players who participated in the Spanish ACB Basketball League during the 2014-2015 season. In the final data sample, we included 193 players from the 2014-2015 season who played at least 10 games for an average of at least 10 minutes in the ACB League. These criteria were imposed to provide the authors with data from players who participated consistently during the season to ensure the legitimacy of the study's results (Berri et al. (2007); Berri and Krautmann (2006); Cooper et al. (2009)). The data were obtained from the official database of the league (`www.acb.com`).

2.2 Performance-Regularity index

With the aim of obtaining a measure that takes into account performance (the average) and regularity (the way in which the average is obtained), the authors use the PR index (see Salmerón-Gómez and Gómez-Haro (2016)), which responds to the following expression:

$$PR = \overline{x} + \frac{\overline{x}}{S_x} = \overline{x} + CV_x^{-1},$$

where $CV_x = \frac{S_x}{\overline{x}}$ is the coefficient of variation, of which S_x is the standard deviation and $\overline{x}$ the average of any statistical variable that might be interesting to a coach/manager.

To illustrate the usefulness of this index, consider the following examples.

Example 1 *We suppose that the scores of two players in 80 matches is available[1]:*

P1: 22, 18, 23, 23, 18, 10, 25, 22, 13, 16, 19, 16, 17, 4, 19, 15, 18, 7, 22, 16, 28, 11, 15, 7, 29, 23, 22, 9, 31, 16, 21, 32, 10, 19, 13, 19, 6, 14, 19, 16, 2, 11, 18, 15, 11, 17, 16, 15, 16, 15, 11, 18, 6, 23, 9, 24, 8, 19, 10, 19, 12, 20, 11, 15, 22, 18, 8, 5, 10, 7, 10, 3, 18, 20, 3, 7, 3, 12, 21, 12.

P2: 8, 11, 12, 11, 13, 14, 13, 10, 13, 15, 11, 13, 9, 15, 11, 13, 15, 14, 11, 15, 16, 12, 13, 11, 16, 12, 11, 11, 9, 10, 13, 11, 13, 9, 12, 14, 14, 14, 12, 12, 11, 12, 7, 12, 11, 14, 10, 14, 11, 16, 10, 13, 13, 11, 12, 12, 15, 11, 10, 15, 15, 12, 13, 15, 16, 8, 11, 12, 16, 13, 12, 14, 11, 9, 12, 11, 14, 12, 11, 14.

It is observed that the first player has an average of 15.2875 and the second has an average of 12.2875. Considering this measure of central position, it seems clear that the first player is preferable to the second.

However, if we analyse how the players obtain these average values, we note that the first player has a standard deviation equal to 6.701962, while the second has a standard deviation equal to 2.038824. At the same time, the minimum and maximum values of the first player are 2 and 32, while those of the second are 7 and 16. Therefore, it seems clear that the second player is more regular than the first, thus some coaches may think that second player is preferable. This aspect is not reflected when using the mean as a measure that summarizes the information.

Finally, the PR values for the first and second player are 17.56855 and 18.31426, respectively. Therefore, once the measurement of performance and regularity have been combined into a single measure, the second player is preferable to the first. □

Example 2 *Game score (GmSc) is a metric developed by basketball statistician John Hollinger with the objective of calculating a player's statistical performance in a basketball game. This measure is obtained from the expression:*

> *Game Score = Points Scored + 0.4 · Field Goals — 0.7 · Field Goal Attempts — 0.4 · (Free Throw Attempts -– Free Throws) + 0.7 · Offensive Rebounds + 0.3 · Defensive Rebounds + Steals + 0.7 · Assists + 0.7 · Blocks — 0.4 · Fouls made -– Turnovers.*

If this measure is calculated for players in the ACB season 14-15 database, the two players with the highest averages are M. Todorovic (MT11, Bilbao) and A. Panko (AP7, Fuenlabrada). The GmSc values obtained by these players are:

MT11: 7.4, 22.7, 10.6, 4.3, 8.4, 14.2, 6, 18, 11.6, 11.6, 8.3, 9.9, 2.5, 7.3, 9.5, 7, 16.7, 19.7, 13.4, 18.3, 7.2, 14.8, 18, 7.3, 13.8, 12.9, 9.2, -0.4, 17.9, 12.2, 15.4, 11.6, 8.7, 19.3, 16, 14.3, 8.3.

