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/i20705948v10n2p666

On the credibility of basketball scoring efficiency

By Pulgarín, Arias-Nicolás, Jiménez

Published: 15 November 2017

Electronic Journal of Applied Statistical Analysis
Vol. 10, Issue 03, November 2017, 666-676
DOI: 10.1285/i20705948v10n2p666

On the credibility of basketball scoring efficiency

Antonio A. Pulgarín*[a], José P. Arias-Nicolás[a], and Héctor V. Jiménez[b]

[a] *University of Extremadura, Department of Mathematics, 10003 Cáceres (Spain)*
[b] *University of Extremadura, Department of Financial Economics and Accounting, 10003 Cáceres (Spain)*

Published: 15 November 2017

Our aim deals with appraising the scoring efficiency of a player in terms of points scored per hundred possessions. A Bayesian approach to the problem, should reflect not only individual scoring skills, but also taking into account the collective performance. In this wide context, credibility theory becomes an adequate mechanism deciding whether scoring efficiency calculation to be more or less plausible. We model the scoring per possession process by means of the conjugated family Multinomial-Dirichlet in order to obtain a net scoring efficiency credibility formula.

keywords: Credibility factor, Multinomial-Dirichlet, scoring efficiency.

1 Introduction

An adequate framework estimating productivity is used by the sport industrial organizations for management applications (Mertz et al., 2016) in order to offer general managers or head coaches the possibility of measuring players for the decision making in business or team improvement.

Certain performance indicators of the game has been advocated to provide a measure of the offensive productivity of a basketball player on the court, being advanced statistics more accurate tools to evaluate in depth the offensive skills either of a player or team, in contrast with the classical box score statistics that could be very misleading.

*Corresponding author: aapulgar@unex.es

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

Comparing statistics is more effective at the level of opportunities to score points in the game, -namely possessions- instead of entire games or minutes played, since teams play at significantly different paces. The most commonly-held view is that a possession is all the time a team holds the ball before the other team gains it, roughly speaking, an uninterrupted control of the ball. There are formulas that can do a pretty good job of estimating the total number of possessions in a game from standard box score statistics.

There are only three ways for a possession to end: 1) a field goal attempt (FGA) that is not rebounded by the offense (either a make or a defensive board of a miss), 2) a turnover (Tov), or 3) some free throws (FTA).

Free throws ending possessions are not as well-recorded. Some free throws come in pairs, and the first of the pair cannot end the possession. Oliver (2004) found through several score sheets, that in the range between 40-50% accounts the fact that when a player scores a basket and is fouled, they shoot a free throw, which is not a possession. This is also true of flagrant fouls and technical fouls, while three free throws make up one possession when a player is fouled shooting a 3-pointer. The consensus value for college basketball is due to Pomeroy (2004) by assuming 47.5%. As to the NBA, Hollinger (2002) typically uses 44%. Depending on the chosen prefactor the formula may slightly differ, although the practical difference is probably meaningless (Zimmermann, 2016).

Here is one place where defining possession becomes tricky, because there are two different definitions used by various analysts. The defining characteristic of this concept is whether assuming or not that the control continues after an offensive rebound. Team's plays do not necessarily end with a shot since an offensive rebound extends a possession. However, it is not the case whether computing a possession used by a player. The main difference from individual and team possession is then, that an offensive rebound starts a new possession instead of continuing the previous one. NBA.com Stats (2017: Accessed 03-03) already keeps statistics by a player using this definition.

Calculating the number of possessions by a single player are merely estimate box score stats, defined as time before an attempt to score is made or a turnover is recorded (ESPN.com Insider, 2017: Accessed 03-03; NBA.com Stats, 2017: Accessed 03-03, -glossary available online at `http://stats.nba.com/help/glossary`):

$$\mathrm{Poss} = \mathrm{FGA} + 0.44 \times \mathrm{FGA} + \mathrm{Tov} \tag{1}$$

Our next task will deal with the indicator (Usg%) representing an estimate of the percentage of team possessions used by a player while he was on the floor.

