# Players' importance in basketball and the generalized Shapley value

**Rodolfo Metulini**[1], **Giorgio Gnecco**[2]

December 18th, 2021

CMStatistics 2021

14th International Conference of the ERCIM WG on Computational and Methodological Statistics

[1]Department of Economics and Statistics (DISES) - University of Salerno

[2]AXES Research Unit, IMT - School for Advanced Studies, Lucca

# Introduction

- To help coaches and staff on the strategy to adopt in basketball, the analysis of i) play-by-play and ii) tracking data with Statistics and Machine Learning techniques is now a common practice.
- We evaluate player's importance as his/her average marginal contribution to the utility of an ordered subset of players, through a generalized version of the Shapley value.
- A novel multi-step approach is proposed, which involves the estimation of the generalized characteristic function for each lineup (as the probability that lineup has to win the game).
- A novelty, allowed by explicitly considering single lineups, is represented by the possibility of forming best lineups, based on players' generalized Shapley values conditional on specific constraints (e.g., an injury or a coach's decision).
- We show the application of our method to real National Basketball Association (NBA) data. Generalized Shapley values are computed for Utah Jazz players, along with an application of a greedy algorithm for lineups management.

# Players' (and team) performance

- Literature devoted to analyse team (e.g., **Moreno and Lozano**; **2014**; **Hofler and Payne**; **2006**; **Metulini et al.**; **2018**) and players performance (box-score, e.g., **Cooper et al.**; **2009**; **Fearnhead and Taylor**; **2011**; **Page et al.**; **2013**, shooting variables, e.g., **Piette et al.**; **2013**; **Metulini and LeCarre**; **2020**; **Sandri et al.**; **2020**, synthetic metrics, e.g. **Terner and Franks**; **2020**) is wide and constantly evolving.
- Measures for evaluating single player's contribution to the team:
  - Plus-Minus (PM): points scored by a team minus points scored by its opponent during the time the specific player is on the court:
    1. Regression-based versions accounting for the other players on the court (Adjusted PM, APM) (**Rosebaum**; **2004**),
    2. Box-score PM (BPM): includes other players' statistics among the explanatories and controls for the team strength (e.g., **Kubatko et al.**; **2007**; **Ilardi**; **2007**),
    3. Regularized APM (RAPM): accounts for the presence of multicollinearity in APM (e.g., **Sill**; **2010**; **Engelmann**; **2017**),
    4. Real Plus Minus (RPM): normalizes the measure by the number of offensive and defensive possessions.
- ... Despite recent PM versions move in the direction of i) not just accounting for scoring factors, and ii) solving for multicollinearity, those issues still deserve more attention (**Terner and Franks**; **2021**).

# Players' (and team) performance (cont'd)

- Measures for evaluating single player's contribution to the team (cont'd):
  - Win-Shares (WS): the contribution for team success of its individuals. WS48 (WS per 48 minutes) expresses the WS values in a per-minute basis.
  - Wins Above Replacement (WAR), also referred to WAR Player (WARP, **Pelton**; **2019**), seeks to evaluate a player by comparing the performance of a team made up of him/her and four average players with the performance of a team made up of four average players and one replacement-level player.
  - Value Over Replacement Player (VORP): points per 100 team possessions that a player contributed above a replacement level player (**Myers**; **2020**). Similar to a WARP with Plus-Minus instead of wins.
- ... A player WS score is positively influenced by the amount of time he/she is on the court (**Sarlis and Tjortjis**; **2020**).
- ... WARP and WS48 outperform WS as they are expressed on a per-minute basis. However, likewise WARP, VORP suffers from issues of multicollinearity (**Sarlis and Tjortjis**; **2020**).

# A cooperative game framework to measure player's contribution

- Rarely players' and team are jointly analysed by considering them within a cooperative game where each play represents a *pathway* through which players (and the team) move to the goal (i.e., score the basket, avoid the opponents to score the basket, ...).
- The utilization of the Shapley value has not massively been percolated to team sports (an exception the works on Soccer by **Auer and Hiller**; **2015**; **Hiller**; **2015**) and has never been used in basketball, except for the conference article by **Yan et al.**; **2020**.
- ... Shapley value is a solution concept used to assign a reward to each player in a coalition, according to the marginal contribution he makes to it (its axiomatic characterization is expressed in terms of simple properties - **Michalak et al.**; **2014** - that can be easily transferred to basketball), whereas other industry-standard measures lack such a game-theoretical interpretation.
- ... By construction, our metric, which is based on a generalization of the Shapley value, i) is better motivated than the Shapley value for the case of basketball data analysis, ii) does not suffer for multicollinearity (at least, not directly), iii) considers both scoring and non scoring, offensive and defensive factors.