[1] Simulated data from two normal distributions, P1 (player 1) with an average of 15 and a standard deviation of 7, and P2 (player 2) with an average of 12 and a standard deviation of 2.

Table 1: Analysis for the number of fast breaks made

Player	Team	Average (Ranking)	Stan. Dev.	PR (Ranking)
F. Causeaur	Baskonia	0.7027027 (1)	0.9087496	1.475966 (4)
B. Newley	Gran Canaria	0.6875 (2)	0.7803018	1.5685693 (1)
P. Pumprla	Obradoiro	0.666666 (3)	0.7772816	1.5243567 (2)
N. Richotti	Tenerife	0.666666 (4)	0.9574271	1.3629773 (5)
R. Neto	Murcia	0.61764706 (5)	0.6969503	1.503861 (3)

AP7: 6.8, 7, 11.5, 1.4, 12.7, 10.6, 1.9, -0.2, 7.6, 23.1, 9.2, 22.7, 16.9, 11.9, 6.2, 12.1, 9.2, 25.8, 13.5, 9.9, 23.4, 17.8, 10.1, 17.2, 3.2, 18.7, 16.3, 13.7, 17.8, 15.1, 6.4, 7.7, 24.2, -0.3.

It is observed that the average of MT11 is 11.72703, while that of AP7 is 12.09118. Therefore, the award for best player will go to AP7.

However, if we take into account the way in which the previous values have been obtained (the standard deviations are 5.189019 and 7.127342, respectively), the PR value of MT11 is 13.987, while that of AP7 is 13.78763. In other words, in this case MT11 would take first place. □

Example 3 *One of the variables available in the database compiled is the number of counterattacks made by each player. This variable may be of interest if you want to detect the most dangerous players in a fast game.*

The average values, standard deviations and calculations of PR are in Table 1. It can be observed that the first-place player considering the average value, F. Causeaur, would fall to fourth place when considering PR. If the average is used, we note a tie between P. Pumprla and N. Richotti that is clarified when considering the regularity of each player.

Finally, note that the most regular of the five players, R. Neto, experiences a two-position improvement when using PR. □

Furthermore, because the coefficient of variation is a non-dimensional measure, PR is not interpretable as a measurement unit. However, this index measures the performance of a player in any statistical variable based on the arithmetic mean; therefore, a higher mean indicates higher performance. Similarly, the index measures player regularity because it is based on the inverse of the coefficient of variation; therefore, a high value indicates high regularity. Therefore, it is clear that the higher the PR value, the better the player's performance and regularity (Salmerón-Gómez and Gómez-Haro (2016)).

2.3 Variables

As mentioned, the PR index can be applied to any statistical section considered to be of interest. In the present work we will use:

$$PerfPts = \frac{ACB\ Performance - Points}{Minutes} = \frac{Rest}{Minutes},$$

where:

Rest = Rebounds + Assists + Steals + Fouls received + Blocks — Missed throws – Turnovers -– Blocks received — Fouls made.

Thus, the dependence on the ACB's rating of points scored is eliminated; furthermore, dividing the results by minutes played removes the possible accumulation effect of playing more minutes. This measure is similar to the TENDEX index (Heeren, 1992), which is widely used to evaluate player performance.

With this result, the following interpretation is possible:

- If $PerfPts > 0$, then $Rest > 0$; that is, the number of positive actions (unrelated to the annotation) made by the player exceeds the number of negative actions.
- If $PerfPts < 0$, then $Rest < 0$; that is, the number of negative actions (unrelated to the annotation) made by the player exceeds the number of positive actions.

In short, if coaches or managers are interested in the scoring capacity of a player, then they only have to check the points scored in each match. On the other hand, if they want to rely on the overall performance of the player, they should use the player's overall rating (ACB Performance). However, as we noted before, this evaluation depends on the excess of the annotation. In our database, in those cases where $ACB\ Performance > Points$, we establish $Points/(ACB\ Performance)$, obtaining an average of 0.6446. That is, 64.64 % of the player's overall rating is due to his score. For this reason, if you are interested in measuring the overall performance of a player regardless of the points they have scored, in the sense of knowing whether their positive actions exceed their negative ones, then consider using $PerfPts$ instead of $ACB\ Performance$.