Remark. *Extracting (*Usg%*) from* `NBA.com` *Stats has the advantage that records are obtained from exact data instead of estimates (counting all offensive team possessions while a player is on the court).*

By setting (Poss) as in (1), usage percentage will allow us to set:

$$\mathrm{Poss}^{\mathrm{Tm}} = 100 \times (\mathrm{Poss} \,/\, \mathrm{Usg\%}) \tag{2}$$

In the sequel, superindex (Tm) will refer to stats by the team "solely" when the player is on the floor.

1.1 Offensive rating

Evaluating points scored per 100 possessions, is known as "offensive rating" or "offensive efficiency". We shall refer to player offensive efficiency as

$$\text{OEff} = 100 \times (\text{Pts} \,/\, \text{Poss}) \tag{3}$$

The concept of offensive rating can have two very different meanings depending on who is providing the number. On the one hand, NBA.com's offensive rating is a reflection of how many points the team scored per 100 possessions when the player was on the court, namely:

$$\text{OffRtg} = 100 \times (\text{Pts}^{\text{Tm}} \,/\, \text{Poss}^{\text{Tm}}) \tag{4}$$

By combining (2) and (4) we may calculate:

$$\text{Pts}^{\text{Tm}} = \text{OffRtg} \times (\text{Poss} \,/\, \text{Usg\%}) \tag{5}$$

Offensive rating (ORtg) published in Basketball Reference (2017: Accessed 03-03) is a complex formula designed by Dean Oliver (further details can be found in Kubatko et al. (2007)) that estimates individual points produced, -through made shots, assists, and offensive rebounds-, per 100 player total possessions (the total number of how many times he ends his team's possession).

We do not disagree at all with the fact that (OffRtg) should reflect part of the impact of the scoring efficiency of a player in the game (in fact, we emphasize that this mechanism is a convenient indicator), but we strongly support the idea that (OEff) should have a weight in the calculation, and it would be desirable measuring "how credible" is this indicator.

1.2 Credibility theory

Credibility theory was originally developed in actuarial sciences to determine risk premiums, as a convex linear combination of the the individual experience and a prior belief on the collective. Bailey (1945) showed that credibility formulas may be derived from Bayes theorem, and further Bayesian techniques were introduced in a big way in the late 1960s when Bühlmann (1967) laid down the foundation to the empirical Bayes credibility approach, which is still being used extensively.

We claim for a simply computed credibility formula of scoring efficiency for a player combining both the individual information and the collective belief.

By compounding both concepts as an unified approach, we are betting for the factor (OEff) when the player has an outstanding participation in the game. Otherwise, if the contribution of the player is residual, some "belief" on the collective offensive efficiency (OEff^{Tm}) should be a more accurate indicator for the player. Several authors suggest to compare (OEff) always with (Usg%).

Definition 1. *A credibility formula of scoring efficiency for a player is a convex linear combination*

$$\text{SEff} = C \times \text{OEff} + (1 - C) \times \text{OEff}^{\text{Tm}} \tag{6}$$

*being C an increasing function on (*Usg%*) called credibility factor, bounded from below by 0 and from above by 1.*

2 Method

Our object of study is the categorical variable X determining the number of points ($s = 0, 1, 2, 3$) scored by a player in a given possession. We are supposing that distribution parameters are $\mathbf{p} = (p_0, p_1, p_2, p_3)$, with $p_0 + p_1 + p_2 + p_3 = 1$, where $p_s = P[X = s]$ (at prior unknown) determines the propensity of the team (it may depend of certain strategies conditioned while the player is on the court) to score s-points in an offensive possession when the player is on the floor.

We may identify X with the multivariate random variable

$$\mathbf{X} = (X_0, X_1, X_2, X_3) \sim \text{Multinomial}(\mathbf{p}),$$

where X_s denotes the number of possessions in which the player scores s-points. Note that X is a random variable specifying each individual outcome, while the multinomial distribution $\mathbf{X}$ specifies the number of outcomes of each of the categories (adequacy of multinomial model to determine point scores within a possession is justified in Parker (2010)).