# A multi-step procedure

Our methodological strategy is the following:

1. A logistic regression model, based on all NBA data at hand, is estimated, where the Outcome (win=1, defeat=0) is regressed on the well-known *four Dean's factors* (**Kubatko et al. 2007**) at game level.
2. First step estimated coefficients are used to derive the probability to win associated with each lineup (i.e., replacing the at-game level Dean's factors with those computed at-lineup level).
3. Based on lineups' probabilities, we compute two versions (unweighted and weighted) of the generalized Shapley value for each player.

# Generalized Shapley value

In cooperative game theory, a "generalized" coalitional game is defined as a pair $(N, v)$ (**Nowak and Radzik**; **1994**), where:

- $N = \{1, 2, ..., n\}$ is the player set (cardinality $= n$) .
- $v$ is the generalized characteristic function, which assigns to every ordered coalition $T$ extracted from the set $N$ a certain worth $v(T)$ reflecting the "abilities" of such an ordered coalition.

For any ordered coalition $T$ and any player $i$, let $T(i)$ denote the ordered (sub)coalition formed by the players that precede $i$ in $T$

The generalized Shapley (or Nowak-Radzik, NR) value of player $i$ is the average of the marginal contribution of that player when he enters an ordered subcoalition (of any cardinality) of an ordered coalition $T$ with cardinality $|T| = n$:

$$\phi_i^{NR}(N, v) = \frac{1}{n!} \sum_{T \in \mathcal{T} \text{ with } |T|=n} \left(v((T(i), i)) - v(T(i))\right) . \quad (1)$$

- This definition differs from the classical one of a coalitional game, whose characteristic function is defined on the set of unordered coalitions (**Maschler et al.**; **2013**, ch. 17).

# Generalized Shapley value (cont'd)

A generalization of the Shapley value to the case of ordered coalitions, which was proposed by **Nowak and Radzik**; **1994** (NR) is more suitable than the Shapley value itself:

- Only five players of each team can play simultaneously.
- This means that players that "virtually" enters a coalition after the fifth player must have zero worth.
- The introduction of the constraint above leads to multiple values for the coalition made of all players in the team (grand coalition), which makes the Shapley value inapplicable (no issues arise instead for the application of the NR generalized Shapley value)

# Estimated generalized Shapley value

- In the case of basketball, we have non-zero worth only on coalitions with cardinality $m = 5$, whereas $n > m$ is the total number of players rotating in the court.
- Let $\mathcal{L}_i$ be the set of observed lineups (ordered coalitions with cardinality$= 5$) in which player $i$ appears. Then, one gets the following estimate of his "empirical" generalized Shapley value:

$$\hat{\phi}_i^{NR}(N, \upsilon_k) = \frac{5}{n} \frac{1}{5|\mathcal{L}_i|} \sum_{L \in \mathcal{L}_i} (\hat{\upsilon}_k(L) - 0) = \frac{1}{n|\mathcal{L}_i|} \sum_{L \in \mathcal{L}_i} \hat{\upsilon}_k(L). \qquad (2)$$

  - $\hat{\upsilon}_k(T(i)) = 0$ because of zero worth when cardinality $< 5$.
  - The inclusion of the factor $\frac{1}{5}$ is needed since, for any specific quintet, each player has the same probability of being the fifth to join all the other members of that quintet.
  - The other factor $\frac{5}{n}$ expresses the probability that player $i$ enters in one of the first 5 positions.