Example 4 *To illustrate the previous comments, in Tables 2 to 4 we show the top 10 results for $PerfPts$, $ACB\ Performance$ and $Points$ according to the average value. The value of PR is also shown in each case.*

We note that:

- *Between $PerfPts$ and $Points$ no players have repeat values, while between $ACB\ Performance$ and $Points$ there are 3 players in common (AP7, SJ89 and NM11), two of which occupy the first and third places as best scorers.*
- *The best rated player and top scorer, A. Panko (AP7), disappears from the first 10 positions of $PerfPts$. More specifically, he is in position 82. Thus, to use $ACB\ Performance$ as a measure of the best player would lead us to consider AP7 as the best of all when the reality is that he is really only the best scorer. This spurious conclusion would not be obtained when using $PerfPts$ to find the most complete player.*

□

Table 2: Top 10 $PerfPts$ ranking players according to the average value

Player	Team	Average	PR
L. Williams (LW21)	Bilbao	0.1633195227	0.919358721
L. Sikma (LS43)	Tenerife	0.1636652081	1.057956344
F. Vázquez (FV17)	Unicaja	0.1710020063	0.976591536
M. Begic (MB15)	Baskonia	0.1773810983	0.925443469
T. Satoransky (TS13)	Barcelona	0.2031295922	1.098117661
A. Tomic (AT44)	Barcelona	0.2153813572	1.164359733
W. Tavares (WT22)	Gran Canaria	0.2198994012	1.334549548
C. Suárez (CS12)	Unicaja	0.2241470273	1.482121470
M. Todorovic (MT11)	Bilbao	0.2295077007	1.648506223
A. Lima (AL23)	Murcia	0.2556038018	1.776634528

Table 3: Top 10 $ACB\ Performance$ ranking according to average value

Player	Team	Average	PR
D. Díez (DD33)	Gipuzkoa	14.5666667	16.0419251
F. Reyes (FR9)	R. Madrid	14.8292683	16.7218419
N. Martín (NM11)	Estudiantes	15.0294118	16.8901270
P. Ribas (PR5)	Valencia	15.0384615	16.9589327
A. Tomic (AT44)	Barcelona	15.1136364	17.0300420
L. Sikma (LS43)	Tenerife	15.7272727	18.1848227
S. Jelovac (SJ89)	Zaragoza	15.7941176	17.6969444
A. Lima (AL23)	Murcia	16.5882353	18.7361580
M. Todorovic (MT11)	Bilbao	18.0000000	20.3714795
A. Panko (AP7)	Fuenlabrada	18.8823529	20.9143661

Table 4: Top 10 *Points* ranking according to average value

Player	Team	Average	PR
A. Waczynski (AW21)	Obradoiro	12.470588	14.322455
N. Richotti (NR5)	Tenerife	12.484848	14.693775
K. Kuric (KK24)	Gran Canaria	12.805556	15.137267
M. James (MJ3)	Baskonia	12.925926	15.396248
N. Martín (NM11)	Estudiantes	12.970588	15.159259
V. Stojanocski (VS19)	Andorra	13.769231	15.864339
S. Burtt (SB15)	Fuenlabrada	13.882353	15.672912
S. Jelovac (SJ89)	Zaragoza	14.029412	16.534554
A. Mumbrú (AM15)	Bilbao	14.352941	17.261202
A. Panko (AP7)	Fuenlabrada	18.617647	21.505001

2.4 Sensitivity to extreme data

Because the arithmetic mean and standard deviation are sensitive to extreme data, PR has the same problem. It is clear that these extreme data are based on the concepts that regularity aims to quantify; therefore, it is necessary to distinguish between high and low data and abnormally high and low data. Undoubtedly, the latter would distort the analysis. Within descriptive statistics, there are techniques that detect the existence of such data; however, following the premise that the rates proposed should be easy to calculate and understanding that an analysis of such samples is not easy, this sample is analysed after ruling out 5% of the larger and smaller data.