Let $n = \text{Poss}$ be the total number of individual possessions used by a player during a game, and consider the sample $S = \{s_1, \ldots, s_n\}$ consisting of points scored by the payer in each one of the n possessions, then $\overline{S} = \text{OEff}/100$. The likelihood function based on the sample $S = \{s_1, \ldots, s_n\}$ is given by,

$$l(S|\mathbf{p}) = \prod_{k=1}^{n} p_{s_k} = p_0^{n_0} p_1^{n_1} p_2^{n_2} p_3^{n_3}, \tag{7}$$

where n_s indicates the number of possessions in which the player scores s-points, then $n_0 + n_1 + n_2 + n_3 = n$ and we denote $\mathbf{n} = (n_0, n_1, n_2, n_3)$. Recall that the probability mass function of a multinomial distribution is proportional to (7):

$$P[\mathbf{X} = \mathbf{n}] = \begin{cases} \dfrac{n!\,(p_0^{n_0} p_1^{n_1} p_2^{n_2} p_3^{n_3})}{n_0! n_1! n_2! n_3!} & , \text{ when } n_0 + n_1 + n_2 + n_3 = n \\ 0 & , \text{ otherwise} \end{cases}$$

Bayes approach relies on the knowledge of prior distributions characterizing such probabilities. The main empirical support for the Dirichlet distribution is that it provides a better fit to the aggregate scoring frequency distribution than that given by the assumption that all individuals have the same distribution (Multinomial-Dirichlet assumption can be read from Lee et al. (1968)).

Dirichlet distribution to be prior for parameter $\mathbf{p}$ means that

$$\mathbf{p} \sim \text{Dirichlet}(\mathbf{a}), \text{ i.e. } \pi(\mathbf{p}) = \frac{\Gamma(a)\left(p_0^{a_0-1} p_1^{a_1-1} p_2^{a_2-1} p_3^{a_3-1}\right)}{\Gamma(a_0)\Gamma(a_1)\Gamma(a_2)\Gamma(a_3)}$$

and $E_{\mathbf{p}}[p_s] = a_s/a$, where $\mathbf{a} = (a_0, a_1, a_2, a_3)$ are the hyperparameters of the prior distribution (this term is used to distinguish the parameters of the model for the underlying system under analysis) and $a = a_0 + a_1 + a_2 + a_3$.

Bayes theorem, allows us to obtain a posterior distribution of probability by means of the likelihood function and the prior distribution (see Lamas et al. (2015)):

$$\pi(\mathbf{p}|S) = \frac{l(S|\mathbf{p})\pi(\mathbf{p})}{E[l(S|\mathbf{p})]} = \frac{\Gamma(a+n)\left(p_0^{n_0+a_0-1} p_1^{n_1+a_1-1} p_2^{n_2+a_2-1} p_3^{n_3+a_3-1}\right)}{\Gamma(n_0+a_0)\Gamma(n_1+a_1)\Gamma(n_2+a_2)\Gamma(n_3+a_3)}, \text{ then}$$

$$\mathbf{p} \sim \text{Dirichlet}(\mathbf{n}+\mathbf{a}), \text{ and therefore } E_{\mathbf{p}|S}[p_s] = (n_s + a_s)/(n+a).$$

Now we define principles for the calculation of the player scoring efficiency. Recall that a loss function is a mapping $L : \mathbb{R}^2 \to \mathbb{R}$ which attributes to each par (s, x) the error assumed in a possession for a player "expecting" to scored x meeting with a score s.

Definition 2. *A principle for the calculation of the player scoring efficiency from a loss function* $L{:}\mathbb{R}^2 \to \mathbb{R}$, *is the functional* $x(\mathbf{p})$ *minimizing the expected loss* $E_X[L(s, x(\mathbf{p}))]$.

If we consider the standard squared-error loss function $L(s, x) = (s - x)^2$, then by deriving $E_X[L(s, x(\mathbf{p}))]$ over x we have $dE_X/dx[(s - x)^2] = -2E_X[s] + 2x$ (the second derivation is $2 > 0$, thus a minimum), and therefore $x(\mathbf{p}) = E_X[s] = p1 + 2p2 + 3p3$ which is known as the net efficiency principle (see Heilmann (1989)).