# Generalized characteristic function

- A player marginal utility is computed based on the values assumed by a generalized characteristic function $v(.)$ - that measures the cohesion (performance) of each ordered combination of players with him in the court.
  - **Yan et al.**; **2020** also highlighted the importance of correctly specifying the characteristic function.
- We consider two alternatives for $\hat{v}_k(L)$, denoted respectively by $\hat{v}_1(L)$ and $\hat{v}_2(L)$.

$$\hat{v}_1(L) = P(Win)_{(L)} \tag{3}$$

$$\hat{v}_2(L) = P(Occ)_{(L)} P(Win)_{(L)} \tag{4}$$

- The two corresponding estimated generalized Shapley values are, respectively, the "unweighted generalized Shapley value"($\hat{\phi}_i^{NR}(N, \hat{v}_1)$) and the "weighted generalized Shapley value" ($\hat{\phi}_i^{NR}(N, \hat{v}_2)$).

# Logistic model for estimating P(Win)

Single lineups do not play the full match, thus making not straightforward to determine an estimate for the probability to win for each quintet.

- We adopt a strategy based on estimating logistic regression at game level to obtain $\beta$ coefficients.
- We use the four Dean's factors (**Oliver**; **2004;Kubatko et al.**; **2007**) as explanatory features, which are well-known and agreed in the literature ($R^2 \sim 0.9$)[3]:
  1. effective field goal percentage (eFG%): $\frac{(FG+0.5*3P)}{FGA}$,
  2. turnover percentage (TOV%): $\frac{TOV}{(FGA+0.44*FTA+TOV)}$,
  3. offensive rebound percentage (ORB%): $\frac{ORB}{(ORB+OppDRB)}$,
  4. free throws percentage (FT%): $\frac{FT}{FGA}$.
- Studies using Statistics/Machine Learning techniques to estimate game's outcome (Artificial Neural Networks, Naive Bayes Classifier, Support Vector Machines) generally achieve a 70-80% accuracy.

[3] To account for both teams' features the four Dean's factors are actually eight (we use notation *off* when referring to the considered team, *def* when referring to the opponent team).

# Logistic model for estimating P(Win) (cont'd)

Logistic regression model, to be estimated in step 1, reads, for game $i$, as:

$$log \frac{P(Y_i = 1 \mid \boldsymbol{X_i})}{P(Y_i = 0 \mid \boldsymbol{X_i})} = \boldsymbol{X}_i \beta \tag{5}$$

- The left part of the equation is the log-odds of $Y_i$ conditional to $\boldsymbol{X}$.
- $\boldsymbol{Y}$: the response binary variable representing the outcome of the games, $Y_i \in \{0, 1\}$ , $i = 1, ..., g$, where $g$ is the number of games.
- $\boldsymbol{X}_i$: the $i-$th row of the design matrix $\boldsymbol{X}$ with $g$ rows and $p$ columns ($p$=8, the eight Dean's factors used as explanatory variables, *eFG%_off*, *eFG%_def*, *TOV%_off*, *TOV%_def*, *ORB%_off*, *ORB%_def*, *FT%_off*, *FT%_def*, computed at game level).
- $\beta$: the vector containing the $p$ regression parameters associated with the explanatory variables. These parameters have to be estimated from the data.

# Logistic model for estimating P(Win) (cont'd bis)

- So, in the second step, on dataset $\tilde{\boldsymbol{X}}$, where the eight Dean's factors are expressed at single lineup level[4] we predict the probability to win the game $P(Win)_{L_j}$ on each lineup $L_j$ by using the vector $\hat{\beta}$ estimated in step one:
    - let $\tilde{\boldsymbol{X}}_{\boldsymbol{j}}$ be the $j$-th row of the matrix $\tilde{\boldsymbol{X}}$ with $l$ rows (number of different considered lineups) and $p$=8 columns (the eight Dean's factors computed at lineup level), the probability to win the game for the lineup $L_j$ is:

$$P(Win)_{L_j} = \frac{exp(\tilde{\boldsymbol{X}}_{\boldsymbol{j}}\hat{\beta})}{1 + exp(\tilde{\boldsymbol{X}}_{\boldsymbol{j}}\hat{\beta})}, \ j = 1, ..., l\,. \tag{6}$$

[4] An average of the Dean's factors is taken for the opponent lineups, considering that they are not fixed

# Estimating players marginal utilities

- In the third step the generalized Shapley values are computed for each player, by using Equation 2 along with the winning probabilities provided by Equation 6 computed for each lineup.

- Finally, a greedy algorithm is applied for lineups management:
  1. First, we choose, among $n$ players, the one with the largest generalized Shapley value (or a "preferred" player), which is the "most important player";
  2. Then, we choose the player with the largest generalized Shapley value considering all the other players based on the subset of lineups in which the most important player was in;
  3. repeat, until the first five most important players have been chosen.