However, to analyse the stability of the calculations in this section, we compared the results after discarding the highest and lowest 5% of the data and obtained the results by considering all of the data for the 193 players studied. For this comparison, we calculated the mean absolute error (MAE) and the mean square error (MSE) using the following formulas:

$$MAE = \frac{1}{193} \cdot \sum_{i=1}^{193} |PR_i - PR_i^T|, \quad MSE = \frac{1}{193} \cdot \sum_{i=1}^{193} (PR_i - PR_i^T)^2,$$

where PR_i denotes the PR value for the i-th player considering all data, and PR_i^T denotes the PR value for i-th player when the highest and lowest 5% of data are discarded.

In this case, we observed that the $MAE = 0.166165$ and the $MSE = 0.056099$. We did not observe large differences between the two methods of calculating PR. This result likely occurred because dividing the result by the number of minutes played softened any abnormally extreme values.

2.5 Statistical analysis

Discriminant analysis identifies the characteristics that differentiate two or more groups of individuals. Thus, membership in one group or another is used as the dependent variable. In this case, the dependent variable was the PR division (high, low or medium) determined by a cluster of k-means, while the independent variables are those that are assumed to differ among these groups. Because these are continuous quantitative variables, the use of this technique is not recommended for discrete variables (Pérez (2005)), but it is useful in this research context (Ibáñez et al. (2008); Sampaio et al. (2006)). In this work, we considered the following independent variables: number of games played (Games), minutes played (MinGame), 2-point shots scored per minute (2P), 2-point shots attempted per minute (2PA), 3-point shots scored per minute (3P), 3-point shots attempted per minute (3PA), 1-point shots scored per minute (1P), 1-point shots attempted per minute (1PA), defensive rebounds per minute (DRB), offensive rebounds per minute (ORB), assists per minute (AST), steals per minute (STL), turnovers per minute (TOV), blocks per minute (BLK), blocks received per minute (BLKR), fouls per minute (PF), fouls received per minute (PFR), and the plus/minus value per minute (plus/minus).

In addition, the previous analysis was completed taking into account the player's position. Specifically, we conducted an analysis of variance (ANOVA) of a factor with the null hypothesis that the mean PR value will be the same for players regardless of their position (1-point guard, 2-shooting guard, 3-forward, 4-power forward and 5-center). The sample was classified into these five categories based on information provided by the ACB League website, `www.acb.com`.

3 Results

3.1 Discriminant analysis and k-means method

First, with the objective of categorizing PR in classes that are identified with low, medium and high values, after the highest and lowest 5% of the sample was excluded, we performed a k-means analysis. This analysis classified the players into three groups so that the players that make up the same group are as similar as possible to each other and different from those of the other groups. The centres (average values of PR) and members of each cluster are presented in Table 5. Cluster 1 identifies high performance/regularity players (representing 34.19% of the sample), cluster 2 contains medium values of performance/regularity (representing 47.66% of the sample) and cluster 3 consists of low values of performance/regularity (representing 18.15% of the sample).

To obtain information about the individual significance of each variable in the discriminant function, we used stepwise inclusion. In this case, the eigenvalues of the two discriminant functions that form the model are very uneven; the first explains 99% of the available data, and the second explains only the remaining 1%. Moreover, the Wilks' lambda of the second function has an associated p-value equal to 0.458. That is, the second function is not significant enough to have an associated p-value greater than 0.05;

Table 5: Centres and numbers of members of each cluster

Cluster	Centres	Number of members
1	0.844	66
2	-0.279	92
3	-1.282	35

Table 6: Centroids function of the groups

PR levels	Function 1	Function 2
1	1.838	0.113
2	-0.357	-0.162
3	-2.613	0.214

therefore, only the information provided by the first function will be used (if it is shown to discriminate by having a Wilks' lambda less than 10^{-3} and, therefore, a p-value less than 0.05).

Table 6 shows the centroids in each of the discriminant functions. The first function distinguishes the players with low performance and regularity (whose centroid is on the negative side) from the players with high performance and regularity (whose centroid is on the positive side). Players with medium performance and regularity are in the central area. For this reason, this feature provides information only for the extreme PR index levels.

Finally, Table 7 shows the standardized coefficient matrix for each function. The variables not included in this table were eliminated by the analysis because they did not support discrimination among groups. The first column shows that the negative centroid is associated with players with low performance and regularity, and those players are characterized by a greater number of 3-point shots attempted (3PA), turnovers (TOV) and personal fouls made (PF). The positive centroid is associated with players with high performance and regularity, and those players are characterized by high values for defensive rebounds (DRB), assists (AST) and steals (STL).