3 Results

Definition 3. *Player scoring efficiency for the observed sample* S *is the value (*SEff*) minimizing the expected loss function* $E_{\mathbf{p}|S}[L(x(\mathbf{p}), \text{SEff}/100)]$.

Following analogous procedures of derivation as to the above net efficiency principle, we have that

$$\begin{aligned}
\text{SEff}/100 =& E_{\mathbf{p}|S}[x(\mathbf{p})] \\
=& \frac{(n_1+a_1) + 2(n_2+a_2) + 3(n_3+a_3)}{n+a} \\
=& \left(\frac{n}{n+a}\right)\left(\frac{n_1+2n_2+3n_3}{n}\right) + \left(\frac{a}{n+a}\right)\left(\frac{a_1+2a_2+3a_3}{a}\right) \\
=& \left(\frac{\text{Poss}}{\text{Poss}+a}\right)\left(\frac{\text{Pts}}{\text{Poss}}\right) + \left(1 - \frac{\text{Poss}}{\text{Poss}+a}\right)\left(\frac{a_1+2a_2+3a_3}{a}\right)
\end{aligned}$$

by establishing a credibility formula

$$\text{SEff} = C \times \text{OEff} + (1 - C) \times \text{OEff}^{\text{Tm}} \tag{8}$$

with credibility factor

$$C = \text{Poss}/(\text{Poss}+a) \tag{9}$$

and collective belief

$$\mathrm{OEff}^{\mathrm{Tm}} = 100 \times (a_1 + 2a_2 + 3a_3)/a \tag{10}$$

3.1 Hyperparameters

In accordance with the general estimation principle of "the larger the sample the better", a can be understood as the "precision", in the sense that when a_s is large enough with respect to a, then a_s is likely to be near p_s, otherwise when a_s is small with respect to a, then p_s is distributed more diffusely.

Taking into account that C should be an increasing function on Usg% $= 100 \times (\mathrm{Poss}\,/\,\mathrm{Poss}^{\mathrm{Tm}})$, from (9) next situations hold

$$\begin{array}{lllll} C = 1 & \iff & (\mathrm{Poss}^{\mathrm{Tm}} - \mathrm{Poss}) = 0 \text{ and } \mathrm{Poss} > 0 & \iff & a = 0 \text{ and } \mathrm{Poss} > 0 \\ C = 0 & \iff & (\mathrm{Poss}^{\mathrm{Tm}} - \mathrm{Poss}) > 0 \text{ and } \mathrm{Poss} = 0 & \iff & a > 0 \text{ and } \mathrm{Poss} = 0 \end{array}$$

Thus, a naive approach may assert

$$a_s = \lambda(m_s - n_s), \text{ with } 0 < \lambda < 1 \tag{11}$$

m_s denoting the number of offensive possessions scoring the team s-points while the player is on the floor. Therefore, we may obtain from (10) a collective belief

$$\mathrm{OEff}^{\mathrm{Tm}} = 100 \times (\mathrm{Pts}^{\mathrm{Tm}} - \mathrm{Pts})/(\mathrm{Poss}^{\mathrm{Tm}} - \mathrm{Poss}) \tag{12}$$

not depending on λ.

From the Bayesian point of view, using data establishing the prior might seem inappropriate, however several authors support this approach as a useful approximation to the preferred method of hierarchical modeling, since intuitive interpretation of hyperparameters provides situations avoiding complex statistical calculus.

Accordingly, if $\lambda \to 0$, then from (8) SEff = OEff. Otherwise $\lambda \to 1$, implies SEff = OffRtg. Hence, λ can be understood as a factor determining whether (SEff) to be more or less concentrated around (OEff) or (OffRtg). To this aim, a non-informative Uniform$(0, 1)$ distribution, whose mean is $\overline{\lambda} = 1/2$, can be used to estimate the value of λ.