# Data

- Play-by-play of all NBA games (both regular seasons and play-offs) for 17 seasons (from 2004–2005 to 2020–2021), available thanks to an agreement with BigDataBall (UK) (`www.bigdataball.com`).
- Start/end of the period, made/missed 2 points shots, made/missed 3 points shots, made/missed free throws, offensive/defensive rebounds, assists, steals, blocks and fouls, for each game and for both teams, associated with the lineup of both the two teams.
- All features for both $\boldsymbol{X}$ and $\tilde{\boldsymbol{X}}$ have been computed from play-by-play dataset.
- We have retrieved computed WS, WS48, BPM and VORP values from Basketball Reference website (`www.basketball-reference.com/`).

# Logistic results

- Since the dependent variable is binary, logistic model is used, instead of linear regression with Ordinary Least Squares (as in **Kubatko et al.**; **2007**) to prevent us from having the estimated probability of success not in the range [0,1] (**Wooldridge**; **2010**).
- In order to let the effect on the Outcome of Dean's factors comparable, we have normalized all the features using a z-score transformation.
- According to the logistic regression results on the full set of games ($g = 21,735$) from all 17 seasons (play-offs included), we use the estimated vector $\hat{\beta}$ to be used to determine $P(Win)$ for each lineup $L_j$ in the second step of our analysis:

$$[10.255, -10.255, -1.850, 1.749, 0.998, -0.998, 0.780, -0.810]'$$

- ROC curve used as a validation confirms the high level of classification accuracy (AUC $= 0.951$). The hit-rate stands to 0.903.
- A 10-fold cross validation supports the previous evidences (average AUC of 0.946)

# Players' generalized Shapley values

- We retrieve the winning probabilities for all relevant lineups of Utah Jazz in 2020/21 regular season (52-20, best record).
- 107 different lineups (11 different players) on the court more than 4 minutes in regular season (up to a total of 426 different lineups from this set of 11 players, and covering 86% of the total time of play).
- We compute the estimate for winning probability for each lineup (according to equation 6), then we determine the (two versions of) generalized Shapley value for each player:

$$UWGS_i = \frac{1}{11|\mathcal{L}_i|} \sum_{L \in \mathcal{L}_i} \hat{v}_1(L), \quad (7)$$

$$WGS_i = \frac{1}{11|\mathcal{L}_i|} \sum_{L \in \mathcal{L}_i} \hat{v}_2(L). \quad (8)$$

.

- Bootstrap confidence intervals are computed (200 reps).
- Players' ranks on different Bootstrap samples show the robustness of our measures to changes in the sample of lineups considered.
- A comparison of our measures with industry-standard counterparts shows a strong linear relation.

# Lineup management

Applications of a greedy algorithm to form best lineups, using WGS.

1. Coach choice of having Rudy Gobert on the court.

| Player | ($1^{st}$ step) | ($2^{nd}$ step) | ($3^{rd}$ step) | ($4^{th}$ step) |
|---|---|---|---|---|
| Donovan Mitchell | 0.0551 | - | - | - |
| Mike Conley | 0.0535 | 0.0821 | - | - |
| Joe Ingles | 0.0383 | 0.0547 | 0.0941 | - |
| Royce O'Neale | 0.0512 | 0.0707 | 0.0924 | 0.1076 |
| Bojan Bogdanović | 0.0485 | 0.0664 | 0.0922 | n.a. |
| Jordan Clarxson | 0.0337 | 0.0231 | n.a. | n.a. |
| George Niang | 0.0398 | n.a. | n.a. | n.a. |
| Miye Oni | 0.0000 | n.a. | n.a. | n.a. |
| Trent Forrest | n.a. | n.a. | n.a. | n.a. |
| Derrick Favors | n.a. | n.a. | n.a. | n.a. |