3.2 Variance analysis

Given the statistics that discriminate between players with low and high performance and regularity, it could be assumed that $PerfPts$ favours players with a certain profile and harms others; for example, it might be assumed that inside players (centers) would have a benefit over outside players (point guards, shooting guards and forwards). Table 8 shows the mean and standard deviation of PR according to position.

Table 7: Standardized coefficients of discriminant functions

	Function 1	Function 2
3PA	-0.865	0.561
DRB	0.879	0.243
AST	0.962	-0.551
STL	0.448	0.242
TOV	-0.456	0.463
PF	-0.56	-0.742

Table 8: Mean and standard deviation of PR based on players' role

Position	Players (N)	Mean	Standard Deviation
Point-guard	42	0.0306	0.7833
Shooting guard	38	-0.6386	0.8315
Forward	34	-0.2255	0.7631
Power forward	46	0.0015	0.78704
Center	33	0.5345	0.7189

A test of homogeneity of variances did not reject the hypothesis that the variance is the same for all five subgroups (which was necessary to address equality of means requirement) because its associated p-value (0.995) is greater than 0.05. However, it rejected the null hypothesis of equal means because the p-value associated with the ANOVA is less than 10^{-3} and, therefore, less than 0.05. This result shows that the PR value differs depending on the player's position. Because there are significant differences among positions, it is appropriate to determine which positions have different performance and regularity. To perform this analysis, we used the Bonferroni inequality method, which showed significant (the null hypothesis of equal means is rejected for having associated p-value less than 0.05) differences between: point guards vs. shooting guards; shooting guards vs. power forwards and centers; forwards vs. centers; and power forwards vs. centers.

While the differences between forwards and centers and between shooting guards and power forwards and centers may seem obvious, the other two differing groups deserve special attention and are discussed in the next section.

In addition, the averages shown in Table 8 indicate that the PR values are highest for centers (players who particularly stand out on the rebound) and lower for shooting guards (players who base their game on the outside shot).

4 Discussion of the results

The objective of this study was to analyse what factors better discriminate between the actions of players during a season by using a new indicator that combines performance and regularity, PR. The results show that players with more 3-point shots, turnovers and fouls committed are associated with low performance and regularity in $PerfPts$. This profile is clearly associated with outside players, although significant differences between the guard positions are also observed. This is likely because of the increased number of assists (especially) and steals by point guards. Recall that these factors are associated with high performance and regularity.

The significant differences observed between power forwards and centers are explained by the greater prominence of power forwards in open shots (3-point shots, for example) in modern basketball. Because this statistical factor discriminates negatively, it explains why power forwards have the lowest average PR index values.

It has also been shown that centers have a higher level of performance and regularity. This is because, after the points scored variable was removed, these players have the easiest time accumulating shares of the remaining statistical features. This might suggest that the use of this index to rank the performance and regularity of a player is biased. However, in Table 8, the average PR values by position are provided. For example, a player with an average PR of zero could be a guard with high (above average) regularity and performance, a power forward with medium regularity and performance or center with low (below average) regularity and performance. That is, the results obtained allow a player to be classified as having high or low performance and regularity regardless of his position.

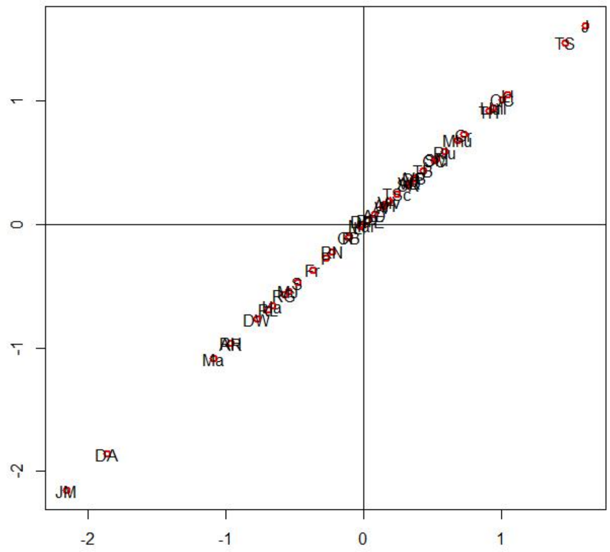

Figure 1: Point-guards PR index

On the other hand, sports professionals, the media and general supporters may believe that players with high performance and regularity play more games and have more playing time per game because of their allegedly higher quality. However, these factors do not discriminate among quality levels. In the case of games played, this may be because, in modern basketball, the number of minutes played is greatly divided, and it is common for the whole roster to play in most games.