By virtue of (9), our credibility factor yields

$$C = 2 \times \mathrm{Poss}\,/(\mathrm{Poss} + \mathrm{Poss}^{\mathrm{Tm}}) \tag{13}$$

Finally, by compounding (8), (12) and (13) we are obtaining our desired scoring efficiency credibility formula:

$$\mathrm{SEff} = 100 \times (\mathrm{Pts} + \mathrm{Pts}^{\mathrm{Tm}})/(\mathrm{Poss} + \mathrm{Poss}^{\mathrm{Tm}}) \tag{14}$$

4 Discussion

This paper explores an individual factor of productivity appraising the scoring efficiency of a player. There are several works analysing offensive efficiency (Chen et al., 2013; Skinner, 2010) or comparing individual and team abilities to construct an advanced statistical framework for data collection to decision making (Bruce, 2016). In this sense, our proposed model can be used to evaluate player effective opportunities to score points during the game.

NBA player with the highest OffRtg = 153.1 during the Regular Season 2015/16 was Coty Clarke (BOS), however this value is not significant since he participated only in 3 games (with 2 MP in average).

We have refined our study just for relevant players. During the Regular Season 2015/16, the number of games disputed by a player in average was of 55, the number of minutes played per game was of 20, and the credibility factor was of 0.3 (this value is very important since it distinguishes active players during the game, -in some sense similar to the usage percentage). We have reduced our study only for players whose stats are above the average of these indicators.

Table 1: Top25 (SEff) players during the NBA Regular Season 2015/16

Rnk	Player	Tm	Poss	Poss^{Tm}	Pts	Pts^{Tm}	C	OffRtg	SEff	Diff
1	Stephen Curry	GSW	25.7	80.3	30.1	93.7	0.48	116.7	116.8	0.1
2	Klay Thompson	GSW	20.2	76.5	22.1	88.3	0.42	115.4	114.1	-1.3
3	Draymond Green	GSW	15.1	81.2	14	94.5	0.31	116.4	112.8	-3.6
4	J.J. Redick	LAC	13.9	62.1	16.3	69.2	0.37	111.5	112.6	1.1
5	Kevin Durant	OKC	25.7	84.3	28.2	95.6	0.47	113.4	112.5	-0.9
6	Kawhi Leonard	SAS	18.6	72.1	21.2	78.7	0.41	109.2	110.2	1
7	LeBron James	CLE	24.8	79.7	25.3	89.6	0.47	112.4	110	-2.4
8	Chris Paul	LAC	19.6	72.9	19.5	81.4	0.42	111.7	109.2	-2.5
9	Enes Kanter	OKC	11.6	49.6	12.7	54.1	0.38	109.1	109.2	0.1
10	Jonas Valanciunas	TOR	11.9	57.5	12.8	62.7	0.34	109.1	108.8	-0.3
11	Kevin Love	CLE	16.3	69.4	16	77	0.38	111	108.6	-2.4
12	LaMarcus Aldridge	SAS	17.2	66.2	18	72.2	0.41	109.1	108.2	-0.9
13	Russell Westbrook	OKC	25.6	81.8	23.5	92.4	0.48	113	107.8	-5.2
14	Dirk Nowitzki	DAL	17.5	68.6	18.3	73.9	0.41	107.7	107	-0.7
15	Patrick Patterson	TOR	7.2	37.3	6.9	40.7	0.32	109	106.9	-2.1
16	Kyle Lowry	TOR	21.3	81.9	21.2	88.9	0.41	108.6	106.8	-1.8
17	Damian Lillard	POR	25.6	82.6	25.1	89.3	0.47	108.1	105.8	-2.3
18	Terrence Ross	TOR	9.6	52.5	9.9	55.8	0.31	106.3	105.8	-0.5
19	James Harden	HOU	28.8	88.6	29	95	0.49	107.2	105.6	-1.6
20	Jae Crowder	BOS	13.7	74.5	14.2	78.9	0.31	105.9	105.6	-0.3
21	Avery Bradley	BOS	15.5	77.5	15.2	82.9	0.33	107	105.5	-1.5
22	Karl-Anthony Towns	MIN	17.7	71.7	18.3	76	0.4	106	105.5	-0.5
23	DeMar DeRozan	TOR	23.6	79.5	23.5	85.1	0.46	107.1	105.4	-1.7
24	Hassan Whiteside	MIA	13.2	64.4	14.2	67.4	0.34	104.7	105.2	0.5
25	Kemba Walker	CHO	20.9	78.9	20.9	83.9	0.42	106.4	105	-1.4

In the top of the ranking is located the 2015/16 MVP Stephen Curry (GSW) with the SEff = 116.8 practically the same that his OffRtg.