2. Unavailability of Mike Conley due to injury (he played 51 out of 72 games)

| Player | ($1^{st}$ step) | ($2^{nd}$ step) | ($3^{rd}$ step) | ($4^{th}$ step) |
|---|---|---|---|---|
| Joe Ingles | 0.0341 | - | - | - |
| Bojan Bogdanović | 0.0316 | 0.0487 | - | - |
| George Niang | 0.0204 | 0.0206 | 0.0809 | n.a. |
| Rudy Gobert | 0.0337 | 0.0385 | 0.0529 | 0.0925 |
| Royce O'Neale | 0.0307 | 0.0459 | 0.0579 | 0.0925 |
| Donovan Mitchell | 0.0292 | 0.0440 | 0.0596 | n.a. |
| Jordan Clarxson | 0.0139 | 0.0131 | 0.0258 | n.a. |
| Derrick Favors | 0.0225 | 0.0370 | 0.0580 | n.a. |
| Miye Oni | 0.0004 | 0.0005 | n.a. | n.a. |
| Trent Forrest | 0.0002 | n.a. | n.a. | n.a. |

Conditional WGS of the algorithm steps for each player are reported.
"n.a.": he never played together with the chosen (until that step) player(s).

# Concluding remarks

- We contribute to the literature aimed at measuring player's contribution, by proposing an approach that:
  1. gathers advantages (and avoids disadvantages) of the industry-standard ones;
  2. permits assessing lineups' management;
  3. is based on a game-theoretical approach.
- This work targets to provide managers, coach and the staff with a robust measure of player's marginal utility along with a strategy for the management of the lineup, that can be used by them for:
  1. replacing a player with one with a larger Shapley, in case the current estimate of the winning probability with him on the court is too low;
  2. the choice of a specific lineup, that could be guided by players' generalized Shapley values or by their ranks, conditional on constraints (e.g., the presence/absence of a player in the lineup).
- As further developments:
  1. Employ a version of the generalized Shapley value that:
     i) excludes some coalitions, to account for impossible lineups,
     ii) permits the analysis of player's contribution at single game level.
  2. Assess the impact of additional features, e.g., ball's position, retrieved by the use of computer vision and machine learning (**Giuffrida et al.**; **2019**), to model the generalized characteristic function.

Auer, B. R., & Hiller, T., On the evaluation of soccer players: a comparison of a new game-theoretical approach to classic performance measures. Applied Economics Letters, 22(14), 1100-1107 (2015)

Cooper, W. W., Ruiz, J. L., & Sirvent, I., Selecting non-zero weights to evaluate effectiveness of basketball players with DEA, European Journal of Operational Research, 195, 563–574 (2009)

Engelmann, J., Possession-based player performance analysis in basketball (adjusted+/–and related concepts). In Handbook of statistical methods and analyses in sports (pp. 231-244). Chapman and Hall/CRC (2017)

Fearnhead, P., & Taylor, B. M., On estimating the ability of NBA players. Journal of Quantitative Analysis in Sports, 7(3), article number 11 (2011)

Giuffrida, D., Benetti, G., De Martini, D., & Facchinetti, T., Fall detection with supervised machine learning using wearable sensors. In 2019 IEEE 17th International Conference on Industrial Informatics (INDIN) (Vol. 1, pp. 253-259). IEEE (2019)

Hiller, T., The importance of players in teams of the German Bundesliga in the season 2012/2013–a cooperative game theory approach. Applied Economics Letters, 22(4), 324-329 (2015)

Hofler, R. A., & Payne, J. E., Efficiency in the National Basketball Association: a stochastic frontier approach with panel data. Managerial and Decision Economics, 27(4), 279-285 (2006)

Kubatko, J., Oliver, D., Pelton, K., & Rosenbaum, D. T., A starting point for analyzing basketball statistics. Journal of Quantitative Analysis in Sports, 3(3), article number 1 (2007)

Ilardi, S., Adjusted plus-minus:an idea whose time has come. 82games.com (2007)

Maschler, M., Solan, E., & Zamir, S., Game Theory. Cambridge University Press (2013)

Metulini, R., Manisera, M., & Zuccolotto, P., 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 (2018)

Metulini, R., & Le Carre, M., Measuring sport performances under pressure by classification trees with application to basketball shooting. Journal of Applied Statistics, 47(12), 2120-2135 (2020)

Michalak, T. P., Szczepánski, P. L., Rahwan, T., Chrobak, A., Brânzei, S., Wooldridge, M., & Jennings, N. R., Implementation and computation of a value for generalized characteristic function games. ACM Transactions on Economics and Computation, 2(4), article number 16 (2014)