Therefore, this paper encourages collaboration between scholars and practitioners since its results could allow people who make decisions for sports teams – namely, coaches and managers - to have the best possible information to make correct decisions. From a theoretical point of view, this paper uses a new indicator that provides support for new research in sports management. That is, it shows that it may be more appropriate to use PR than the average to summarize the information regarding a certain statistical section if you want to take into account the player's regularity in addition to his performance. From a practical point of view, this work establishes thresholds for classifying the players' performance by taking into account their position and shows aspects of play that lead to high/low values of the indicator studied. This information can be used by coaches, managers or agents to make rosters. In this sense, in Figure 5, it can be seen that A. Lima is the center with best PR value (2.1417, the average value for centers is 0.5345) which (probably) marked his signing for Real Madrid (last champion league) next season. We show, in Figures 1 to 5, the PR values for all players analysed according to their role on the court.

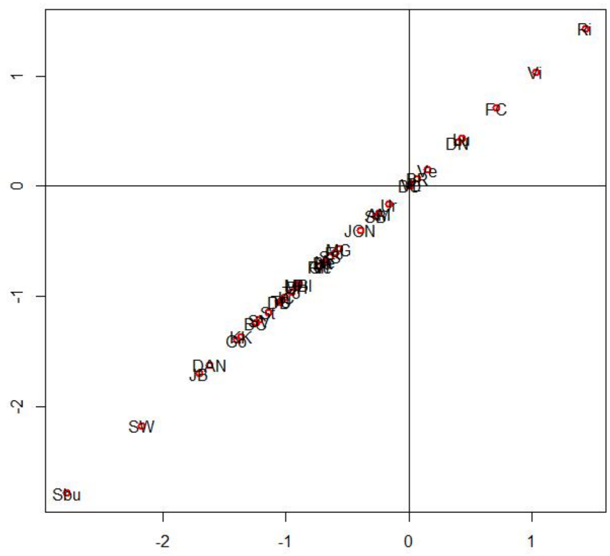

Figure 2: Shooting guards PR index

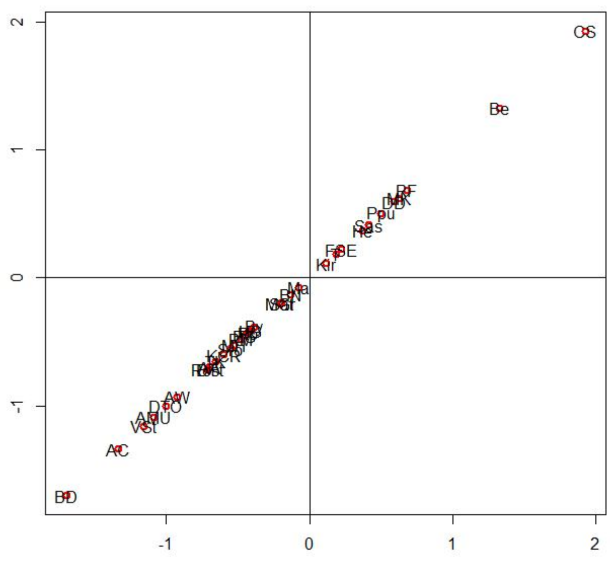

Figure 3: Forwards PR index

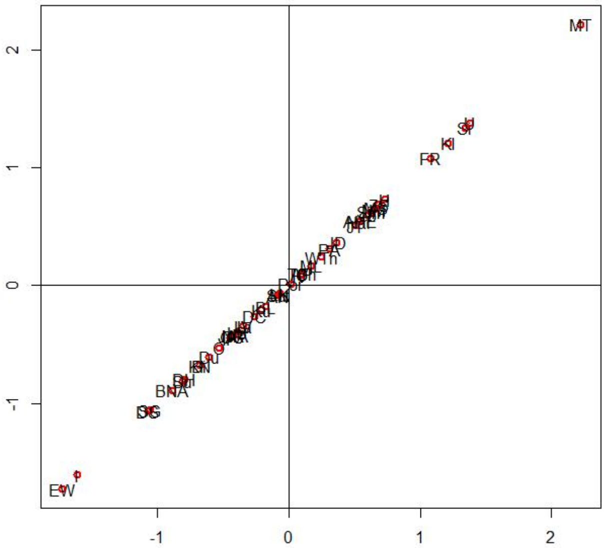

Figure 4: Power-forwards PR index

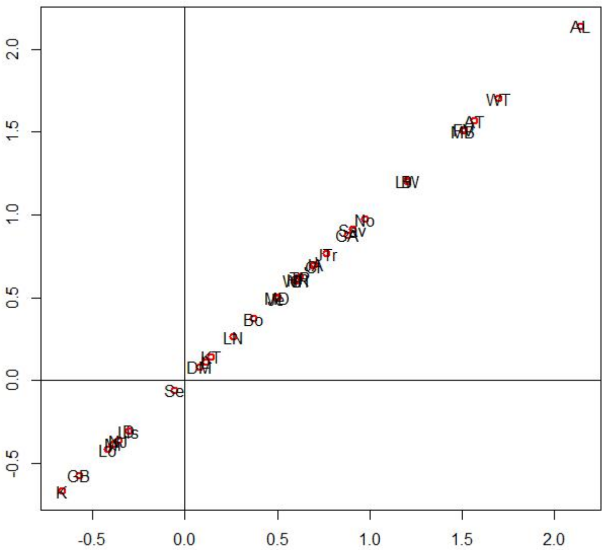

Figure 5: Centers PR index

5 Conclusions

The identification and development of new performance indicators for players and teams has great value for the decisions of sports professionals, coaches and managers (Alamar (2013); Özmen (2016); Çene (2018)). The objective of this article is to determine the factors that differentiate the performances of basketball players after introducing the concept of regularity. As we have explained, the analysis of dynamic performance in team sports such as basketball is highly complex and requires the ongoing review of existing indicators in an attempt to obtain relevant information that can improve decision-making among coaches and managers. When a coach or manager of a basketball club wants to determine whether a given player has the profile of a scorer, he or she may simply consider the player's points per game. However, this process can be very time consuming when one wants to determine the player's quality without taking scoring into account. To that end, we propose subtracting the total points scored from those scored by the player and dividing the result by the player's minutes played. This method avoids the accumulation effect that results when a greater number of minutes are played (Salmerón-Gómez and Gómez-Haro (2016)).