Recall that by "Big-Three" we are regarding the three all-star players by the leaders of the Eastern Conference: Cleveland Cavaliers (CLE) and the Western Conference: Golden State Warriors (GSW) respectively, namely: Irving-James-Love and Curry-Green-Thompson. The outstanding (GSW) Regular Season with the 73-9 record, coincides with the fact that the (SEff) podium is justly composed by the (GSW) Big-Three. The stats numbers shows to Draymond Green (GSW) as an excellent basketball player in a successful team. However, with a OffRtg = 116.4, -similar to Curry-, our credibility formula punishes him -3.6 points. This is the case of a low-credibility factor $C = 0.31$ indicating that he uses few offensive possessions. Although Green doesn't score too much (14 Pts per game in average), he keeps a remarkable 3rd position (only one position lost) since (GSW) scores at the level of SEff = 112.8 when he is on the court.

As to the (CLE) Big-Three, Kyrie Irving does not appear in Table 1 (although we have included him at the scatted diagram of Figure 1) since he played only 53 games during the Regular Season because a long-term injury from the previous season. With a high-credibility factor $C = 0.45$, our formula reduces -3.3 points his OffRtg = 110.4 by passing to SEff = 107.1.

In the proposed credibility model there are some overrated players and players deserving a higher offensive rating. Accordingly, Russell Westbrook (OKC) is the player whose SEff = 107.8 (13th position) presents the biggest variation (-5.2) with respect to his OffRtg = 113 (5th position). Since he has a high-credibility factor $C = 0.48$, this overrating identifies a player not producing too many points in accordance with the big amount of possessions that he consumes (lower offensive efficiency than rating).

The most underrated players in accordance with our credibility formula are J.J. Redick (LAC) by passing from OffRtg = 111.5 (8th position) to SEff = 112.6 (4th position), and Kawhi Leonard (SAS), OffRtg = 109.2 (10th position) to SEff = 110.2 (6th position). Their production of points per possession used are excellent (Redick with a credibility factor of $C = 0.37$ and Leonard with $C = 0.41$), and their offensive rating would deserve to be increased (higher offensive efficiency than rating).

A player deserving special mention is Hassan Whiteside (MIA) by climbing 25 positions in the ranking (from 49th to 24th) with a SEff = 105.2.

Now, by reducing Playoffs study again just for players whose stats are above the league average (MP = 20, $C = 0.3$ and GP = 9), we find once more that Kawhi Leonard (SAS), OffRtg = 107.8 (10th position) and SEff = 108.3 (4th position) is one of the most underrated players (Jonas Valanciunas (TOR) has a variation of 1.3 which is even bigger than Leonard's, although his SEff is not higher), while Russell Westbrook (OKC), OffRtg = 112.2 (4th position) and SEff = 106.3 (10th position) is the most overrated one.

By chance, (CLE) Big-Three is located in the first three positions of the Playoffs' ranking and they are the 2015/16 NBA Champions. Stephen Curry (GSW) falls down up to the 9th position by passing from SEff = 116.8 in Regular Season to SEff = 106.5 in Playoffs. By comparing the Big-Threes evolution from Regular Season to Playoffs in the scatted diagrams from Figure 1, it is shown the deflation by the (GSW) against