Moreno, P., & Lozano, S., A network DEA assessment of team efficiency in the NBA. Annals of Operations Research, 214(1), 99-124 (2014)

Myers D. 2020. About box plus/minus (BPM). https://www.basketball-reference.com/about/bpm2.html

Nowak, A., & Radzik, T., The Shapley Value for n-person games in generalized characteristic function form. Games and Economic Behavior 6(1), 150–161 (1994)

Oliver, D., Basketball on paper: rules and tools for performance analysis. Potomac Books, Inc. (2004)

Page, G. L., Barney, B. J., & McGuire, A. T., Effect of position, usage rate, and per game minutes played on NBA player production curves. Journal of Quantitative Analysis in Sports, 9(4), 337-345 (2013)

Pelton K. 2019. The WARP rating system explained. http://sonicscentral.com/warp.html.

Piette, J., Anand, S., & Zhang, K., Scoring and shooting abilities of NBA players. Journal of Quantitative Analysis in Sports, 6(1), article number 1 (2013)

Rosenbaum, D., Measuring how NBA players help their teams win. 82Games.com (http://www.82games.com/comm30.htm) (2004)

Sandri, M., Zuccolotto, P., & Manisera, M., Markov switching modelling of shooting performance variability and teammate interactions in basketball. Journal of the Royal Statistical Society: Series C (Applied Statistics), 69(5), 1337-1356 (2020)

Sarlis, V., & Tjortjis, C., Sports analytics–Evaluation of basketball players and team performance. Information Systems, 93, article number 101562 (2020)

Sill, J., Improved NBA adjusted+/-using regularization and out-of-sample testing. In Proceedings of the 2010 MIT Sloan Sports Analytics Conference (2010)

Terner, Z., & Franks, A., Modeling player and team performance in basketball. Annual Review of Statistics and Its Application, 8 (2021)

Wooldridge, J. M., Econometric analysis of cross section and panel data. MIT press (2010)

Yan, T., Kroer, C., & Peysakhovich, A., Evaluating and rewarding teamwork using cooperative game abstractions. Advances in Neural Information Processing Systems, 33 (2020)

# Logistic regression with Maximum Likelihood

| | *Dependent variable:* |
|---|---|
| | outcome (win: 1, defeat: 0) |
| eFG%_Off | 10.255*** |
| | (0.147) |
| eFG%_Def | -10.225*** |
| | (0.146) |
| TOV%_Off | -1.850*** |
| | (0.038) |
| TOV%_Def | 1.749*** |
| | (0.037) |
| ORB%_Off | 0.998*** |
| | (0.026) |
| ORB%_Def | -0.998*** |
| | (0.026) |
| FT%_Off | 0.780*** |
| | (0.029) |
| FT%_Def | -0.810*** |
| | (0.029) |
| Constant | 0.063*** |
| | (0.023) |
| Observations | 21,735 |
| Log Likelihood | -6,304.493 |
| Akaike Inf. Crit. | 12,626.990 |
| McFadden pseudo $R^2$ | 0.581 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 |

Go back to slide

# Validation

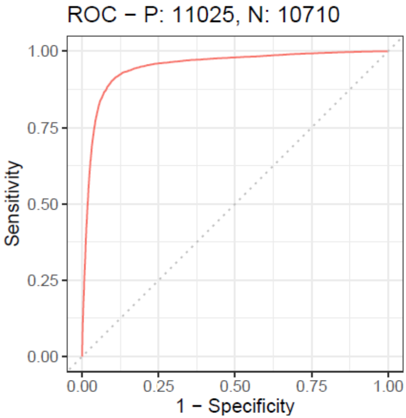


Figure: Receiving Operation Characteristic curve computed from the full sample of 21,735 games: 11,025 positives (outcome=1) and 10,710 negatives (outcome=0).