Throughout this process, the average value of the index is used as a measure of central tendency that summarizes the observed values. However, this measure has limitations, including sensitivity to extreme values. This creates a situation in which players can obtain the same performance values even when the ways in which these results are produced differ. To address this issue, we use an index that combines performance and regularity.

Based on the performance and regularity of ACB League players in the 2014-2015 season, this paper analysed the statistics that distinguish players with high performance and regularity from those with low performance and regularity. The identification of these factors could be used by coaches and managers to classify new players who have not been analysed in terms of high or low performance and regularity. The utility of these results for sports professionals is obvious. Because the PR index is a new and simple indicator that takes into account not only players' performance but also their regularity, it provides useful information to help coaches and team managers develop strategies for selecting or renewing players in subsequent seasons.

Finally, there are a number of limitations that may need to be addressed in future research, such as determining whether playing at home or away affects a player's performance and regularity, determining the influence of the second competition (for players who compete during the season's European competitions, such as the Euroleague) or the time of the competition (first or second round, regular season or playoff, etc.). Future research should also expand the sample of players to search for greater regularity in the mean values of PR according to the position of the player. This extension could be made by considering data from other ACB seasons and/or other leagues and observing whether the values differ materially.

Acknowledgement

This work was partially supported by the Spanish Ministry of Education [project ECO2013-47009-P]; and the Regional Government of Andalusia [project P11-SEJ-7988]. Authors want to thank the members of the research group, "Innovation, Sustainability and Business Development" (SEJ-481), and Quantitative Methods Department for their help and comments.