Table 2: Top25 (SEff) players during the NBA Playoffs 2016

Rnk	Player	Tm	Poss	Poss^{Tm}	Pts	Pts^{Tm}	C	OffRtg	SEff	Diff
1	Kevin Love	CLE	15.1	68	14.7	80.1	0.36	117.8	114.2	-3.6
2	Kyrie Irving	CLE	24.2	80.7	25.2	93.4	0.46	115.7	113.1	-2.6
3	LeBron James	CLE	26.1	85.3	26.3	97.6	0.47	114.4	111.2	-3.2
4	Kawhi Leonard	SAS	20.4	73.4	22.5	79.1	0.43	107.8	108.3	0.5
5	Draymond Green	GSW	16.5	87.3	15.4	96.8	0.32	110.9	108.1	-2.8
6	Klay Thompson	GSW	22.7	81.4	24.3	87.8	0.44	107.9	107.7	-0.2
7	Shaun Livingston	GSW	8.9	49.4	8.2	54.4	0.31	110.1	107.3	-2.8
8	LaMarcus Aldridge	SAS	20	72.5	21.9	77.1	0.43	106.3	107	0.7
9	Stephen Curry	GSW	25	79.6	25.1	86.3	0.48	108.4	106.5	-1.9
10	Russell Westbrook	OKC	29.5	87.5	26	98.2	0.5	112.2	106.3	-5.9
11	Kevin Durant	OKC	29.8	93.4	28.4	102.5	0.48	109.7	106.3	-3.4
12	CJ McCollum	POR	21.5	93.5	20.5	99.7	0.37	106.6	104.6	-2
13	Al-Farouq Aminu	POR	14.6	79.3	14.6	82.8	0.31	104.4	103.7	-0.7
14	Damian Lillard	POR	28.1	94.9	26.5	99	0.46	104.3	102	-2.3
15	Tony Parker	SAS	12.2	56	10.4	59.1	0.36	105.6	101.9	-3.7
16	Dwyane Wade	MIA	22.8	71	21.4	74	0.49	104.2	101.6	-2.6
17	Joe Johnson	MIA	13.3	72.7	12.1	75.3	0.31	103.6	101.6	-2
18	Kyle Lowry	TOR	22	82.7	19.1	85.4	0.42	103.3	99.8	-3.5
19	Jonas Valanciunas	TOR	13	60.5	13.8	59.5	0.35	98.4	99.7	1.3
20	Maurice Harkless	POR	12	57.7	11	58.2	0.34	100.9	99.4	-1.5
21	Jeff Teague	ATL	16.9	63.5	14.5	65	0.42	102.4	98.9	-3.5
22	Kent Bazemore	ATL	14.1	73.8	11.9	74.4	0.32	100.8	98.2	-2.6
23	Goran Dragic	MIA	18.3	71.2	16.5	71.1	0.41	99.9	97.9	-2
24	Al Horford	ATL	14	72.5	13.4	71	0.32	97.9	97.6	-0.3
25	Paul Millsap	ATL	18.7	82.4	16.7	81.3	0.37	98.7	97	-1.7

(CLE).

From the wealth of the technological information age, the problem of computing the exact number of possessions occurred during a game has been already solved. We have the play-by-play logs already available, that could be used to determine how many possessions are used by a player. Furthermore, our model could be extended by incorporating pair or group performance in the model, enabling the analysis of scoring efficiency when a pair or three players are on the court at the same time (we have discussed above the importance of the Big-Threes). A future research should include defensive efficiency in the model to offer a complete individual evaluation, now a key limitation because data referring to those abilities (Goldsberry and Weiss, 2013).

Acknowledgement

Authors acknowledge the careful reading by the referees which with their deep comments and suggestions greatly improved this paper.