Go back to slide

# Shapley values

| Player (i) | n lineups | %time | UWGS (rank) | WGS*100 (rank) |
|---|---|---|---|---|
| Royce O'Neale | 78 | 77.6 | 0.0446 (4) | 0.0380 (5) |
| Bojan Bogdanović | 73 | 74.0 | 0.0439 (6) | 0.0382 (4) |
| Rudy Gobert | 71 | 74.2 | 0.0487 (2) | 0.0454 (2) |
| Donovan Mitchell | 67 | 66.5 | 0.0445 (5) | 0.0389 (3) |
| Joe Ingles | 65 | 54.3 | 0.0452 (3) | 0.0359 (7) |
| Jordan Clarxson | 61 | 44.1 | 0.0360 (8) | 0.0242 (8) |
| Mike Conley | 49 | 54.1 | 0.0504 (1) | 0.0510 (1) |
| Derrick Favors | 36 | 25.7 | 0.0324 (9) | 0.0200 (9) |
| George Niang | 30 | 24.7 | 0.0413 (7) | 0.0368 (6) |
| Miye Oni | 4 | 3.5 | 0.0005 (10) | 0.0004 (10) |
| Trent Forrest | 1 | 1.0 | 0.0002 (11) | 0.0002 (11) |

Table: Generalized Shapley values for the 11 selected players of the Utah Jazz team in the 2020/21 regular season (with rank in brackets). *n lineups* is the number of different lineups where that player was in; %*time* is the percentage of time that player was on the court, with respect to the time played by all the 107 considered lineups; the expressions of the two generalized Shapley values are detailed in Equations (7) and (8).

Go back to slide

# Bootstrap confidence intervals

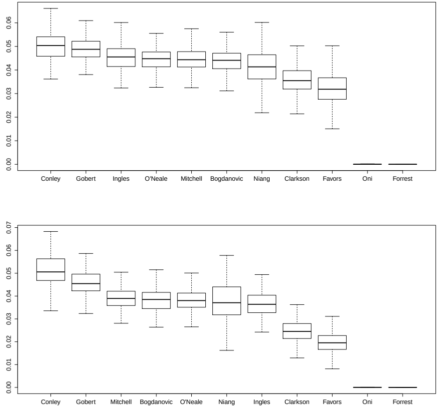


Figure: Box plots with $1^{st}$, median and $3^{rd}$ quartiles (boxes) and the 99% confidence intervals (wiskers) from $n_r = 200$ bootstrap samples, for the 11 Utah Jazz players during season 2020/21. UWGS (top chart), WGS (bottom chart).

Go back to slide

# Players' ranks

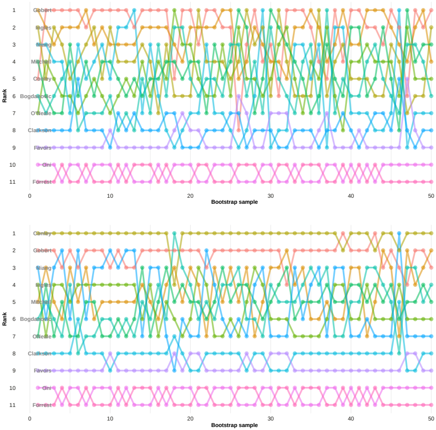


Figure: Bump chart reporting the rank of Utah Jazz players according to UWGS (top) and to WGS (bottom) and the bootstrap samples (first 50, for clarity). Season 2020/21.

Go back to slide

# Comparison with industry-standard measures

| Pearson | | | | | | |
|---|---|---|---|---|---|---|
| | WS | WS48 | BPM | VORP | UWGS | WGS |
| WS | 1.00 | .786 | .851 | .929 | .822 | .850 |
| WS48 | | 1.00 | .841 | .750 | .627 | .710 |
| BPM | | | 1.00 | .943 | .794 | .833 |
| VORP | | | | 1.00 | .751 | .784 |
| UWGS | | | | | 1.00 | .968 |
| WGS | | | | | | 1.00 |

| Kendall's Tau | | | | | | |
|---|---|---|---|---|---|---|
| | WS | WS48 | BPM | VORP | UWGS | WGS |
| WS | 1.00 | .709 | .709 | .836 | .855 | .708 |
| WS48 | | 1.00 | .709 | .655 | .709 | .636 |
| BPM | | | 1.00 | .764 | .709 | .636 |
| VORP | | | | 1.00 | .691 | .582 |
| UWGS | | | | | 1.00 | .782 |
| WGS | | | | | | 1.00 |

Table: Pearson correlation (top) and Tau-Kendall rank correlation among the two generalized Shapley values and the 4 industry-standard measures adopted for players' contributions. Utah Jazz players. Season 2020/21.

Go back to slide