References

Alamar, B. (2013). *Sports analytics: A guide for coaches, managers, and other decision makers.* Columbia University Press.

Berri, D. J., Brook, S. L., and Schmidt, M. B. (2007). Does one simply need to score to score? *International Journal of Sport Finance*, 2:142–148.

Berri, D. J. and Krautmann, A. C. (2006). Shrinking on the court: testing for the incentives effects of guaranteed pay. *Economic Inquiry*, 44(3):536–546.

Çene, E. (2018). What is the difference between a winning and a losing team: insights from euroleague basketball. *International Journal of Performance Analysis in Sports*, 18(1):55–68.

Cooper, W., Ruiz, J., and Sirvent, I. (2009). Selecting non-zero weights to evaluate effectiveness of basketball players with dea. *European Journal of Opera-tional Research*, 195:563–574.

Deshpande, S. and Jensen, S. (2016). Estimating an nba player's impact on his team's chances of winning. *Journal of Quantitative Analysis in Sports*, 12(2):51–72.

Drust, B. (2010). Performance analysis research: Meeting the challenge. *Journal of Sports Sciences*, 28(9):921—-922.

García, J., Ibáñez, S., Gómez, M., and Sampaio, J. (2014). Basketball game-related statistics discriminating acb league teams according to game location, game outcome and final score differences. *International Journal of Analysis in Sport*, 14:443–452.

Heeren, D. (1992). *Basketball Abstract.* Prentice Hall.

Hollinger, J. (2004). *Pro Basketball Forecast.* Washington, DC: Brassey's Sports.

Ibáñez, S., Sampaio, J., Feu, S., Lorenzo, A., Gómez, M., and Ortega, E. (2008). Basketball game-related statistics that discriminate between team's season-long success. *European Journal of Sport Science*, 8(6):369–372.

Martínez, J., Ruiz, M., Casals, M., and López, F. (2017). Regular point scoring by professional basketball players. *Electronic Journal of Applied Statistical Analysis*, 10(3):759–772.

Metulini, R., Manisera, M., and Zuccolotto, P. (2018). Modelling the dynamic pattern of surface area in basketball and its effects on team performance. *Journal of Quantitative Analysis in Sports*, 14(3):117–130.

Owen, P., Ryan, M., and Weatherston, C. (2007). Measuring competitive balance in professional team sports using the herfindahl-hirschman index. *Review of Industrial Organization*, 31:289–302.

Özmen, M. (2016). Marginal contribution of game statistics to probability of winning at different levels of competition in basketball: Evidence from the euroleague. *International Journal of Sport Science and Coaching*, 11:98–107.

Page, G., Fellingham, G., and Reese, C. (2007). Using box scores to determine a position's contribution to winning basketball games. *Journal of Quantitative Analysis in Sports*, 3(4):1–18.

Pérez, C. (2005). *Métodos Estadísticos Avanzados con SPSS.* Thomson.

Rosenbaum, D. (2004). *Measuring How NBA Players Help Their Teams Win.* `http://www.82games.com/comm30.htm`.

Salmerón-Gómez, R. and Gómez-Haro, S. (2016). Expanding horizons on performance measurement and regularity in professional basketball. *RICYDE, Revista internacional de ciencias del deporte*, 45(12):234–249.

Sampaio, J., Janeira, M., Ibáñez, S., and Lorenzo, A. (2006). Discriminant analysis of game-related statistics between basketball guards, forwards and centers in three professional leagues. *European Journal of Sport Science*, 6(3):173–178.

Sampaio, J. and Maças, V. (2012). Measuring tactical behaviour in football. *International Journal of Sports Medicine*, 33:395–401.

Schmidt, M. and Berri, D. (2001). Competition and attendance: The case of major league baseball. *Journal of Sports Economics*, 2:147–167.

Schmidt, M. and Berri, D. (2002). Competitive balance and market size in major league baseball: A response to baseball's blue ribbon panel. *Review of Industrial Organization*, 21:41–54.

Zuccolotto, P., Manisera, M., and Sandri, M. (2017). Big data analytics for modeling scoring probability in basketball: The effect of shooting under high-pressure conditions. *International Journal of Sports Science and Coaching*, 14(3):117–130.