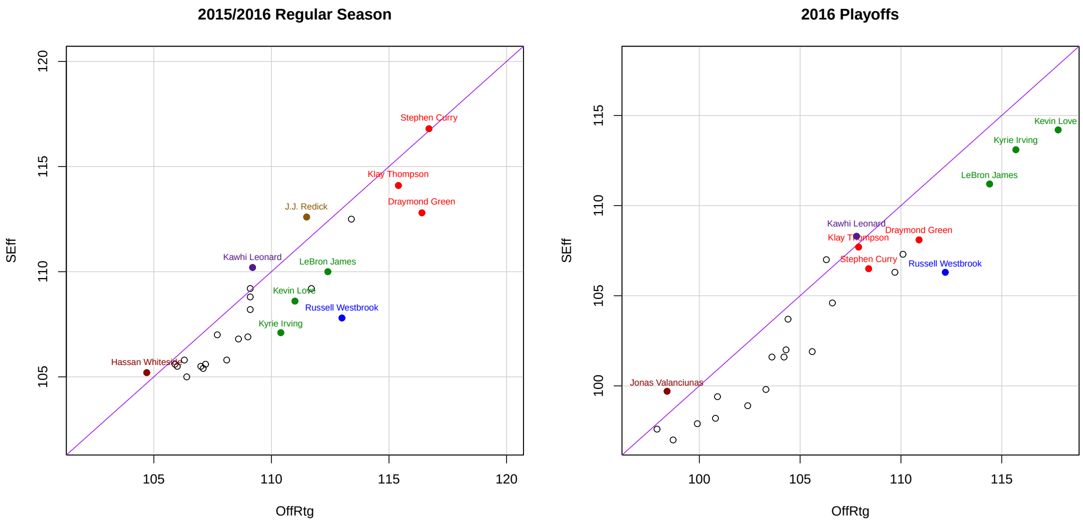

Figure 1: Scatter diagram showing (SEff) behavior versus (OffRtg)

References

Bailey, A. (1945). A generalized theory of credibility. *Proceedings of the Casualty Actuarial Society*, 32: 13–20.

Basketball Reference (2017: Accessed 03-03). Available online at: `http://www.basketball-reference.com`.

Bruce, S. (2016). A scalable framework for NBA player and team comparisons using player tracking data. *Journal of Sports Analytics*, 2(2): 107–119.

Bühlmann, H. (1967). Experience rating and credibility. *Astin Bulletin*, 4: 199–207.

Chen, X., Chen, Y. and Wilbur, K.C. (2013). There's no 'I' in 'team': Estimating NBA players' offensive production. *SSRN Electronic Journal*. Available online at: `http://dx.doi.org/10.2139/ssrn.1861192`.

ESPN.com Insider (2017: Accessed 03-03). Available online at: `http://insider.espn.com/nba/hollinger/statistics`.

Goldsberry, K., and Weiss, E. (2013). The Dwight effect: A new ensemble of interior defense analytics for the NBA. *MIT Sloan Sports Analytics Conference*.

Heilmann, W.R. (1989). Decision theoretic foundations of credibility theory. *Insurance: Mathematics and Economics*, 8: 77–95.

Hollinger, J. (2002). *Pro basketball prospectus: 2002 Edition (Pro basketball forecast)*. Potomac Books.

Kubatko, J., Oliver, D., Pelton, K. and Rosenbaum, D. (2007). A starting point for analyzing basketball statistics. *Journal of Quantitative Analysis in Sports*, 33(3): 1–22.

Lamas, L., Santana, F., Heiner, M., Ugrinowitsch, C. and Fellingham, G. (2015). Modeling the offensive-defensive interaction and resulting outcomes in basketball. *PloS one*, 10: e0144435.

Lee, T.C., Judge, G.G. and Zellner, A. (1968). Maximum likelihood and Bayesian estimation of transition probabilities. *Journal of the American Statistical Association*, 63: 1162–1179.

Mertz, J., Hoover, D.L., Burke, J.M., Bellar, D., Jones, L.M., Leitzelar, B. and Judge, L.W. (2016). Ranking the greatest NBA players: A sport metrics analysis. *International Journal of Performance Analysis in Sport*, 16(3): 737–759.

NBA.com Stats (2017: Accessed 03-03). Available online at: `http://stats.nba.com`.

Oliver, D. (2004). *Basketball on paper: Rules and tools for performance analysis.* Brassey's Inc., first edition.

Parker, R.J. (2010). Modeling basketball's points per possession with application to predicting the outcome of college basketball games. *Unpublished working paper*. Available online at: `http://www.basketballgeek.com/downloads/ryan_bach_essay.draft.pdf`.

Pomeroy, K. (2004). National efficiency. Available online at: `http://kenpom.com/blog/national-efficiency`.

Skinner, B. (2010). The price of anarchy in basketball. *Journal of Quantitative Analysis in Sports*, 6(1): 3–6.

Zimmermann, A. (2016). Basketball predictions in the NCAAB and NBA: Similarities and differences. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 9(5): 350–364.